The Effects of Internal and External Factors on the Non-standard Usages of English by the Emerging Workforce of Central Switzerland: a spoken corpus study

By Susanne Oswald-Trefz

A thesis submitted to

The University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Department of English Language and Linguistics

The University of Birmingham

June 2021

UNIVERSITY OF BIRMINGHAM

Abstract

This thesis explores the spoken output of 18-20 year old Central Swiss residents by means of a learner corpus compiled by conducting interviews. Descriptive statistical analysis was used to investigate the effect of demographic, identity-related and educational variables on non-standard usages (NSU). The collection of extensive participant metadata made the variable analysis possible. Education policy changes concerning English as foreign language instruction in primary school facilitated the evaluation of early and late starters and their use of twenty-five NSU. Further, an online survey was conducted to measure the acceptability of spoken output containing NSU features. It has been concluded that the increased exposure to multiple languages in home environments can have a positive effect on the accuracy of English use. Internal adverse associations appeared to negatively affect the amount of output and accuracy of the participants' spoken English. In addition, educational paths and proficiency levels were found to correlate with accuracy, however, length of study or length of stays abroad did not. With few exceptions, the early starters outperformed the late starters in terms of accuracy. The survey revealed that there is a broad acceptance of NSU with typical first language interference features and a surprising fifty percent acceptance rate of zero third person singular use in everyday spoken situations.

Acknowledgements

I would like to dedicate this thesis to my loving family who have supported me throughout the entire process and my parents who not only allowed but encouraged me to explore.

Further, I would like to thank Dr Paul Thompson for his exceptional expertise and guidance throughout the doctoral process.

Thank you to SHL for allowing me to pursue my academic interests in an encouraging environment.

Table of Contents

Figures

Tables

List of Abbreviations and Terms

CEFR - Common European Framework of Reference for Languages

CLIL - Content and Language Integrated Learning

EDK - Swiss Conference of Cantonal Ministers of Education

L1 - First language and mother tongue used interchangeably

L2 - Second language

NSU - Non-standard usage

phw - Per hundred words

SD - Standard deviation

SLA - Second language acquisition

CIA - Contrastive interlanguage analysis

CEA - Computer-aided error analysis

Corpora

BASE - British Academic Spoken English Corpus

BNC - British National Corpus

BNC2014 - Spoken British National Corpus 2014

CANCODE - Cambridge and Nottingham Corpus of Discourse in English

COLT - Bergen corpus of London teenage language

CSC - Central Swiss Corpus

ELFA - English as a Lingua Franca in Academic Settings

ICCI - International Corpus of Crosslinguistic Interlanguage

ICLE - International Corpus of Learner English

LINDSEI - Louvain International Database of Spoken English Interlanguage

LINDSEI-CH - Louvain International Database of Spoken English Interlanguage Chinese

LINDSEI-CZ - Louvain International Database of Spoken English Interlanguage Czech

LINDSEI-GE - Louvain International Database of Spoken English Interlanguage German

LINDSEI-JP - Louvain International Database of Spoken English Interlanguage Japanese

LINDSEI-PL - Louvain International Database of Spoken English Interlanguage Polish

LINDSEI-TR - Louvain International Database of Spoken English Interlanguage Turkish

LOCNEC - The Louvain Corpus of Native English Conversation

LOCNESS - Louvain Corpus of Native English Essays

MICASE - Michigan Corpus of Academic Spoken English

MLE - Havering Multicultural London English Corpus

Swiss English Database

TTC - Toronto Teen Corpus

VOICE - Vienna-Oxford International Corpus of English

Computer programmes used

Microsoft

SPSS 25

AntConc

Soundscriber

Dragon

Potplay

Terms

Foreign – this study was conducted in Switzerland and the term foreign refers to non-Swiss in origin.

Native language – the term was used instead of first language on the consent form and participator profile so as not to confuse participants who speak multiple languages when the first language they learned might not be what they now consider their native language.

Non-standard usage – the term non-standard usage (NSU) was used to label English utterances which did not conform to the English standards used in Switzerland.

Numerus clausus – numerus clausus means 'closed number' in Latin and refers to university study restrictions. The political decision to apply a numerus clausus is made for medical studies in Switzerland when the number of applications exceeds the capacity.

Matura - A Matura is a Swiss exam which enables students to enter Swiss universities. It is equivalent to British A-levels or an International baccalaureate.

Multilingualism – the term is used interchangeably with plurilingualism with the latter being preferred in Swiss literature.

Footnotes are occasionally used throughout this paper to give internet references and make additional comments which would interrupt the readability of the text.

Chapter 1 **INTRODUCTION**

1.1   Introduction

English has taken its place in multilingual Switzerland. In this land of linguistic complexity, many factors influence foreign language acquisition and use. In addition to the most obvious factors, global and economic pressures, the introduction of English in primary school education has not only challenged long-standing nationalist values but has influenced the spoken English of its youth. The investigation of that spoken English is the basis for this doctoral study.

Vocational education is a key factor in the dual education system in Switzerland where up to seventy percent of the 18–20-year-olds enter the workforce as certified skilled or semi-skilled workers. Because of this unique educational system and the nation's linguistic diversity, I believe that this emerging workforce has the potential to influence current and future English use in Switzerland and have therefore compiled a spoken corpus to allow the investigation of internal and external factors which affect the spoken English of this age group.

1.2   Motivation

In interactions with Swiss youth through my capacity as a teacher of English to speakers of other languages, my curiosity was heightened to find out which factors were influential in shaping the learner's spoken output. At the same time, I became familiarised with the multifaceted Swiss school system through my three children when each chose a different scholastic route within the system. When I discovered that my youngest child would be in the first class of students to have English in the third grade, I became interested in understanding how this would influence the spoken English of her generation and could directly observe the differences between her teacher's approach to teaching English opposed to the teachers of her

two older brothers. To date, there has been little agreement on whether the popular notion of 'the earlier the better' applies to second language (L2) language learning or if early English learners will be 'indistinguishable' to later learners as Singleton (2018, p. 56) suggests. To gain a more thorough understanding, the following section elaborates on the Swiss system in which this study was conducted.

1.3   Background: the Swiss context

1.3.1   Swiss multilingual ideology

Switzerland has an interestingly rich and complex linguistic environment. Since the formation of the Swiss confederation in 1291, when three separate rural communities joined alliance, the Swiss have been refining linguistic and democratic coexistence. Gradually further communities united until Switzerland took its final shape in 1815 when the boundaries, which hold the current 26 cantons and four languages, were defined. At the time the Swiss Federal Constitution of 1848 was ratified, Herder's (1794) beliefs of one language, one culture, one nation were widespread throughout Europe (Demont-Heinrich, 2005). To justify being different, Grin (1998 p. 2) suggests that Switzerland 'had to legitimise its existence' by embracing its multilingualism and creating a 'Swiss national myth' to reinforce the uniqueness of linguistic diversity. Since then, the Swiss have cherished the notion that their democracy is built on the idea or myth that multilingualism and the cooperation and tolerance for others is the very essence of being Swiss.

Not all acknowledge this idea of a multilingual society where tolerance and respect are prominent. Rosenberger (2009) offers another viewpoint, suggesting, 'the myth of linguistic harmony may [thus] be said to cover up a rather unpleasant reality of ignorance and indifference…' (Rosenberger, 2009, p. 112). Indifference is similarly recognised by Dürmüller (1994) in the political arena when he notes that some politicians have the view that cultural

alliance may be stronger with neighbouring countries who speak the same language than with Swiss compatriots who speak a different language (Dürmüller, 1994, p. 47). If indifference or harmony prevail in a Swiss person's perception of their multilingual coexistence, could be as individual as their own identity. Schoch (2000) gives an insightful perspective of the Swiss acknowledgement of linguistic diversity within its society.

> …the cultivation of historically generated cantonal particularisms in the face of national trends towards standardization compelled the Swiss very early on to seek compromises and maintain a liberal and tolerant stance. One dimension of this is the recurrent and challenging experience, known to every Swiss child, of realizing that there are people who, though linguistically different, are unquestionably also part of its own world, and must therefore be respected (Schoch, 2000, p. VI).

1.3.2   Language situation

Switzerland is a 41,285 km^2 landlocked country consisting of three language regions and four national languages. Within each of these regions, variations unique to Switzerland are found. Although bordered by three countries with the same languages, except for a short period, Switzerland has never politically belonged to the bordering countries (Grin, 1998, p. 2). The largest language region is the German speaking area where 62.6 % of the population resides. The second largest language region is French at 22.9%, which is followed by Italian at 8.2%. Romansh is the fourth national language and is spoken by 0.5% of the population. It is isolated in five areas within the canton of Grison. Furthermore, 25% of the population claim a language other than one of the four national languages as one of their main languages[1]. Although the right to speak one's own language is embedded in the Swiss Federal Constitution of 1999, Romansh is considered an official language only in the canton of Grisons. (Bundesverfassung

---

[1] (www.bfs.admin.ch accessed December 30, 2019)

der Schweizerischen Eidgenossenschaft from April 18th, 1999). Figure 1 below shows the geographical distribution of languages.

Figure 1-1 Geographical distribution of the languages of Switzerland (2000)[2]

Within the German-speaking region, roughly 85% of the residents speak a Swiss German dialect (Lüdi and Werlen, 2005, p. 37). Swiss German is used in all aspects of daily life with variance mainly accounted to demographic location, not social hierarchy. Standard German is used for most written communication and in formal spoken settings including education. Weinreich (2011) contributes the success of Swiss German to the 'Lack of dependence on the German social pyramid' as well as Swiss independence and 'self-sufficiency of social values' (Weinreich, 2011, pp. 94-95). Although Swiss German is a spoken language with no orthographic standards, the evolution of the digital age of text messages and emails has brought about an increase in the use of Swiss German in written form (Petkova, 2012, p. 5).

---

[2] source www.bfs.admin.ch

Additionally, use of Swiss German has increased in the media with weather forecasts, many television programmes and even publicly broadcast political discussions being conducted in dialect (Siebenhaar, 2006, p. 482).

This situation in the German speaking area represents a diglossia where a language community uses two different language varieties each with distinctive functions. Speakers of the four national languages tend to view this diglossia in different ways. On one end of the spectrum, many Swiss German speakers view it as a form of bilingualism where Standard German is considered another language (Brohy, 2005, p. 135). Standard German is not learned naturalistically but must be learned in school similar to a foreign language. Almost 80% of those questioned in a small-scale study by Hägi and Scharloth (2005, p. 4) viewed Standard German as a foreign language.

At the other end of the spectrum, Ribeaud (2013), a French speaking Swiss journalist, argues that Swiss German is not even a national language. Ribeaud suggests that although Swiss German is considered a mother tongue, its widespread use in the public domain excludes all others not belonging to the language community. He further suggests that Standard German should replace Swiss German in communication and media. While this viewpoint may reflect French speakers' general dislike of non-standard language, Swiss German speakers place great value on dialect and indeed this value supersedes that of Standard German in what Watts calls the 'ideology of dialect' (Watts, 1999, p. 67). Whereas Swiss German speakers generally comprehend the numerous regional dialect variations, understanding Swiss German can be a daunting task for everyone else. Even the neighbouring German residents have great difficulty understanding Swiss German dialects.

It could be argued that the dominance of Swiss German can interfere with or even hinder the flow of communication between the language regions. Swiss German speakers outnumber their compatriots by seven to three and therefore arguably represent the most powerful linguistic group in Switzerland. Although Standard German is learned in schools as a main or foreign language throughout Switzerland, Swiss German speakers show reluctance in speaking this form when interacting with their compatriots (Siebenhaar, 2006, p. 483). This situation is complicated by the broad use of Swiss German as a spoken medium in public and private affairs. The linguistic and cultural differences between the regions are considered a known but tolerated difference of opinion.

In 2002 at the brink of the internet becoming relied upon as a channel of communication, Dürmüller already observed that 'People in Switzerland are now often more familiar with English than with the languages of their compatriots' (Dürmüller, 2002, p. 116). To investigate this hypothesis that English was increasingly being used as a means of communication between Swiss nationals with different mother tongues Oswald (2014) conducted a questionnaire with 383 participants in Central Switzerland. The results confirmed Dürmüller's observations with 70% of participants claiming to speak English on a regular basis and over 50% said they used English alone or in combination with nation languages to communicate with their compatriots. Further increases in the use of English seem imminent because of earlier introduction of English in primary school, expanding Content and Language Integrated Learning (CLIL) programmes in vocational schools and the extended use of English in higher education.

A reluctance to speak each other's language might revolve around the sheer absence of necessity. With the exception of a handful of bilingual communes, communal governments conduct all business in one of the national languages (Grin and Korth, 2005, p. 69). Thus, local

governments are monolingual units. Each of the twenty-six cantons allocate varying degrees of independence to the communes and therefore not only cantonal but also communal differences are to be expected. The cantons are also primarily monolingual with three cantons bilingual and one trilingual (Grin and Korth, 2005, p. 69). In addition to governing communal responsibilities, 'Each canton has its own constitution, and its own parliament, government and courts.' (Information services of the Federal Chancellery, 2014, p. 14). A consequence of the communes and cantons having autonomy in many language and educational decisions is that there is not a federal minister of education. Instead, there is a council which coordinates education policy and emphasises certain key values. This autonomy allows local governments to make policy decisions which are tailored to the local population. Of the 8.6 million Swiss residents, 25% are of foreign nationality which adds to the language complexity in Swiss schools. The following Figure illustrates the municipalities coloured according to the proportion of foreigners in 2016. As can be seen, cities and highly populated areas tend to have more foreigners than rural areas. Central Switzerland, the geographical area studied, has been indicated within the red circle.
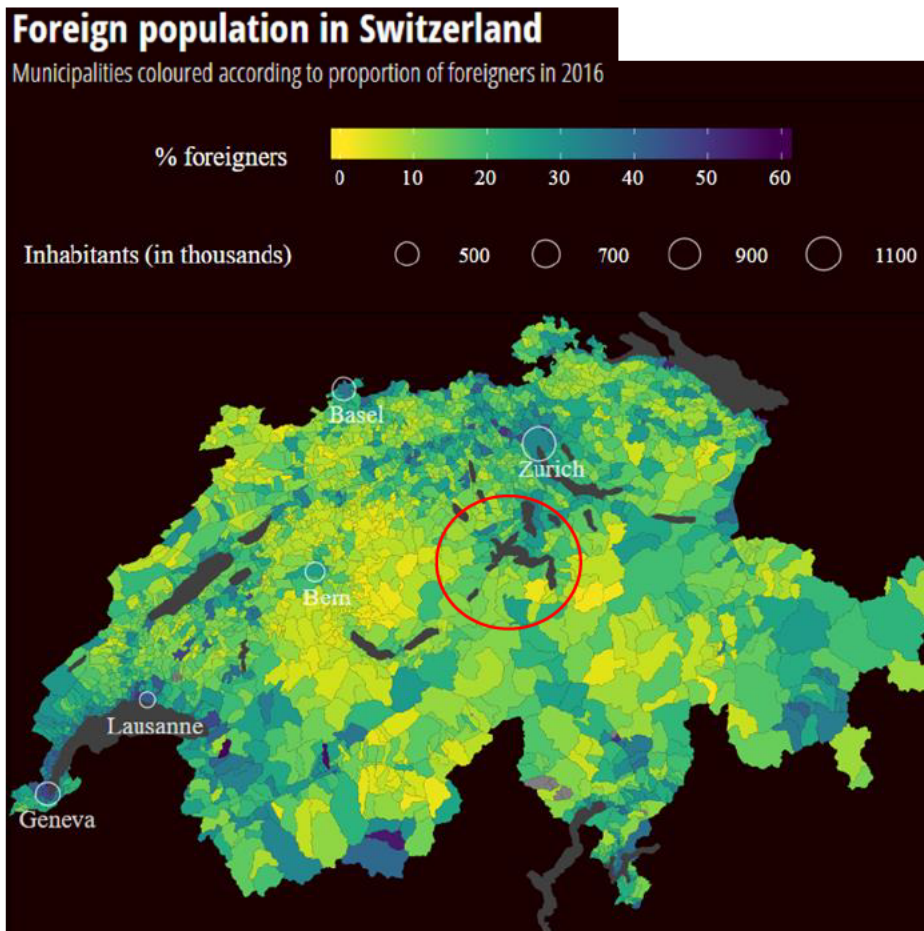
Figure 1-2 Foreign population in Switzerland percentage of foreigners 2016[3]

This map is key in understanding the linguistic diversity within Switzerland because the foreign population of an area affects many municipal decisions, especially on education. It also affects the constitution of learners in foreign language classrooms and within the 2,202 unique municipalities of Switzerland, decisions on education are made at the local level. This study examines one small area; however, it represents a cross section of Swiss diversity, and its results can be used to make further hypotheses which relate to similarly constituted regions.

Whereas many countries have traditionally gathered statistical information about language use based on ethnic heritage or white/non-white status, in Switzerland, the practice of classifying

---

[3] Source with interactive map: https://www.swissinfo.ch/eng/society/migration-series-part-1-_who-are-the-25-foreign-population-in-switzerland/42412156

residents as being either Swiss or foreign nationals for statistical data collection and research is more widespread. It is a reoccurring and prominent factor in numerous research projects on language acquisition in Swiss classrooms (see Gnos, 2012; Haenni-Hoti et al., 2011).

1.3.3 Education system and current language policies

This research is concerned with recent changes in foreign language policy as an external influence on spoken English. An overview of the educational system is, therefore, necessary to understand the role of foreign language policy. The educational system is characterized by diversity. Language policies vary at cantonal and communal levels due to their autonomy. Language education also varies depending on the type of school and the students' academic ability.

All students are given equal education until the fourth to sixth grade; whereupon scholastic achievement decides who has the opportunity to continue directly onto a school which prepares students for tertiary education. On completion of lower-secondary school at the age of sixteen, once again, if one has met the requirements, an upper-secondary or Matura school is an option. Once students receive a Matura, all university courses of study are open to them except medicine, which has a *numerus clausus*. The majority of the remaining students find employment as apprentices.

The educational statistics for 20-year-olds (Bundesamt für Statistik, 2010) recorded that after lower-secondary school the majority, 69%, chose a profession and received on the job training as an apprentice for three to four years while attending vocational school 1-2 days a week. Another 20% attended upper-secondary or Matura school to gain entrance into a tertiary level school and pursue a corresponding career. The remaining 11% had no further qualifications

beyond the obligatory nine years of education. In 2010 an estimated 80% of Swiss youth entered the workforce at the age of 18-20. Since 2010, numerous educational reforms have come into effect. This combined with a general shift towards tertiary education has reduced the number of students perusing an apprenticeship and nearly doubled the number of students following a general education path which leads to tertiary education. As a result, it is currently estimated that 65-70% of Swiss youth between the age of 18-20 will enter the workforce in 2021. It is this sector of society that this study is concerned with and the English that this emerging workforce speaks. Figure 1-3 below gives a simplified illustration of the educational system.

THE SWISS EDUCATION SYSTEM

CONTINUING EDUCATION AND TRAINING (CET)

TERTIARY

PhD Doctorate

Master (5 years)

Bachelor (3 years)

College of Higher Education Diploma

Advanced Federal Diploma of Higher Education

Federal Diploma of Higher Education

8
7
6
5
4
3
2
1

UNIVERSITIES INCL. FEDERAL INSTITUTES OF TECHNOLOGY

UNIVERSITIES OF TEACHER EDUCATION

UNIVERSITIES OF APPLIED SCIENCES

COLLEGES OF HIGHER EDUCATION

PREPARATION FOR FEDERAL DIPLOMA AND ADVANCED FEDERAL DIPLOMA OF HIGHER EDUCATION[4]

Baccalaureate

Specialised Baccalaureate + Specialised School Certificate

Specialised School Certificate

Federal Vocational Baccalaureate + Federal VET Diploma

Federal VET Diploma

Federal VET Certificate

Transition 2: Federal Vocational Baccalaureate or Specialised Baccalaureate → Universities (additional exam)

Transition 1: Baccalaureate → Universities of applied sciences (practical training)

UPPER SEC.

4
3
2
1

BACCALAUREATE SCHOOLS

SPECIALISED BACCALAUREATE

UPPER SECONDARY SPECIALISED SCHOOLS

FVB 1

FVB 2

FVB 2

VOCATIONAL EDUCATION AND TRAINING (APPRENTICESHIP)[3]
With Federal Vocational Baccalaureate (FVB) as an option: FVB 1 during the apprenticeship, FVB 2 after the apprenticeship

VOCATIONAL EDUCATION AND TRAINING[3]

BRIDGE YEAR COURSES

LOWER SEC.

11
10
9

LOWER SECONDARY LEVEL[2]
Performance-based grouping

PRIMARY

8
7
6
5
4
3
2
1

PRIMARY LEVEL[1]
Including kindergarten or first learning cycle

First learning cycle

Kindergarten

Compulsory education

Children and adolescents with special educational needs
Support through special measures from birth up to age 20

The responsibility for the precise arrangement of special needs education during compulsory education (special schools, special classes, special support integrated in regular classes, etc.) lies with the cantons and is not shown on the chart.

© EDK CDIP CDEP CDPE, August 2019

ISCED | International Standard Classification of Education 2011
- ISCED 8
- ISCED 7
- ISCED 6
- ISCED 4
- ISCED 34 + 35
- ISCED 2
- ISCED 1
- ISCED 020

[1] Two years of kindergarten or the first two years of a first learning cycle: included in compulsory education in the majority of cantons

[2] Lower secondary level: 4-year scuola media in the Canton of Ticino (pursuant to exception clause in Art. 6 HarmoS Agreement). Transition to Baccalaureate Schools possible after 10th school year

[3] Vocational education and training (apprenticeship): training in a host company + instruction at a VET school + branch courses. For certain professions, a VET diploma can be obtained in a full-time school programme (e.g. in trade or IT schools).

[4] Federal examination / Federal diploma of higher education = ISCED 6
Advanced federal examination / Advanced federal diploma of higher education = ISCED 7

Figure 1-3 The Swiss educational system[4]

11

Foreign language instruction has long played an important role in Swiss education. Traditionally, instruction of another national language began in secondary school, although not comprehensively throughout Switzerland. In 1975, learning at least one Swiss foreign language commencing in the fourth or fifth grade became an educational objective for all students (Hega, 2001, p. 217). Subsequent recommendations by The Swiss Conference of Cantonal Ministers of Education (EDK) have included learning English as an educational objective for all students. In 1998, the EDK published a 'Gesamtsprachenkonzept' (comprehensive language policy) with the primary goal to promote multilingualism. The EDK recommended that all students learn at least one other national language and English, beginning with the first language in the fifth grade in primary school and the second in the seventh grade in secondary school (Baumgartner, 2012, p. 9). Moreover, binding learning objectives were introduced with language competency references corresponding with the Common European Framework of Reference for Languages (CEFR). The new comprehensive policy neglected to recommend which language should be taught first and as a result opened a heated discussion which has strained tolerance between language regions. In short, some German speaking cantons preferred teaching English before French, whereas the French speaking cantons preferred teaching German before English. The repercussions of this loyalty issue of choosing to teach a non-national language (English) prior to a national language can still be heard in political debates concerning educational policy.

The goals of multilingualism have since been broadened. In 2004 the EDK published new recommendations which led to a national referendum in the attempt to harmonise obligatory education under a concordat called *HarmoS*. Within the new agreement, one national foreign language and English are both taught in primary school until the end of obligatory school; the first foreign language beginning in the third grade and the second foreign language in the fifth

---

4 Source: http://www.edk.ch accessed 08.09.2019

grade. Additionally, a second national foreign language is offered in secondary school (EDK, 2019).

Further developments include the project Lehrplan 21 which was conceived between 2010-2014 and is the first joint curriculum for the primary schools in the German speaking regions. It was completed under the guidance of the Swiss-German Conference of Cantonal Ministers of Education (D-EDK) and approved in 2014. However, each of the twenty-one participating cantons retains legal standing to make its own decisions on various aspects. These changes have also affected the English language classrooms in the form of new methodologies, didactics and learning objectives (Lehrplan 21, 2021).

Figures 1-4 and 1-5 below show the geographical areas and the cantons' decisions on which order the languages are instructed. As can be seen, the entirety of the French speaking region teaches German first, whereas only the German speaking cantons which border the French speaking region have decided to teach French first. Interestingly, the canton of Ticino decided to teach French as the first foreign language in the third grade, German as the second language in the seventh grade and English as a third language in the eighth grade. The canton of Grisons unanimously decided to teach English as the second foreign language in the fifth grade, as can be seen in Figure 1-4. However, each of the communities has autonomy in choosing which languages are taught when. Thus, throughout the canton, depending on the community, German, Italian or Romansh is taught as a first foreign language (Gross, 2017, p. 20).

First Obligatory Foreign Language
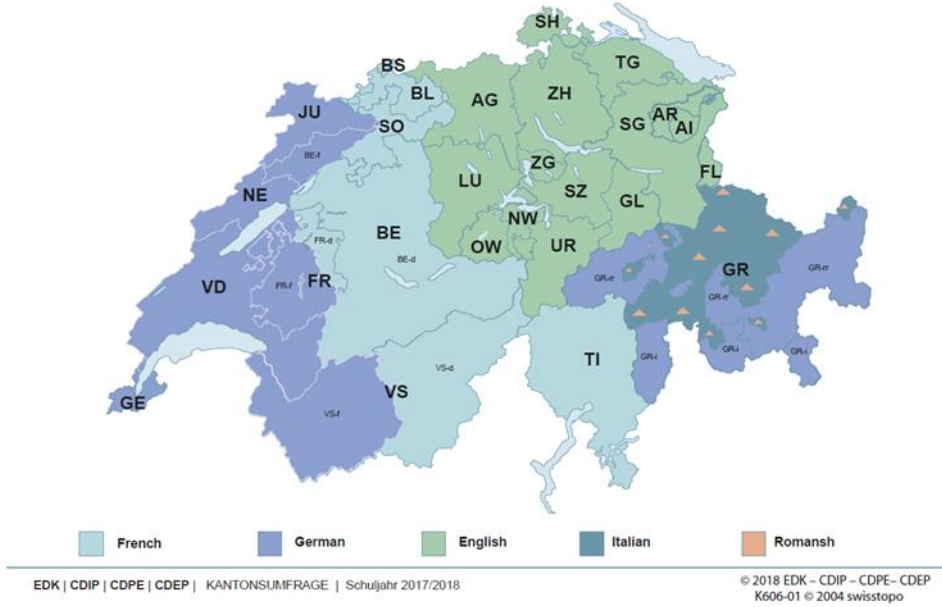
Figure 1-4 First obligatory foreign language school year 2017/2018[5]

Second Obligatory Foreign Language
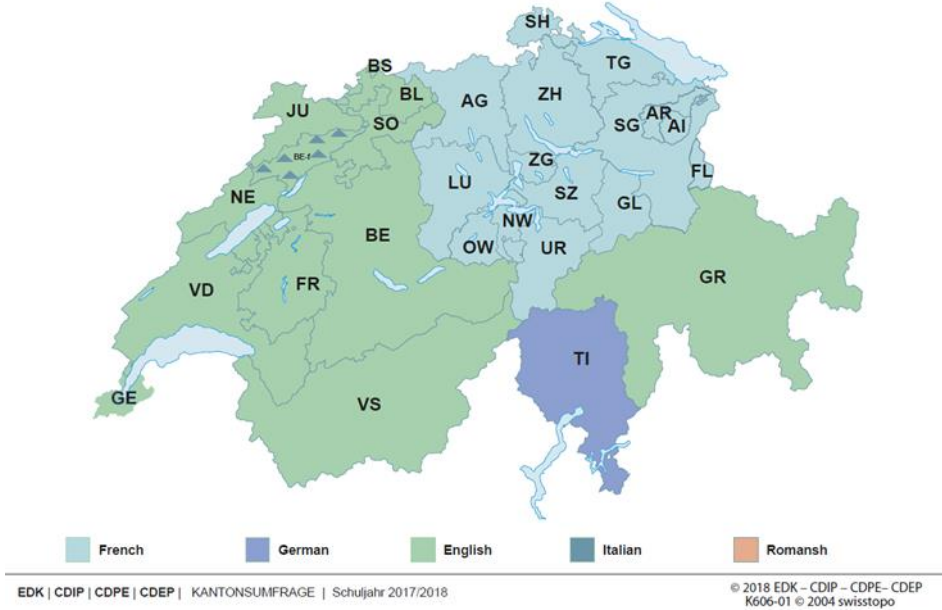
Figure 1-5 Second obligatory foreign language school year 2017/2018[6]

---

[5] Source: http://www.edk.ch accessed 01.15.2019
[6] Source: http://www.edk.ch accessed 01.15.2019

Turning briefly to this study, the complexity of foreign language policy in Switzerland hindered the evaluation of the spoken language of the entire country. Consequently, the decision to concentrate on the spoken output of Central Switzerland was made. Central Switzerland belongs to one of four regional conferences of the Swiss Conference of Cantonal Directors of Education (EDK). Central Switzerland is defined as including the cantons Lucerne, Uri, Schwyz, Obwalden, Nidwalden, and Zug as seen in Figure 1-6 below.
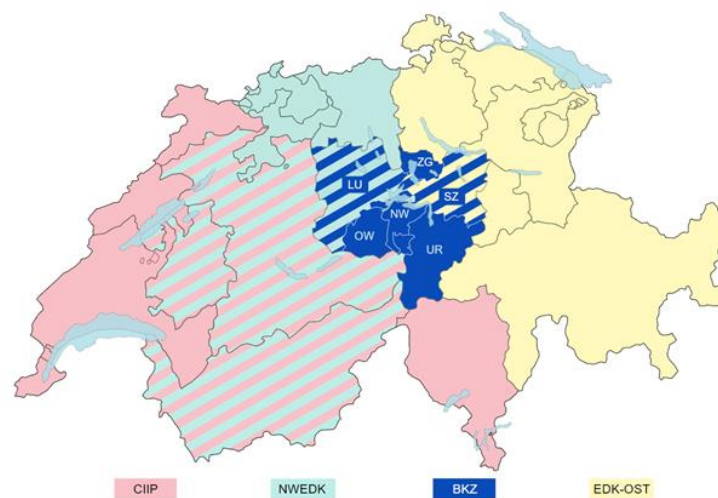
Figure 1-6 The four regional conferences of Swiss Conferences of Cantonal Director of Education (EDK)[7]

The shift from commencing English instruction in secondary school to primary school has had numerous repercussions. Most significantly, primary school teacher education has needed to be expanded to include obligatory English or French language methodological and didactic training. Prior to the language policy change, foreign language skills were not a requirement for becoming a primary school teacher. Currently to qualify as a Swiss primary school teacher, one must receive a bachelor from one of the Swiss universities of teacher education where one is trained to be able to teach seven primary school subjects. This is a sharp contrast to the qualifications needed to teach in a Swiss secondary school. Secondary school teachers must

---

[7] Source https://www.bildung-z.ch/bkz/organisation accessed 07.05.2021

have a master's from a Swiss university of teacher education and specialise in only four subjects (Pädagogische Hochschule Luzern, 2020). Another difference between teacher qualifications is the required English proficiency levels. Primary school teachers are required to have a B2 level compared to the C1 level of secondary teachers. However, it is not a necessity to verify proficiency with an international certificate (Hardmeier, 2017, p. 4). Thus, before the decision to teach English in primary school, English was taught in secondary school by foreign language specialists to teenagers and is now being taught by less educated generalists to children. It was hypothesised that this new situation would have an effect on the students' language output and investigating the difference would be worthwhile.

As mentioned above, the idea of a multilingual Switzerland is embedded in the Swiss cultural identity and is anchored in the new foreign language policy where the language objectives are not perfect bi or trilingualism, but rather training for functional multilingualism which is defined as:

> 'Functional multilingualism strives for a diverse, dynamic repertoire with differently advanced skills in different areas of competence or languages in order to be able to act linguistically successfully in different situations.' (translated from: Kanton Luzern Bildungs- und Kulturdepartement, 2020).

Didactic guidelines strengthen this by suggesting that 'The communicative intention always comes before formal correctness.' (translated from: Kanton Luzern Bildungs- und Kulturdepartement, 2020). The cantonal guidelines take into account that errors are to be expected and dealt with depending on the learning situation. For example, when promoting the flow of speech, corrections should be made cautiously. The guidelines further suggest that a student's individual interlanguage should be developed by encouraging conscious transfer between languages. In turn, false analogies (e.g., over-generalization) are considered or even expected to occur with the resulting errors being used as an indicator of the current proficiency

level (Kanton Luzern Bildungs- und Kulturdepartement, 2020). This study will document some of these analogies.

The Swiss approach is reminiscent of English as a lingua franca (ELF) advocates who discourage non-native English speakers from conforming to native speaker norms and put emphasis on the communicative success over native-like grammatical correctness and pronunciation. This could be deemed acceptable if it were not the case that native English institutionalized exams such as Cambridge are used as formal exit exams for Swiss students in apprenticeships and many other scholastic areas. At present, students begin with a playful implicit learning approach for four years in primary school and then are faced with explicit form focused learning for an additional three years in secondary school. The transition from primary to secondary foreign language learning was researched by Pfenninger and Lendl (2017). Although they did not find a general problem with the transition, they found a need for better coordination and communication between schools and acknowledgement of the value of implicitly learned language skills to maintain motivation. Furthermore, the primary teacher's mastery of English seems to have been a determining factor in the student's perception of their own success. The fossilization of non-standard usages in teenage and adult learners is a common discussion theme and concern in Swiss language school teacher rooms. Thus, the Swiss school system could be unjustly creating difficulties for students who want to reach the highest language proficiency levels by delaying attention to grammatical accuracy past the stage where it could be cognitively processed. It is a widely held view that if students are not exposed to impeccable language and gently guided to recognise variations from the beginning, the ability to master English could be hindered.

The viewpoint of promoting awareness of grammatical correctness in the classroom does not necessarily transfer to everyday life where perfection is not required. As Werlen points out, oral interaction skill is conceptualized as the ability of functionally communicating with an interlocutor, making oneself understood in a context-adequate manner rather than as grammatical correctness and phonological accuracy (Werlen, 2006, p. 11).

Thus far, the complexity and depth of the Swiss linguistic landscape and its coordinating education system has been shown to aid in understanding the parameters in which this study is set. In sum, Swiss residents live in a country with four national languages and 25% non-Swiss inhabitants, are expected to learn at least two foreign languages commencing in primary school and function within primarily monolingual bureaucratic municipalities. The following section presents the research aims and specific questions that were designed to contribute to the existing knowledge of the English spoken in Central Switzerland.

1.4   Aims of thesis

The overarching aim of this thesis is to provide better understandings of the spoken English of the emerging Swiss workforce. As mentioned above, a better understanding of spoken English could have important practical implications for foreign language policy in the Swiss context. To date, a Swiss spoken English learner corpus has not been compiled and empirically investigated. The study is guided by three main research questions.

RQ 1: What is the relationship between the use of non-standard English features and speakers' demographic, identity-related and educational variables?

In order to answer this research question, samples of spoken language produced by young Swiss speakers of English were collected, non-standard usages identified and compared with the subsequent variables. No previous Swiss study has provided information on the effects of these variables on the spoken output of the English of this age group. The evaluation of demographic information in this study can reveal if English language learning in Switzerland is influenced by nationalities, mother tongues and languages spoken at home. Furthermore, identity-related variables have been known to influence spoken output (Cook, 2002). Lastly, educational variables about the length of instruction and proficiency levels reached can give insight into the results of current practices and offer pertinent themes for future discussion.

This study also seeks to assess the possible impact of recent changes in Swiss language policy. The changes involved commencing English language instruction in primary school as the first foreign language in grade three in 2005. In the previous policy, English was the second foreign language taught after French and instruction began in the seventh grade.

RQ 2: What effect does early English, taught with a productive focused curriculum, have on the number and type of non-standard features used?

Previous studies have focussed on the effect early English learning has had on the performance of primary school children's French and German proficiency and on the transition from the primary to secondary school language classroom. However, no research has examined and compared the outcome of an additional four years of English instruction has on spoken English. Not only the number of lessons was increased, but the teaching methods shifted from deductive to inductive (Pfenninger and Lendl, 2017, p. 451). The results of this research question can give empirical evidence that early English is beneficial or highlight areas in need of improvement.

In order to address this research question, half of the samples of spoken language mentioned above were compiled from participants who began learning English in secondary school and the remaining half was compiled of samples from participants who began learning English four years earlier in primary school. The non-standard features found in each half were then compared to determine what correlations exist between the age when English instruction began, and number or type of non-standard features used.

This thesis also intends to determine the extent to which the emerging Swiss workforce accepts non-standard features. Standard language norms are determined by their users and change over time. In the case of English, we have standard norms dictated by native speakers and English varieties which have evolved in non-native speaking counties with their own norms. Bamgbose (1998, p. 3) suggests five factors for determining the status of an innovation: demographic, geographic, authoritative, codification and acceptability, the acceptability factor being 'the ultimate test of admission of an innovation' (Bamgbose, 1998, p. 4). By isolating non-native features of the spoken corpus created in this thesis and exposing Swiss respondents to those features the third research question is posed.

RQ 3: How is the acceptance of commonly used non-standard features perceived by the Central Swiss workforce?

The third research question was addressed by selecting ten sentences with non-standard features from the compiled corpus and then asking respondents to decipher their level of acceptability by means of an online survey. Acceptability of non-standard language is the beginning of change, and the results can contribute to the documentation of that change.

1.5   Outline of the thesis

Chapter 1 has introduced the Swiss context in which this study was conducted. The uniqueness of languages in this quad-lingual country and its education system were highlighted to illustrate the opportunity that foreign language policy change has provided to investigate the effects of early English on spoken output.

Chapter 2 reviews pertinent literature on corpus linguistics and types of corpora which emphasise the current research gap of a Swiss learner corpus to explore spoken output. It also provides an overview of previous learner corpora studies and the methodology used to investigate learner corpora such as error-tagging text to reveal error frequencies for comparison and evaluation. Standards and the acceptance of non-standard usages conclude the chapter.

The methodology used to create and explore the corpus is explained in detail in Chapter 3. Interviews as means of data collection based on the compilation of the LINDSEI corpus are clarified as well as the profiles of participants with the extensive metadata which was collected. Lastly, the transcription and annotation process of the 25 non-standard usages is described. Then, in Chapter 4 the demographic, identity-related and educational variables are explored and their statistical relevance in relation to non-standard usages revealed before each non-standard usage is thoroughly investigated and discussed in Chapter 5.

In Chapter 6, the survey conducted to access the acceptability of non-standard sentences found in the corpus is presented and the acceptability rates discussed before Chapter 7 discusses and concludes the results, implications and contribution of this thesis.

Chapter 2 **LITERATURE REVIEW**

2.1   Introduction

This sociolinguistic study seeks to understand the spoken output of a group of young adults who are at the beginning of their careers by investigating the frequency of non-standard usages in their spoken language production. In Switzerland, it is likely that this age group have completed their formal foreign language instruction. The interest lies in better understanding what their rucksack of knowledge to master spoken discourse in English includes and how the contents differ depending on measurable variables including the timing and length of learning English in the classroom. A broader understanding could highlight advantages and facilitate the alleviation of disadvantages in the education system, with the goal of gaining knowledge which can be collected, computed and examined for commonalities and dissimilarities.

This chapter reviews the literature of past and current research and theories so that the contribution of this study can be established. Firstly, corpus studies will be addressed. Thereafter, the definition and impact of standardisation will be discussed before turning to the speaker and their position in second language acquisition as well as language change through acceptance of non-standard features.

2.2   Corpus linguistics

'Language cannot be invented; it can only be captured.'

(Sinclair, 1997, p. 31)

This study is situated in the field of sociolinguistics and corpus linguistics by examining variables and their effect on language output through a corpus construct. The key interests of Sociolinguistics are the social impact on how language is performed and linguistic variation

within language(s) (Hinrichs and Bohmann, 2020, p. 283). Corpus linguistics on the other hand 'is concerned with collecting and designing corpora and using corpora and corpus linguistic methods to study language' (Aijmer, 2020, p. 6). Hence, the choice of creating and using a corpus to investigate the language variation in Swiss youth is warranted. As Hunston (2002, p. 20) reminds us, a corpus is more reliable than native speaker intuition in telling us what a language is like.

Modern-day electronic corpora started in the 1960s with the first American and British national corpora of one million words each. Today, with advancements in digitalisation, immediate access to digitalised written text has become standard as throughout the world anyone with a smartphone has access to search engines with trillions of words. Besides advancing the understanding of how language works, written corpora have become a driving force for language exploration with dictionaries and coursebooks alike relying heavily on corpus linguistics research to inform on the current state of English use. Although 'many language scholars and teachers believe that the spoken form of the language is a better guide to the fundamental organisation of the language than the written form' (Sinclair, 1991, p. 16), and spoken language is considered the most usual means of using language, (Kennedy, 1998, p. 20) spoken corpora tend to be less common. This can mainly be attributed to the resources needed to produce it (Adolphs and Knight, 2010, p. 41).

Spoken corpora are used in many fields of investigation; language teaching and learning, learner language, second language acquisition, English as a lingua franca, discourse analysis, translation studies, forensic linguistics, pragmatics, and social linguistics (O'Keeffe and McCarthy, 2010). This study uses a learner corpus to investigate spoken output and the variations of Standard English within. Gablasova et al. (2017, pp. 132-133) propose two

approaches to investigating variation. The first is a formal approach in the tradition of Labov (1972a) to explore different ways of expressing the same meaning. The second which is used in this study is the functional approach in the tradition of Biber (1988) and focuses on the frequency of linguistic features. The importance of frequency is advocated by Gablasova et al. (2017, pp. 132-133) as a valuable resource for understanding and comparing spoken learner corpora. Furthermore, Leech (2011, p. 5) argues that information about frequency is a benefit that cannot be obtained by other means. Leech (2011) draws our attention to three uses of frequency. First, 'raw frequency' is simply how many times something occurs in the corpus. This is used in frequency lists to compare the general composition of a corpus. Second, 'normalized frequency' is used to compare the occurrences to other corpora by calculating the occurrences within a set number of words. In this study, the normalized frequency is set at per one hundred words to facilitate comparison with other like corpora. Thirdly, 'ordinal frequency' allows comparison of two or more features such as the relation of *going to* and *will* future. Ordinal frequency facts are considered highly valuable for pedagogical reasons as pointed out by Leech (2011 p. 8).

Turning to corpus design, 'A corpus can show nothing more than its own contents' (Hunston, 2002, p. 22). Keeping this in mind, Thompson (2005) reminds us that a spoken corpus needs to be carefully compiled with the purpose at the forefront and with the size and degree of transcription complexity designed to match that purpose. Perhaps the most influential account on how to design a corpus can be found in the work of Sinclair who pioneered work in corpus linguistics. Sinclair (2005[8]) set forth the following guidelines for spoken corpus design with keywords marked bold by the author:

---

[8] Online source, page number not available http://users.ox.ac.uk/~martinw/dlc/chapter1.htm

1) The contents of a corpus should be selected without regard for the language they contain, but according to their **communicative function** in the community in which they arise.

2) Corpus builders should strive to make their corpus as **representative** as possible of the language from which it is chosen.

3) **Only** those components of corpora which have been designed to be independently contrastive should be **contrasted**.

4) Criteria for determining the structure of a corpus should be **small in number, clearly separate** from each other, and efficient as a group in delineating a corpus that is representative of the language or variety under examination.

5) Any **information** about a text other than the alphanumeric string of its words and punctuation should be **stored separately** from the plain text and merged when required in applications.

6) Samples of language for a corpus should wherever possible consist of **entire documents or transcriptions** of **complete speech events**, or should get as close to this target as possible. This means that samples will differ substantially in size.

7) The design and composition of a corpus should be documented fully with **information about the contents** and arguments in justification of the decisions taken.

8) The corpus builder should retain, as target notions, **representativeness** and **balance**. While these are not precisely definable and attainable goals, they must be used to guide the design of a corpus and the selection of its components.

9) Any **control** of subject matter in a corpus should be imposed by the use of **external**, and **not internal, criteria**.

10) A corpus should aim for **homogeneity** in its components while maintaining adequate coverage, and rogue texts should be avoided.

The above guidelines provide researchers with sound advice which leads to qualitative corpora. However, there are areas of debate. Sinclair preferred corpus text to be minimally annotated or raw as mentioned in number five above. Leech (2005) and McEnery and Hardie (2012, p. 31) on the other hand, advocate annotation as an added value. A further observation of the guidelines is that an appeal for appropriate/sufficient metadata to facilitate research based on

comparison of particular groups is missing. Burnard (2002) acknowledges that inconsistencies in metadata information was a problem when compiling the British National Corpus (BNC). The urgent need to standardise metadata, especially in learner corpora is argued by Granger et al, (2005). As the field of learner corpora grows, Gablasova et al, (2017, p. 137) claim that 'corpus representativeness has direct implications for corpus comparability, the degree to which two corpora are similar.' and argue that validity of corpus findings is dependent on it. In addition, Diaz-Negrillo and Thompson (2013) predict that learner corpora annotation will become more sophisticated in the future. As technological advances are made in voice recognition software, annotation software and multimodal corpora, it appears that an updated set of guidelines would aid in the future comparison of spoken corpora and their findings.

In the following sections selected corpora, their composition and relevant research outcomes pertaining to this study will be highlighted. First, key data points and research of two general corpora are introduced. Then, three specialized teen corpora used for frequency comparison in Chapter 4 are introduced before the spoken learner corpus LINDSEI and its relevance to this study is revealed. Finally, the three leading corpora used in ELF research are briefly introduced.

2.2.1   General corpora

A general corpus can broadly be defined as a larger corpus compiled of a variety of texts. They can contain written and or spoken texts and are often called reference corpora because they are often used as a 'baseline' of comparison for more specialized corpora (Hunston, 2002, pp. 14-15). With their larger size, general corpora aim to be as representative and balanced as possible. This can be achieved by including language samples from a wide variety of sources and genres to embody sufficient language to investigate a range of linguistic features (Reppen and

Simpson-Vlach, 2020, p. 93). Two of the most widely used general spoken corpora are introduced below.

2.2.1.1 CANCODE Cambridge and Nottingham Corpus of Discourse in English

CANCODE a five-million-word spoken corpus collected between 1995 and 2000 is a native speaker general corpus that was compiled using a genre approach which allows analysis of different perspectives such as speakers and events as opposed to a demographic approach that would represent a language population (Clancy, 2010, p. 84). The events encompass transactional, professional, pedagogical, socialising and intimate settings to represent spoken English (McCarthy, 1998, pp. 8-10). Research stemming from the corpus includes uncovering discrepancies between ESL coursebook dialogue and real spoken language such as the underuse of vague language and discourse markers (Carter, 1998a). Furthermore, CANCODE (McCarthy, 1998, pp. 8-10) was commissioned to describe spoken grammar (Carter and McCarthy, 2006). It has further served as a reference corpus in numerous studies (Tagg, 2009; Gablasova et al., 2017). However, one major drawback of this corpus is that besides highly published frequency lists and examples, the CANCODE corpus is inaccessible outside of Cambridge University Press (Love et al., 2017, p. 322).

2.2.1.2 BNC British National Corpus

The British National Corpus, compiled from 1991-1994, is a general corpus which represents a cross-section of modern British English. One tenth of the one-hundred-million-word corpus is spoken. The spoken section was designed in two parts, one demographically to establish random sampling considering age, gender, social group and geographic region. The other part was context-governed with educational, business, public/institutional, and leisure sub-sections (Burnard, 2007). Relevant recent research includes Hadikin (2014) exploring lexical priming

of *a*, *an* and *the*, in Korean English and Shirato and Stapleton (2007) who investigated Japanese learner vocabulary. Free accessibility has aided in the BNC becoming 'a highly productive resource for linguistic research over the last two decades' (Love, 2017 p. 322). Keeping with the basic construct of the original BNC for comparison and with added improvements, the new 11.5-million-word BNC2014 offers a completely updated version with all new data from first language (L1) British English speakers. This new corpus has filled the researcher's gap of a large, present day, readily available and accessible corpus of spoken British English with orthographically transcribed data. The data was collected through public participation and the use of smartphones. A wide range of viable metadata was also collected to enhance research potential. Due to the availability of the original spoken version when this study began, it was used for comparisons.

2.2.2 Specialized corpora

A specialized corpus as its name implies is created to include a certain type of data or used for a specific purpose. Teenager Corpora falls in this category. 'Most sociolinguistics agree that adolescence is the "focal point for linguistic innovation and change"' (Tagliamonte, 2016, p. 3). Thus, teenager talk represents the language of change. With over eighty percent of the data collected in this study of participants below the age of twenty, frequency comparisons with the following three teenager corpora will aid in the analysis of their spoken output.

2.2.2.1 COLT Bergen Corpus of London Teenage Language

The Bergen Corpus of London Teenage Language COLT (Stenström et al., 2002) was the first spoken corpus of teenagers and is an integral part of the BNC. It represents London teenagers aged thirteen to seventeen and contains 500,000 words. The recordings were made in a similar way as the 1994 BNC spoken corpus with teenagers given tape recorders and the task to record

everyday conversations. The resulting data gave insight into the language and lives of London teenagers of the 1990s (Stenström et al., 2002, p. 211). Research findings included an unexpectedly even use of vague words throughout social classes and gender. The 'unorthodox' use of *be like* in reported speech was noted as well as the use of the invariant tags *eh*, *okay*, *right*, *yeah* and *innit* were found to be typical of teenager speech (ibid., p. 213).

2.2.2.2 MLE Havering Multicultural London English Corpus

The Havering Multicultural London English Corpus MLE (Cheshire et al., 2007-2010) represents the spoken language of seventeen-year-olds with a multicultural language background from outer London. The data originated from the *Linguistic Innovators* project (2004–2007) where two boroughs of inner-city London, Havering and Hackney, investigated the 'social dynamic of language change' (Kerswill, 2013, p. 2). Havering, representing a white Anglo majority with fourteen percent school children with English as a second language. Hackney, in comparison, was a more diverse community with fifty percent school children with English as a second language. The 1.4-million-word corpus comprised of the interviews or casual speech of nighty-eight individuals reported research findings included innovative adolescent speech and a high use of quotative expressions with *go* and *be like*. The intensifier *really* was found to be used by adolescents twice as often as the elderly in the corpus which is in line with previous research (Kerswill, 2013, p. 11).

2.2.2.3 TTC Toronto Teen Corpus

The Toronto teen corpus, compiled between 2002-2006 represents 90 Canadian speakers aged nine to twenty and is twice as large as the COLT corpus with one million words (Tagliamonte, 2016, p. 12). When comparing word frequency with the MLE, Tagliamonte found that *like*, *just* and *so* were in the top 30 of all three corpora despite the time differences of data collection and

British and North American varieties. This phenomenon is not found in the Toronto English Archive 2002-2006 (Tagliamonte, 2006) of Canadian adults aged sixty and above. In terms of generational language change, Tagliamonte (2016) depicts the generational transition from *very* to *really* to *so* with *really* presently being the most common intensifier of all age groups and *so* the intensifier of youth.

Overall, these corpora and the resulting findings illustrate teenager language innovation and possible directions of language change in communities where English is considered the native language. However, there are far more communities around the world where English is spoken as a second or foreign language and that spoken output also has the possibility of shaping the English of the future.

2.2.3   Learner corpora

Granger defines learner corpora as 'electronic collections of (near-) natural foreign or second language learner texts assembled according to explicit design criteria' (Granger, 2008, p. 338). Design criteria include controlling learner and task variables. Although Granger's definition is generally not disputed, viewpoints on wording and degree of representativeness and naturalness differ slightly (see Gilquin, 2015). The Centre for English Corpus Linguistics (CECL) at the Université catholique de Louvain is at the forefront of learner corpora research with currently fifteen corpora created and a great deal of research conducted. The International Corpus of Learner English (ICLE) is a written corpus of essays by upper intermediate and advanced learners, and it has a native English counterpart; the Louvain Corpus of Native English Essays (LOCNESS). The Louvain International Database of Spoken English Interlanguage LINDSEI complements ICLE and is of special interest to this study.

2.2.3.1 LINDSEI The Louvain International Database of Spoken English Interlanguage

The LINDSEI corpus (2010) is a spoken corpus of approximately one million words, comprised of English L2 learners. The eleven sub-corpora were compiled of third- or fourth-year undergraduate university students majoring in English from ten countries. Three-part interviews were conducted to obtain spoken language. The interviews began with a set task with a choice of three topics, then the interviewer conducted a conversation by asking questions about the learner's life before they were asked to describe a picture sequence. Metadata was gathered on age, gender, mother tongue, country of origin, other foreign language knowledge and stays in English speaking countries (Gilquin et al., 2010, pp. 8-12). In addition, development of the transcription and markup guidelines followed a 'minimalistic transcription standard' (ibid. p. 13). The LINDSEI project has continued to expand with currently[9] twenty sub-corpora and five sub-corpora in progress.

The LINDSEI objectives include better understanding learners' spoken skills and to 'provide a solid empirical basis for the design of efficient pedagogical tools' (ibid. p. 4). Thus, creating an environment where the methodological approach of (CIA) Contrastive Interlanguage Analysis (Granger, 1996) can flourish. CIA is a method of finding non-nativelike features and characterising interlanguage by comparison with a native reference corpus. CECL created the Louvain Corpus of Native English Conversation LOCNEC which is the native speaker counterpart to LINDSEI for this purpose. Research using CIA has brought to light useful information about learner language which ideally will flow into pedagogical practice.

LINDSEI Research to date includes the study of lexis, discourse, methodology, pragmatics, pronunciation, grammar and fluency. The methodology used in these studies can generally be

---

[9] Accessed 24.04.2021 https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei-partners.html

placed in two categories which can be used separately or combined. First, there are comparative studies which apply CIA as mentioned above. Comparison is possible against a native corpus, or between two or more L1 groups. Many studies have focused on over and underuse of linguistic features based on native speaker use. Critique of this method stems from the interpretation of representativeness and degree of comparability (Callies, 2015, p. 40) due to the many factors that are involved in speech acts used to compile corpora such as task, level of speaker proficiency and L1 cultural aspects. Furthermore, the construct of comparative corpora in aspects of selection of participants, mode of data collection and transcription can affect representativeness and the validity of results of such undertaken comparisons. Although these areas of potential criticism are forewarned in Sinclair's guideline points 2, 3, and 8 above, careful weighing of all aspects of corpus construct and meticulous evaluation of comparability can rationalize the interpretation of results. Granger (2015) has addressed criticism of CIA and proposed revisions which reflect the 'variationist trend' by suggesting the terms 'Reference Language Varieties' (RLV) which includes reference varieties beyond native speaker and 'Interlanguage Varieties' (ILV) to refer to learner language (ibid. p. 15).

The second widely used methodology is computer-aided error analysis which entails investigating error-tagged text to reveal error frequencies for comparison and evaluation. It has been instrumental in investigating learner interlanguage and giving insight into language used at different proficiency levels (Götz, 2015, p. 191). The use of annotation programmes can greatly facilitate tagging large corpora. However, a general consensus on 'standardization of error typologies' has not been reached (Diaz-Negrillo and Fernandez-Dominguez, 2006, p. 86) and thus is an area of potential improvement. Challenges to consider when investigating an error-tagged corpus include error identification; grammatical vs. pragmatic, overt vs. covert (Corder, 1971), errors vs. mistakes (Corder, 1967) and which norm is used as a reference

(Callies, 2015, p. 41). When all factors are considered computer-aided error analysis remains a methodology which 'enables researchers and language testers to describe language proficiency on a quantitative level by way of characterising the frequencies, types and contexts of errors that learners commit at a certain proficiency level' (Götz, 2015, p. 192). This study avoids the term 'error' in describing language diverting from standard norms as non-native and prefers the more inclusive term 'non-standard usage' (NSU). This terminology widens the spectrum of use to common standards or norms used in Englishes and encompasses all English speakers.

To exemplify the learner corpus methodology mentioned above, eight studies employing LINDSEI corpora are briefly reviewed below. First, Gráf (2017) shares insight in the story of the Czech LINDSEI creation. Then four studies and their results applying CIA are highlighted before one example of computerised error analysis is presented. Lastly, two German LINDSEI studies which combine CIA and CEA are reviewed.

LINDSEI Czech Republic sub-corpus

By discussing concerns about the current methodologies used in the field of learner corpora, Gráf (2017) offers insight and suggestions on possible improvements in the design and construction of LINDSEI sub-corpora to increase reliability and exploitation. Although he states that 'LINDSEI is an invaluable source of highly authentic learner data', he observed that more information was needed about the scope of the interlocutor's role in questioning and seemed under the impression that the interviewer's role might be to elicit certain grammatical structures. However, the elicitation of certain grammatical structures has to the knowledge of the author not been mentioned in Granger's publications. On the contrary, she claims that 'purely experimental data resulting from elicitation techniques does not qualify as learner corpus data' (Granger, 2013, p. 5). Gilquin et al. also point out that although the LINDSEI

corpus is not constructed of fully natural data, the interviews, apart from the picture description, are close to natural learner language and although the picture description is 'heavily constrained', the 'learners could use words of their own choice' (Gilquin et al., 2010, p. 6). Specific interview elicitation techniques would go against Sinclair's guidelines mentioned in Section 2.1 and place emphasis on language instead of the communicative function. Moreover, elicitation tasks are intended to persuade speakers to produce certain features and are more common in (Quasi-)experimental Second Language Acquisition (SLA) studies (Callies, 2015, p. 42).

Gráf raised several questions related to LINDSEI sub-corpus comparability. Transcription subjectivity due to discrepancies in length of pauses and vowel/syllable length which could distort results was stated as a concern. Furthermore, the suggestion of increased metadata collection and refined proficiency determination criteria appears justified because a random sampling of five learners from each countries' sub-corpus revealed that 64% of the extracts were rated higher intermediate and only 36% advanced when the project had aimed for data from advanced learners (Gilquin et al., 2010, p. 10). To reach proficiency goals, the Czech sub-corpus is being adjusted to replace learners with a proficiency level below a C1 level with learners at a C1/C2 level to assure a truly advanced corpus. Despite initial methodological concerns, the LINDSEI Czech sub-corpus has been successfully used to investigate articles, tenses, accuracy, fluency and disfluency.

LINDSEI Chinese sub-corpus

Quan and Weisser (2015) used a CIA approach to compare learners' syntactic/lexical elements of recycling and replacement 'self-repair' with native speakers. The Chinese component of LINDSEI and the native speaker LOCNEC were used for comparison. The results revealed that

Chinese English learners used more verbs than native speakers in recycling and replacement forms of self-repair. Quan and Weiser define recycling as repeated segments of speech and replacement refers to speech segments which are repaired or corrected. Moreover, an overuse of empty, delexicalized verbs such as *make*, *do*, and *have* was also found. They argue that L1 transfer and typological language differences could contribute to this. It was suggested that L1, English proficiency and age may contribute to the native and non-native speaker differences.

LINDSEI Polish sub-corpus

The Polish LINDSEI sub-corpus was used to investigate the underuse of idiomatically opaque phrasal verbs by Wierszycka (2013). She used a part of speech (POS) tagged version of the sub-corpus against the native speaker LOCNEC corpus to conduct a Contrastive Interlanguage Analysis. Her hypothesis that Polish English learners would substantially underuse phrasal verbs was validated with opaque phrasal verbs being neglected most. After the variables of stays abroad and years of English at school were found to be inconclusive, general avoidance of phrasal verbs or preferences of one-word equivalents was suggested. Most important for this study was the question of varying proficiency levels being raised as possible justification for the phrasal verb underuse.

LINDSEI Turkish sub-corpus

Kizil and Kilimci (2014) investigated structural and functional properties of the recurrent phrases in their corpus-driven Contrastive Interlanguage Analysis study. They used the Turkish sub-corpora of LINDSEI and its native counterpart LOCNEC. The underuse of vagueness markers and hedging devices in the Turkish learners' spoken English was verified and in line with previous research. The study thus was able to produce specific pedagogical recommendations based on their findings.

LINDSEI multiple sub-corpus study

Buysse (2015) investigated *well* as a discourse marker in a Contrastive Interlanguage Analysis of the Dutch, French, German, Spanish and Chinese components of the LINDSEI corpus and Aijmer's (2011) findings of her analysis of the Swedish LINDSEI component and the native speaker LOCNEC corpus. The results showed that all learner corpora except the Chinese contained higher rates of *well* as a discourse marker, particularly when related to speech management. This is an example of using multiple sub-corpora to investigate pragmatic features across L1s and attempt to draw conclusions and determine further research directions. The homogeneity of the corpora in relation to proficiency levels, exposure to English and L1 interference, however, raised the question of validity.

LINDSEI Japanese sub-corpus

An error analysis was conducted with the Japanese LINDSEI sub-corpus by Kaneko (2009). She investigated the use of Japanese L1 in learner speech. The corpus was divided into two proficiency groups and the frequency and type of L1 use was determined. It was found that although less than 0.4% of the corpus was in L1, almost 3% of turns included at least one L1 word. The differences in proficiency groups were marked with the lower group using more L1 interjections and the higher group using more lexical variety. The results of this study indicate a clear relationship between proficiency level and type of L1 language inserted while speaking English.

LINDSEI German sub-corpus

The German sub-corpus of LINDSEI (Brand and Kämmerer, 2006) is the most comparable to the Central Swiss Corpus (CSC) compiled for this study, with the same first language, although most Swiss learners have the syntax and lexis from two varieties of German in their repertoire.

Although a vast amount of research has been conducted in relation to the German sub-corpus, two studies, in particular, are of importance to this study.

Stefanie Dose-Heidelmayer and Sandra Götz (2016)

The progressive in spoken learner language: A corpus-based analysis of use and misuse

This study is an example of qualitative error analysis used to investigate frequency and error types in spoken learner interlanguage. The authors examine the use of progressive by German L1 advanced learners of English. The progressive is a challenging aspect for German L1 learners due to its almost non-existence in L1. Previous studies of written output suggested an overuse of progressive as compensation in line with the 'Aspect Hypothesis' (Anderson and Shirai, 1994). The authors investigated the German error-tagged sub-corpus of LINDSEI-GE and compared it to the native speaker counterpart LOCNEC. Initial findings on frequency revealed that although the German learners use the progressive more frequently in spoken than written output, a highly significant underuse was found when compared to the native corpus. Upon closer investigation, it was found that the German learners were a heterogeneous group with considerable intra-corpus variation and furthermore, there was significant variation in progressive use depending on task type in the data. Next progressive related errors of under and overuse were investigated to reveal that a surprising 80% of errors were related to overuse. This result failed to explain the general underuse of progressive and the authors turned to investigating individual learners which revealed that the errors of overuse cannot be considered a general phenomenon, but rather related to a small portion of individual learners. In sum, this study highlights the need to consider individual learners as well as group performance, the effects of task type when comparing frequency, and proficiency levels.

Robert Fuchs, Sandra Götz, and Valentin Werner (2016)

The present perfect in learner Englishes: A corpus-based case study on L1 German

intermediate and advanced speech and writing

This study is an example of a corpus-based Contrastive Interlanguage Analysis to investigate German L1 learners of English at various proficiency levels against native speakers in written and spoken modes. To perform the analysis of present perfect use, the authors used five corpora; the German sub-corpus of LINDSEI (Brand and Kämmerer, 2006) and its native speaker counterpart LOCNEC, the German component (GICLE) of the International Corpus of Learner English (ICLE; Granger et al., 2009) and its native speaker counterpart (LOCNESS), and the German component of International Corpus of Crosslinguistic Interlanguage (ICCI; Tono, 2012). The corpora were tagged with the CLAWS tagger C7 tagset (Garside and Smith, 1997) before investigation began. This combination of corpora allowed the authors to investigate additional variables such as stays in English speaking countries, years of English at university and differentiate between early and late English learners.

Fuchs et al. (2016) identify present perfect use as a particularly difficult aspect to learn and acknowledge that previous research primarily using observation, elicitation tasks and interviews indicate that acquisition comes late. The German equivalent to present perfect differs slightly in written form and is seldom used in spoken interactions. After a thorough investigation, the authors uncovered that neither do early starters use present perfect more nor do stays abroad increase the use of present perfect in university students. Furthermore, learners of English with L1 German rarely use the present perfect until after school year nine, where an increase of present perfect use is recorded in years ten and eleven to advance towards native speaker use. The general underuse of present perfect in written and spoken modes was verified with proficiency levels having no bearing. Lastly, their findings support the 'Default Past Tense

Hypothesis' that simple past is acquired before present perfect. In sum, this study highlights how the use of multiple corpora can be used to conduct Contrastive Interlanguage Analysis of same L1 learners as well as native speaker corpora to question or confirm the hypothesis.

Overall, these studies provide reasonably consistent evidence of the importance of learner corpora research. Most of the research reviewed here mentioned proficiency as a possible factor in explaining over or underuse of certain spoken language features. However, one of the recurring threads of concern in comparativeness of corpora is the varying proficiency levels of the individual learners and if proficiency is not adequately assessed, generalisability of results is limited (Thomas, 2006). This study integrates the proficiency level factor in the analysis of the multilevel CSC corpus as a source of evidence to describe frequency of non-standard usages which in turn can be used to offer pedagogical insight. 'The analysis of errors provides actual evidence of the areas which learners still need to master, therefore disclosing their pedagogical needs.' (Diaz-Negrillo and Fernandez-Dominguez, 2006).

2.2.4   English as a lingua franca corpora

Another type of corpus focuses on the linguistic exploration of English used as a lingua franca. This is where English is used to communicate when first languages differ. See Section 2.3.3 for further information about the definition of English as a lingua franca. ELF research has explored accommodation, code-switching, new words, aspects of lexicogrammar, resolving miscommunication, establishing rapport, employing communicative strategies and overall communicative behaviour (Seidlhofer, 2010, p. 365). The three major ELF corpora and highlighted below.

2.2.4.1 VOICE Vienna-Oxford International Corpus of English

The one-million-word VOICE corpus from the department of English at the University of Vienna was the first large spoken corpus compiled of non-native English speakers. Directed by Barbara Seidlhofer, the aim of the project is to encourage ELF research by giving open access to linguistic researchers all over the world. The website states the areas of study are corpus linguistics and language variation and informed by the three perspectives conceptual, descriptive and methodological[10].

2.2.4.2 ACE The Asian Corpus of English

ACE is similar to VOICE as it uses the same transcription software and conventions which allows for comparison of European and Asian spoken English (Kirkpatrick, 2016, p. 225). The one-million-word spoken corpus was compiled of speakers from nine Asian countries using naturally occurring interactions (ACE, 2020). Research on identity has suggested that ELF has not only functional but emotive value and that some shared pronunciation features spanning across at least five Asian countries were observed (Kirkpatrick, 2016, p. 227).

2.2.4.3 ELFA The Corpus of English as a Lingua Franca in Academic Settings

The one-million-word ELFA corpus is compiled of spoken English of non-native speakers in an academic setting. It was compiled in 2008 at the University of Helsinki, directed by Anna Mauranen, and is designed to investigate ELF in theoretical, descriptive, and applicational research. In addition to ELFA, the university of Helsinki completed the Written English as a Lingua Franca in Academic Settings WrELFA corpus in 2015, a one and a half-million-word corpus of written academic texts.

---

[10] Source: https://www.univie.ac.at/voice/page/research_perspectives accessed 07.05.2021

2.2.5 Swiss corpora

To my knowledge, the only Swiss English corpora compiled and investigated to date is the Swiss English Database[11] collected in connection with and funded by The Swiss National Science Foundation. The project 'Language Contact and Focussing: The Linguistics of English in Switzerland' investigated the existence of a 'Pan Swiss English' across three of the four national languages; German, French and Italian. The project was based on the hypothesis that focussing was taking place (Rosenberger, 2009, p. 130). Trudgill defines focussing as 'the process by means of which [a] new variety acquires norms and stability' (Trudgill, 2004, p. 88). Results from the three PhD studies Durham (2007), Rosenberger (2009) and Dröschel (2011) did not validate the hypothesis, although the documentation of the language gives a standpoint for future reference. Furthermore, the studies provide valuable in-depth research into the use of English in Switzerland. Although corpora comparability with the CSC might be of concern, the findings resulting from these studies are the only sources of Swiss English currently available for comparison.

Durham (2007) utilized the email portion of the Swiss English Database, compiled of informal medical student association emails written by Swiss nationals in English with German, French and Italian L1s. As a reference, a similar type of database of informal emails from a UK association was used. Durham investigated the existence of a Swiss English variety that was shared by speakers of the three main Swiss languages German, French and Italian. Of the five features examined: non-count plural forms, future tenses, relative pronouns, complementizers and *also*, *as well* and *too*, only the future with *will* vs. *going to* could be considered a possible area of fossilisation (ibid., p. 233-236). Thus, she concluded that a Swiss English variety does not currently exist.

---

[11] The Swiss learner corpus SWIKO directed by Thomas Studer is a small multilingual learner corpus which contains some English. However, the limited publications connected with the corpus are not available in English.

Both Rosenberger and Dröschel used an extended version of the Swiss English Database which was 71% spoken and 29% written. The spoken data included interviews, panel discussions, conferences, general meetings and English classroom lesson recordings and totalled 118,138 words. The written data was compiled from the same emails which Durham used and extended slightly to total 45,330 words.

Rosenberger investigated non-native features and the possibility of Swiss emerging characteristics. Although he found an overuse of the infinitive instead of the gerund, use of *would* as an adverbial, and non-native adverbial placement, as Durham, he found no evidence of a Swiss variety of English (Rosenberger, 2009, p. 211). Dröschel also concluded that evidence of a Swiss English variety does not exist. However, her research approach brought forth eight features to be considered as potential Swiss English characteristics: non-native use of articles, non-native plural marking, reclassification of non-count nouns, non-native use of third person singular -s, adjuncts of backward span, non-native formulation of conditionals, overuse of the *to* infinitive and non-native placement of adverbials. To date, no further corpus studies have been announced by the Swiss National Science Foundation.

To conclude, the use of corpora to study language is many faceted and offers valuable empirical data to advance the understanding of language much further than native speaker instincts ever could. As the digitalisation of capturing spoken and written language evolves, the future of corpus linguistics will surely follow suit with new methodology and research tools. In the next section, standards are explored.

2.3   The standards

English is undoubtedly a world language with the majority of its speakers being bi- or multilingual (Crystal, 1995, p. 106). The number of non-native English speakers has been increasing since Kachru's three circle model was proposed in 1985 (Kachru, 1985), from an estimated 700 million to over two billion non-native users in 2008 (Crystal, 2008). As the use of English has increased worldwide, varietal differences have also increased with the term to describe the language changing from English to Englishes. This denotes the fact that there are many varieties of English. An encompassing definition of variety of language is:

> …any body of human speech patterns which is sufficiently homogeneous to
> be analyzed by available techniques of synchronic description and which has
> a sufficiently large repertory of elements and their arrangements or processes
> with broad enough semantic scope to function in all formal contexts of
> communication. (Ferguson, 1972, p. 30, cited in Wardhaugh, 2006, p. 25)

This implies that a variety is co-defined by the body of speakers who produce the speech patterns and the society which determines the language use in all formal communication. A more generalist or simplified viewpoint is offered by Seidlhofer (2009, p. 43) who defines a variety as a distinct system which can be recognized, identified and verified and then Hudson (1996, p. 22) with a variety defined as 'a set of linguistic items with similar distribution'. These descriptions facilitate disregarding the opinion that 'Traditionally, a variety is the type of language spoken by a precise speech community… and geographical area' (Cogo, 2012, p. 98) and opens the door for broader definitions.

Once a variety has been recognised, it can become standardized. Working towards a plausible definition Crystal lists five 'essential characteristics' of Standard English: the variety does not include local features, linguistic features include grammar, vocabulary and orthography but not pronunciation, the variety carries the highest prestige, it is institutionalized and used in formal education and government and it is understood by all, but not necessarily used by all (Crystal, 1995, p. 110). With the recognition of standard varieties, follows the recognition of non-standard varieties. Britain (2010) explores grammatical variance in spoken English and suggests that Trudgill's (1974) 'guesstimate' that 88% of the population in England use non-standard dialects could be accurate. Interestingly, non-standard use of English is commonly categorised as dialects when used by native speakers and non-standard varieties when used by speakers with another first language. This suggests an elitist attitude toward native speakers which statistically no longer represents the use of English in the world.

There have been many attempts to devise a name for a type of English which could serve as a Central Standard English; World Standard English (McArthur, 1987), International English (Crystal, 1997), Global English (Modiano, 1999; Toolan, 1997) with currently no consensus within the linguistic community. McArthur's (1987, p. 10) idea of a common core where a standard grammar, vocabulary and orthography could be used when communicating internationally or as a lingua franca and a more local variety used within community boundaries remains a futuristic vision. There have also been attempts to standardise simplified versions of English such as Ogden's Basic English (Ogden, 1930) or Quirk's Nuclear English (Quirk, 1985) to no avail. Most would agree with Quirk that 'we must bear a certain responsibility' to uphold English standards (Quirk, 1990, p. 10). This must be balanced with the concept that ownership of a language belongs to its speakers (Widdowson, 1994). However, if we are to communicate

across borders and language communities all parties involved, native and non-native speakers, will have to compromise to achieve intelligibility.

2.3.1   Spoken grammar

The knowledge gained through spoken corpus research has furthered the understanding of spoken grammar. In 1995 Carter and McCarthy 'offered evidence from spoken data that everyday conversations manifested common grammatical phenomena that were marginalised in description and neglected in pedagogy' (Carter and McCarthy, 1995; 2017, p. 1). This was initially done by highlighting four grammatical features of spoken grammar based on corpus investigations. These features which would consequently be considered non-standard in written discourse are **Ellipsis** (e.g., missing pronoun) (I) '*didn't know* …', **Left dislocation** (added information) '*a friend of mine, she's got…*', **Reinforcement the tail slot** (added information) '*Good winter wine that*' and **Indirect speech** '*I was saying*' (ibid., 1995 p. 145-152). From the beginning, Carter and McCarthy advocated exposing learners to natural spoken data to increase comprehension through observation. Mumford (2009) suggests not only exposing learners to natural spoken data but teaching the following forms he considers most useful: phrasal chains; simple sentence structure; non-canonical use of some singular and countable/uncountable forms; 'ellipsis of subjects and auxiliaries; use of declaratives as questions; flexible word order, including headers and tails and fronting of objects; use of lexical chunks, fillers, and placeholders' (Mumford, 2009, p. 139). The listed features are not exhaustive, see Cambridge Grammar of English (Carter and McCarthy, 2006) and Longman Grammar of Spoken and Written English (Biber et al., 2000) for complete coverage. Understanding features of spoken grammar can ultimately aid in deciphering spoken data and developing criteria for learner corpora error analysis.

The concept of perceiving spoken grammar as being the same or different from written grammar models and its consequences are discussed in Leech (2000). He explains that the 'Nottingham School' (work of Carter, Hughes, and McCarthy) approaches spoken grammar as if it were singular in its own right without taking into account written models and if similarities appear all the better, but they should not be expected (Leech, 2000, p. 689). By contrast, Leech et. al. (work of Biber, Johansson, Leech, Conrad, and Finegan) follow a framework in the tradition of Quirk which assumes that spoken and written grammars are similar with variable frequencies of grammar use. In the end, both approaches have brought forth evidence that there are differences in spoken and written grammar that cannot and should not be ignored such as the use of simplistic phraseology and 'frequent reliance on nonclausal fragments' (Leech, 2000, p. 715). This knowledge is valuable in today's English classrooms where communicative competence is often in the foreground.

2.3.2   Proficiency standards

In Swiss primary schools, an implicit and communicative approach is used in foreign language classrooms with a focus on oral production (Pfenninger and Singleton, 2019, p. 215). However, in secondary school, an explicit approach which adheres more closely to native speaker norms is taken with proficiency evaluated according to the Common European Framework of Reference for Languages (CEFR). The Council of Europe created CEFR to identify and classify learner proficiency. Their goals are to:

- promote and facilitate co-operation among educational institutions in different countries;

- provide a sound basis for the mutual recognition of language qualifications;

- assist learners, teachers, course designers, examining bodies and educational administrators to situate and co-ordinate their efforts (Council of Europe, 2020, p. 28).

The common reference levels remain the backbone of the framework and consist of six levels as seen below. For a detailed overview of the levels on the global scale and qualitative aspects of spoken language use, see Appendix 1.

A Basic User	B Independent User	C Proficient User
A1 (Breakthrough) A2 (Waystage)	B1 (Threshold) B2 (Vantage)	C1 (Effective Operational Proficiency) C2 (Mastery)
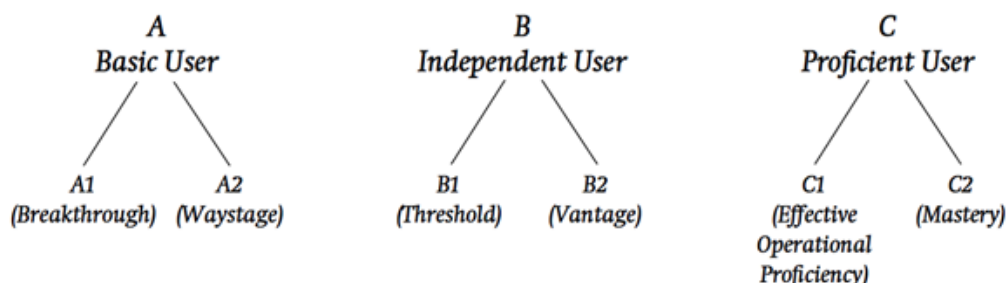
Figure 2-1 Common reference levels (Council of Europe, 2001, p. 23)

Switzerland has actively participated in the creation of the CEFR since 1971, first by hosting a symposium where the developmental work commenced and in 1991 a second symposium produced recommendations for the CEFR (Figueras et al., 2005). The common reference levels have influenced Swiss syllabus guidelines where recommended proficiency levels for each school year are given and all course material is clearly marked with CEFR levels. Furthermore, students at vocational schools are taught to use the reference levels to describe their language proficiency on their curriculum vitae when applying for employment. Traditional Swiss English exit exams for apprentice diplomas in the business field have been replaced by Cambridge ESOL exams. The standardization and close adherence to the reference levels are highly valued by the business world. In addition, proof of CEFR levels is required in most educational settings when the language of instruction differs from the first language or a certain competency level in foreign languages is required to obtain a federally recognized diploma. The use of the CEFR guidelines has been criticized. Firstly, for using can-do statements with a native speaker reference point. This controversial standpoint has been amended in the revised guidelines (Council of Europe, 2020). Furthermore, the lack of 'specific lexical and grammatical details' in the 'can do statements' which could lead to 'quantifiable linguistic descriptors' has been

questioned by Callies and Götz (2015, p. 2) and recommendations for adjustments made (Götz, 2015, p. 210). It is expected that the CEFR will continue to develop and perhaps the next revision will increasingly take into account corpus evidence.

2.3.3 English as a lingua franca

A discussion of English language standards would not be complete without mention of the English as a lingua franca (ELF) movement of the past 20 years. As Meierkord (2013) points out, linguists began to take interest and started to study English as a lingua franca in the 1970s and 80s although the terminology might have differed. An early UNESCO (1953, p. 46) definition identifies a lingua franca as 'a language which is used habitually by people whose mother tongues are different in order to facilitate communication between them'. Interest in ELF research was rekindled first with Jenkin's empirical study of phonology and 'lingua franca core' (Jenkins, 2000, p. 134) and then with Seidlhofer's call for a new paradigm to investigate English as a lingua franca in its own right. Initial investigations while compiling the VOICE corpus suggested the following eight features of ELF.

- Dropping the third person present tense –s
- Confusing the relative pronouns who and which
- Omitting definite and indefinite articles where they are obligatory in ENL, and inserting them where they do not occur in ENL
- Failing to use correct forms in tag questions (e.g., isn't it? or no? instead of shouldn't they?)
- Inserting redundant prepositions, as in We have to study about…)
- Overusing certain verbs of high semantic generality, such as do, have, make, put, take
- Replacing infinitive-constructions with that-clauses, as in I want that
- Overdoing explicitness (e.g., black color rather than just black)

(Seidlhofer, 2004, p. 220)

Since Jenkins' first description of a lingua franca core, 20 years have passed and Mortensen claims that ELF is still in its 'infancy' with 'teething problems' (Mortensen, 2013, p. 26). Besides a lack of conceptual clarity (ibid., 2013) a consensus as to a definition of who is considered an ELF speaker has not been reached as can be seen in Table 2.1 below which highlights some quotes made over the years.

Table 2.1 Various definitions of ELF

Firth (Firth, 1996, p. 240)	'[ELF is] a "contact language" between persons who share neither a common native tongue nor a common (national) culture, and for whom English is the chosen foreign language of communication.'
House (House, 1999, p. 74)	'ELF interactions are defined as interactions between members of two or more different linguacultures in English, for none of whom English is the mother tongue-'
House (House, 2001, p. 2)	'English as a lingua franca is nothing more than a useful tool: it is a "language for communication", a medium that is given substance with the different national, regional, local and individual cultural identities its speakers bring to it. English itself does not carry such identities; it is not a "language for identification".'
Kirkpatrick (Kirkpatrick, 2007, p. 155)	'…it is used as a medium of communication by people who do not speak the same first language.'
Seidlhofer (Seidlhofer, 2011, p. 7)	'English as a lingua franca (ELF) can be thought of as "any use of English among speakers of different first languages for

	whom English is the communicative medium of choice, and often the only option."
Jenkins (Jenkins, 2007, p. 3)	Jenkins (2007, p. 3) states she is in line with Seidlhofer that 'ELF does not exclude native speakers of English'
Samarin (Samarin, 1987, p. 371)	'any lingual medium of communication between people of different mother tongues, for whom it is a second language'.

As the table shows, some believe ELF excludes native speakers and some believe they should be included. Without the basic agreement of who belongs to the focus group, progress might be impeded. From the onset, the ELF paradigm has questioned the necessity of standardized native-speaker norms and has called for considering an alternative model. To date, no such model has emerged.

2.4   The learner

2.4.1   Role of user/learner

With the increased use of English comes the question of L2 identity and the implication it could have on spoken outcome. The following section explores the meaning and implications of user and learner status.

There is great variation in the definition of the terms 'language user' and 'learner'. For example, Firth and Wagner (1989) have a generalist view that language acquisition is an ongoing task where L1 and L2 users never cease to be learners (Firth, 2009). This view could lead to the assumption that becoming an expert user is an unattainable goal or that learner status is everlasting. Cook, on the other hand, defines the term L2 user as 'a person who knows and uses a second language at any level' (Cook, 2002, p. 4). The purpose of learning, therefore, is to use

the language in real life. In contrast, Cook defines learners as people who are learning a language but do not have a present need to use the L2 (Cook, 2002, p. 3). Taking the multilingual construct of Switzerland into consideration where foreign language acquisition is multi-dimensional, adopting Cook's (2002) definitions of users and learners allows us to differentiate between L2 users who learn English to use and L2 learners who learn English to fulfil current scholastic requirements or for later use. Cook (2008) observes user groups, five L2 user groups with examples are summarized below.

Table 2.2 L2 user groups based on (Cook, 2008, pp. 202-204)

Group 1	people using an L2 within a larger community- for example, Swiss students in London speaking English
Group 2	people using an L2 internationally for specific functions- for example, international communications in academics, religion or business
Group 3	people using an L2 globally for a wide range of functions- for example, English as a lingua franca
Group 4	people historically from a particular community (re-) acquiring its language as an L2- for example, to maintain or gain ethnic identity
Group 5	people using an L2 with spouses, siblings or friends

Association with one of these groups could also influence a speaker's view of their identity. Edwards (2009, p. 258) defines identity as 'self-definition by groups and individuals'. The concept of *self* in language learning is linked with motivation in Dornyei's 'L2 Motivational Self System' (Dörnyei, 2009, p. 9) where he identifies three components. The first is the 'ideal L2 self' which represents the person one strives to become. This component includes attitudes towards L2 speakers. A major potential for variance on self-expectations is the individual's conception of who the ideal speaker is. It is difficult to determine whether Swiss students gauge their ideal L2 speaker as their L1 Swiss German speaking teachers, internationally successful

Swiss stars like Roger Federer or native English speakers from music and films. Native English speaker diversity offers a wide range of language examples as ideal; from carefully scripted mainstream Hollywood movies to vernacular loaded gangster rap.

The second component is the 'ought-to L2 self' which includes the characteristics one's self would need to evade negative outcomes in difficult situations. Again, these could greatly vary depending on the interpretation of the ideal L2 speaker. Lastly, the third component is the 'L2 learning experience' which includes the environment where learning takes place and includes elements that would impact the experience. Although this system implies the L2 self as a learner, it can be argued that the same principles could apply to the L2 user who uses English to achieve tangible results in a sales conversation, a business deal or while on holiday.

The decision to perceive one's self as a user or learner of English can reflect the current attitude or situation of a speaker. Attitudes and self-evaluation are closely related to individual identity and are therefore perhaps not as easy to define with the viewpoint that language is learned in the classroom and used outside of it as Mauranen (2011, p. 158) implies. In contrast, Hyltenstam and Abrahamsson (2012, p. 182) argue that learning and using an L2 are different activities which occur alongside each other throughout L2 use and development. An example of L2 learning and using being done simultaneously is the use of English to gain knowledge or communicate on the internet. With this occurring more and more within scholastic environments, the boundaries of the classroom defining L2 learning or use could become obsolete.

The importance of understanding self-perception would not be complete without investigating the extrinsic factors of belonging to language communities. Edwards (2009, p. 27) suggests that

apart from our self-identity, social identities exist which affect our self-perception. The association of belonging to us or them can also influence individual language choices. Kirkpatrick (2007, pp. 10-12) explains these language choices as the 'identity-communication continuum'.

```
                          Language function
                         /       |         \
                        /        |          \
                       ↓         ↓           ↓
                   Identity              Communication
                      |                        |
                      ↓                        ↓
                  Language variety
                   /          \
                  ↓            ↓
broad / basilectal varieties        educated / acrolectal varieties / registers
```
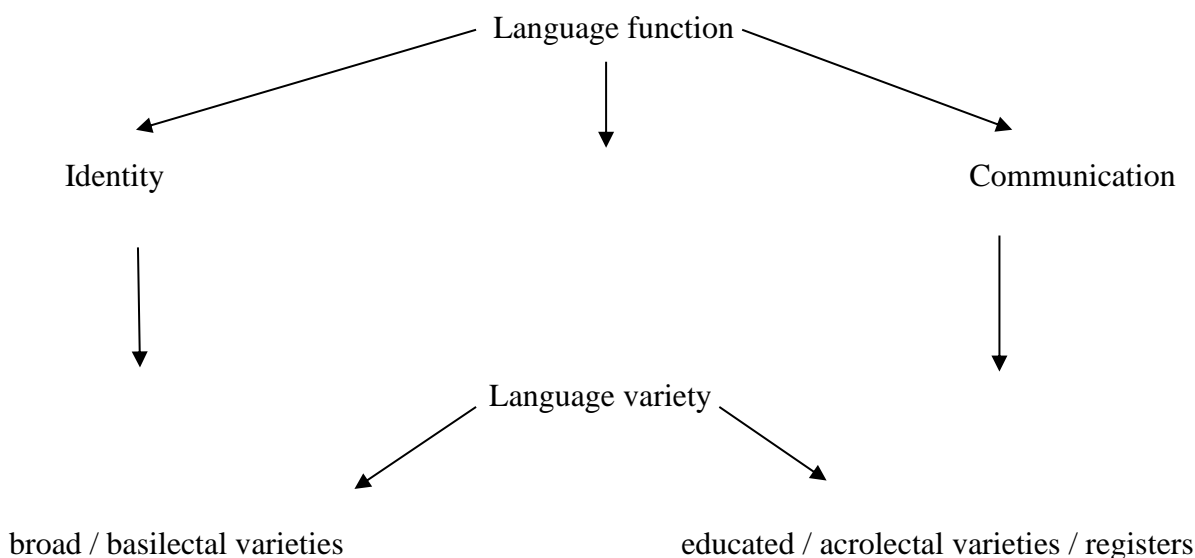
Figure 2-2 The identity-communication continuum (Kirkpatrick, 2007, p. 12)

This Figure illustrates how language choice could depend on the language function with broad varieties being used to signify group identity and educated varieties being used when communication is the main function. In the Swiss context, identity is not only defined by the use of the native language but the ability to communicate in a plurilingual environment at various levels of proficiency and formality while using different varieties. The social identities of the participants who contributed to this study are multiple and complex. The average participant's language choices and how they identify with a language community transpose between at least four language communities; Swiss German, High German, English and French. For example, some participants speak one or two non-Swiss German languages at home, Swiss German outside the home, High German at school and for formal written correspondence,

learns British Standard English at school, consumes American entertainment, uses social media channels in English as a lingua franca and learns business French to pass an exam. With such variance, it is quite plausible that belonging to possible social communities which might influence language identity choices is a constant fluid transition. Furthermore, the speaker might not even be aware of the boundaries of basilectal and acrolectal language varieties or be able to decipher standard from non-standard forms.

2.5   Acceptance of non-standard features

Documenting current acceptance of non-standard features of spoken language could aid in understanding present use and work as a benchmark for further studies. As early as 2009 Modiano suggested that some nativization processes of English were taking place in Europe (Modiano, 2009, p. 216). According to Bamgbose (1998, p. 2) innovations drive the nativization process and there is a need to separate innovations from errors. Thus, allowing new varieties to be recognised as such. Furthermore, acceptability along with demographic, geographical, authoritative and codification are deciding factors of an innovation's status (ibid. p. 3). With increased use of English as a lingua franca within Europe, it is uncertain if a broad European variety of English will emerge, many local varieties, or if Anglo-American standard varieties will prevail.

There have been many studies of acceptance from a native speaker standpoint. For example, the perception of accents (Neuliep and Speten-Hansen, 2013) and perceptions of non-native English (Lindemann, 2005). The acceptance of non-standard forms within a scholastic realm (Van der Walt and Van Rooy, 2002; Tan and Tan, 2008) is also well documented. However, little has been done in investigating non-native speakers' acceptability of non-standard usages, especially in Switzerland. Murray (2003), however, conducted a survey on Swiss teachers'

attitudes to the non-standard variety of Euro English with surprising results. She found that in general native speaker teachers had a higher rate of acceptability for non-standard English than their non-native peers. Furthermore, the acceptability of non-impeding elements ranged from 52-81%, with 'the car of my dentist' being accepted by the majority (81.2%), whereas rule breaking element acceptability ranged from 40-14% with 'that's the film who I saw' deemed acceptable by 13.9% (ibid. p. 157). The high acceptability rate of non-impeding non-standard usages suggests that Swiss teachers lean towards placing higher value on communicative understanding than total correctness. However, Murray's survey also indicates the standpoint that within the Swiss school system native-speaker norms were still the ultimate goal. As part of her investigation into the possibility of an evolving Euro-English, Mollin (2006) conducted a further acceptability survey which expanded on Murray's 2003 survey and will be explored in Chapter 6.

Turning to adolescent native speaker use of non-standard English, Brady (2015) explored the use of non-standard English in a 'working class' area of London, England where standardized norms are historically associated with higher social classes. Ninety percent of the fourteen- to fifteen-year-old students claimed they used non-standard English in their daily lives and could switch to a standard version when needed. The use of non-standard English was associated with a choice to express their identity. The term *choice* is often used in ELF literature when justifying the use of ELF among non-native speakers, see (Sewell, 2013). An unanswered question is if non-native speakers, like the native speakers mentioned in Brady's study, actually have two types of English in their repertoire to choose from or is their use of English situationally governed by chance and not by choice. The results of a survey conducted on Swiss acceptability of non-standard usages from the CSC will be thoroughly explored in Chapter 6.

To sum up, this chapter has introduced the concept of corpus linguistics by reviewing spoken corpus design before exemplifying various types of spoken corpora and the research findings resulting from their investigation. Then, the standardisation of language was explored by examining spoken grammar, proficiency standards and English as a lingua franca. Lastly, acceptability of non-standard English was addressed. The following chapter explains the methodology used to collect and decipher the spoken data investigated in this study.

Chapter 3 **METHODOLOGY**

3.1   Introduction

The subject group of 18–20-year-old Swiss participants was chosen because I argue that they have a heightened potential to alter how English is used within Switzerland and documenting current use can be used as a reference for further research. Murray suggests that alterations and slow changes within a speech community most likely occur when variations are used and accepted by the majority of that community (Murray, 2003, p. 160). Although numerous spoken corpora in other countries such as VOICE (2013), LINDSEI (Gilquin et al., 2010), MICASE (Simpson et al., 2002) are compiled of advanced or self-proclaimed competent users, the significance of researching expert users in Switzerland is questioned. Due to the structure of the Swiss Education System as described in Section 1.2.3, close to 70% of the Swiss workforce conclude formal language learning after completion of an apprenticeship or Matura exams between the ages of 18-20 with a Common European Framework of Reference for Languages (CEFR) English proficiency level between A2-B2. The majority of this segment of the population are not advanced users and will no longer be taught English but use it as a means of communication. The remainder enter higher education or take alternative paths.

To investigate the spoken English of this emerging workforce, the Central Swiss Corpus (CSC) was created. The corpus began as data collection for my master's dissertation (Oswald, 2010). Since then, 10 years have passed, and the CSC corpus has developed into a collection of spoken language that documents the voices of Swiss youth. During this time, the language of young adults from two distinct groups has been gathered. The first group or sub-corpus received foreign language education before the implementation of the current foreign language teaching policy which dictates early English in primary school, and the second group afterwards. Thus, the almost 100,000-word CSC corpus not only offers the opportunity to investigate general

spoken language but also explore the effects of policy change on language output through the comparison of its two sub-corpora.

The corpus is modelled after the LINDSEI-spoken learner corpus (Gilquin et al., 2010). During the preliminary phase of corpus development, the LINDSEI interview material was reviewed and adapted for this thesis. Although the author considered joining the LINDSEI project, it was felt that investigating the spoken language of the emerging workforce would be more representative of how English is used in Switzerland and therefore more significant. The LINDSEI project focused on advanced learners, namely university undergraduates majoring in English in their third or fourth year of university (Gilquin et al., 2010, p. 10). In relation, a mere 3,013 people were studying literature and linguistics in the seven universities in Switzerland in 2018.[12] This number includes all languages. The relatively low number of English majors and the fact that only 20% of the central Swiss population were preparing for tertiary education when this study was initiated, reinforced the decision to use the LINDSEI project as a template, but make appropriate adjustments to fit the Swiss situation.

3.2   Corpus description

The CSC corpus is a collection of spoken data which has been gained by means of interviews conducted and transcribed by the researcher. In total, ninety-six interviews were conducted to produce this unique corpus. To add depth to the analysis of the effects of internal and external factors on non-standard usages, a large number of variables were collected. The extensive gathering of this metadata is supported by Gablasova et al. as they emphasize the importance of taking into account metadata, especially when comparing corpora (Gablasova et al., 2017, p.

---

[12] Federal Statistic office online https://www.pxweb.bfs.admin.ch/pxweb/en/px-x-1502040100_131/px-x-1502040100_131/px-x-1502040100_131.px/?rxid=9a55bede-17b3-4f77-87db-52237fc92f74   accessed 13.06.2020

137), and is in line with Hunston's statement that 'A corpus can show nothing more than its content' (Hunston, 2002, p. 22). Lastly, it was felt that increased information about participants could contribute to the weight of any claims or conclusions drawn.

A brief summary of the variables follows. Detailed information will be given in Section 3.4. Before each interview, the participants filled out a participant profile (see Appendix 2) with the aid of the interviewer. This profile provided the majority of the information.

Age: Participants were chosen between the ages of 18-20. Therefore, representing the age that young Swiss adults enter the workforce or begin tertiary education.

Gender and **nationality**: gender and nationality representation reflect that of the population of the region being studied; Central Switzerland (Federal Statistical Office (FSO), 2018).

Language profile: The mother tongue of the participants, as well as that of both their parents, was noted. Moreover, languages spoken at home with their respective percentages were gathered. Additionally, English proficiency was established by evaluating a section of the interview that was the most comparable. Details will follow in Section 3.5.11. Knowledge of additional foreign languages was asked, along with a self-evaluation of the participants' CEFR competency level in each additional language. This extensive language profiling was deemed beneficial to better understand the effects of plurilingualism and English output in Swiss students for both Brohy (2001) and Haenni-Hoti et al. (2011) have found that children with multilingual backgrounds have an advantage over monolingual students in Swiss schools when learning a third language.

Type of education: The participants were either in or had just finished an apprenticeship or were taking/had completed general education courses which lead to a Swiss Matura. The occupations being studied in vocational education were noted. However, there was no attempt

to collect representative data from the close to 300 vocational career paths available to Swiss youth. Rather, participants were chosen due to availability and with the goal of reaching a sampling reflective of the central Swiss population. The language in which they were taught was also asked.

Years of English study: as mentioned above, as a result of policy changes, the participants had not all had the same type or amount of English instruction. The range was between 2 and 11 years of formal English instruction in a scholastic environment.

Stays in English speaking country: Almost every other participant (48%) had spent some time in an English-speaking country. The country, length of stay and year of stay were recorded. Previous research has found positive correlation between length of stay and fluency, particularly (Götz and Mukherjee, 2018).

Familiarity with interviewer: Although the interviewer was not currently or foreseen to be in a teacher-student relationship with any of the participants, in some cases she was employed at the same school or a former teacher of some of the participants. Therefore, the question of familiarity was asked to investigate any effects on spoken output.

To collect more detailed and comprehensive information, additional variables were collected in the form of five questions about English which were asked to each participant towards the end of the interviews. First, participants were asked which was more important to them; being understood or grammatical correctness. This was asked to determine if the awareness of grammatical importance had an influence on spoken output. Then, participants were asked to whom they thought they would speak English to in the future, native or non-native speakers. This was asked to gauge awareness of the current trend in Switzerland where English is being used more often than a national language to communicate with other Swiss nationals with dissimilar first language (Durham, 2016). Next, participants were asked if their English teachers

to date had been native or non-native English Speakers. In Switzerland, it is the norm in public education for Swiss teachers to teach foreign languages instead of hiring native speakers. So, if any participants had native English teachers the question if there were any noticeable differences in spoken output might be relevant. To determine if identity and self-perception were correlated with performance, two questions were asked. First, if they felt or acted differently when speaking English and if they considered themselves to be a learner or user of English or both.

Number and type of non-standard usages All of the above-mentioned variables were investigated to reveal their significance in relation to the spoken production of non-standard usages. Results of the analysis are reported in Chapter 4.

The term non-standard usage was used to signify a usage that is currently not perceived as standard. The decision to use this term was made to express the possibility of discovering specific language usage or patterns unique to Switzerland. With the initial use of mainly implicit foreign language learning in Swiss primary schools and the dexterity of Swiss German speakers to create new words, it was conceivable that a Swiss English might be emerging. Interestingly the Swiss German term for their dialect is *Mundart* or literally *mouth art*. Using the more common term in learner corpus studies of *errors* was considered too narrow and the term *non-native English* suggests native English superiority. For the most part, the non-standard usages investigated in this study refer to language accuracy. However, occasional disfluencies are included, even though a systematic annotation was outside the scope of the study.

3.3   Structure of CSC corpus

3.3.1   Number of interviews and word counts

In total 96 interviews were conducted. The following Table 3.1 illustrates the number of words and word types including the interviewer's words and transcription mark-up. Descriptive data on word token and word type frequency was obtained using AntConc 3.5.8 (Anthony, 2019).

Table 3.1 Number of interviews and total number of words and word types

Corpus	Number of interviews	Number of words including interviewer	Number of word types including interviewer	Percentage of interviewer words
Whole CSC corpus	96	187,942	4,594	47.6%
Sub-corpus 1	52	98,020	3,175	50.48%
Sub-corpus 2	44	89,922	3,016	44.46%

The object of investigation was the spoken English of the participants, therefore only their output was analysed. Table 3.2 below shows the number of participants' words and word types with all mark-ups removed. Here we see that the sub-corpora are similar but have distinct differences. Although sub-corpus 2 consists of fewer interviews, more words were produced. However, slightly fewer (-106) word types were documented. The greater number of participants in sub-corpus 1 describing different experiences in part 1 of the interview (see Section 3.4 below) could explain this. An additional preliminary indication of differences between the sub-corpora can be observed in the percentages of participant word counts versus the interviewer's. The participants in sub-corpus 2 spoke more than the interviewer compared to the participants in sub-corpus 1.

Table 3.2 Number of interviews and number of participant words in CSC corpus

Corpus	Number of interviews	Number of participant words	Average number of participant words	Number of participant word types	Percentage of participant words
Whole CSC corpus	96	98,489	1,026	4,236	52.40%
Sub-corpus 1	52	48,544	934	2,893	49.52%
Sub-corpus 2	44	49,945	1,135	2,787	55.54%

Included in the word counts above are 1,494 German words generated during 850 code-switching events which equates to 1.52% of the corpus. A further 14 French words generated during 11 code-switching events equalling 0.014% are also included.

3.3.2   Duration of interviews

The duration of the interviews totalled 18 hours and 39 minutes. Time was calculated from the first sentence of part 1 (see Section 3.4 below) and the end of answering the last question in part 3. Any discourse before or after the interview was not included in the transcripts nor calculated in the duration. Table 3.3 below shows the balance of interview duration between the sub-corpora.

Table 3.3 Duration of interviews

Corpus	Number of interviews	Total recorded minutes	Average duration of interviews
Whole CSC corpus	96	1,119	11.66
Sub-corpus 1	52	596	11.46
Sub-corpus 2	44	523	11.89

3.4   Method of interviews

Interviewees were sought by means of posters in schools and through teachers known to the author who allowed students to be voluntarily interviewed during or after class. Most interviews took place in a school setting and several at private homes where students were familiar with the setting.

The criteria for participating in the study was that the person should be between the age of 18-20, have either started learning English in the 3rd or 7th grade, have a minimum English proficiency level of A2 and currently live in Central Switzerland. It was assumed that the participants would have attended primary school in Central Switzerland, but this was not verified. One participant included in the study is known to have moved to Central Switzerland after primary school.

After a short introduction, the participants were given an information sheet and consent form in accordance with the ethical standards set forth and approved by the University of Birmingham (see Appendix 3). Afterwards, the participants filled out a participant profile form to collect the variable information necessary to make comparisons. During this pre-interview phase, the researcher conducted small talk while assisting the participants to complete the forms. This phase lasted approximately 5-8 minutes and was used to build rapport and ease any nervousness on the part of the participants. It was emphasised that the study was interested in what the participants had to say and that their opinions were valued. This approach appeared to work well, as the participants opened up and for the most part spoke quite freely. A small dictation device was then placed on the table and the interviews begun.

The interviews were divided into three parts. In part one, the participants were given the following instructions orally and were allowed to read them from a piece of paper on the table.

Table 3.4 Interview text part 1

Part 1
I'd like to interview you informally on things of interest in your life for fifteen minutes. To get the conversation started could you please choose one of the following topics and think

about what you are going to say. You should aim to be able to talk for 3-5 minutes. The conversation will then continue informally.

Topic 1:   An experience you have had which has taught you an important lesson. You should describe the experience and say what you have learnt from it.

Topic 2:   A country you have visited which has impressed you. Describe your visit and say why you found the country particularly impressive.

Topic 3:   A film/play you've seen which you thought was particularly good/bad. Describe the film/play and say why you thought it was good/bad.

Please don't take any notes as I would like it to be a spontaneous talk.

The three options are identical to those used in the LINDSEI corpus. They gave all participants the opportunity to find something spontaneous to talk about. At the same time, by narrowing the topics, the vocabulary was channelled to talk about experiences, travel or opinions about entertainment. By focusing the subjects, comparison of the two CSC sub-corpora and any LINDSEI sub-corpora was streamlined. After an initial monologue from the participant, the interviewer asked further questions to encourage the participant to elaborate on a subject or simply to continue the conversation. In general, part one was the longest part.

Part two of the interviews was the same picture description used in the LINDSEI interviews. The picture was printed on white paper, handed to the student and instructions below were read aloud.
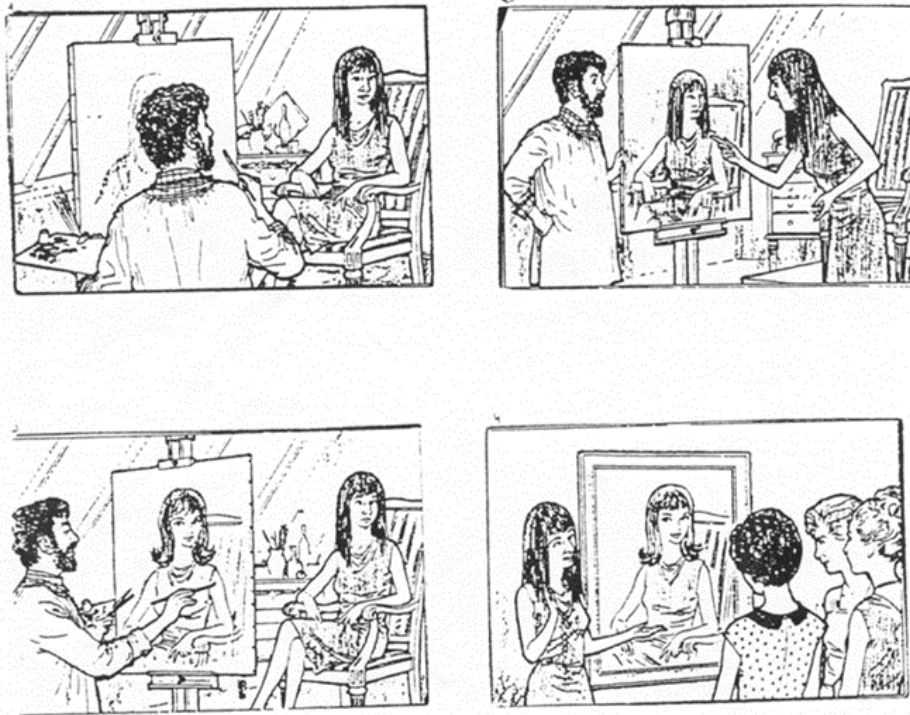
Figure 3-1 LINDSEI picture description, interview part 2

As can be seen above, the pictures enabled the participants to make up a story about the amusing scenes. At the same time, the pictures created many opportunities to highlight a participant's language abilities; from the simple vocabulary of *paint* or *draw* a *picture* or *portrait* to grammatical variance as well as more complex ideological aspects of self-awareness and worth. In addition, the picture descriptions were used to aid in determining the proficiency level of the participants because of this lexical variance and opportunity to showcase grammatical command.

Part three was designed to gather further information about the values and perceptions of each participant. It consisted of the following 5 questions about English in Table 3.5 below. These

types of questions were not part of the LINDSEI interview. Instead, LINDSEI interviews included general questions about life at university, hobbies and travelling abroad right after the monologue in part one.

Table 3.5 Interview part 3

Part 3
1. When you use English, what is more important; grammatical correctness or being understood?
2. Who do you think you will speak English with after you have completed your studies or in the future, native or non-native speakers?
3. Have your English teachers been native or non-native English speakers?
4. Do you feel or act differently when you are speaking English?
5. Do you consider yourself a learner or user of English or both?

Question one was asked to understand the perceived value of grammar. Although unintentional, the question became an additional source to determine the language proficiency of the participants due to the question's use of grammatical as an adjective. Over 40% of participants were not able to form a response with a standard use of the base word *grammar*. Due to the high frequency of non-standard usage this aspect was also tested for acceptability (see Chapter 6 for results).

Table 3.6 Standard and non-standard usages of the base word grammar

Word	Standard usage	Non-standard usage
grammar	20	0
grammatic(s)	0	48
grammatical	20	10
grammatically	19	2

For many participants, the questions in part 3 prompted reflections on their English use which they had seldom been asked to contemplate before. The participants seemed to enjoy voicing their opinions and even thanked me for interviewing them.

3.5   Profiles of participants

In this section, detailed statistical information will be given about the participants in the study. As mentioned above the objective was to find participants that reflected the population of Central Switzerland in order to be able to draw conclusions about the spoken language of this region's emerging workforce. All data was run through SPSS 25[13] to calculate frequency and percentages.

3.5.1   Gender

The CSC corpus represents the exact Figures for Central Switzerland (Federal Statistical Office (FSO), 2018) with 51% of the participants male and 49% female. The sub-corpora differed slightly. Sub-corpus 1 has 54% female and 46% male participants and sub-corpus 2 has 43% female and 57% male.

3.5.2   Age

The age distribution can be seen in Table 3.7 below with the majority aged 18 and an average age of 18.56. There were no substantial differences between the sub-corpora. On average, the participants in sub-corpus 2 were slightly younger than those in sub-corpus 1.

---

[13] IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Table 3.7 Age of participants

Age of participants			
Age	Frequency	Percent	Average age
18	57	59.4	
19	22	22.9	18.56 years
20	17	17.7	

3.5.3 Nationality

The number of foreigners in Switzerland belonging to the permanent resident population has continually increased over the past 30 years to the present 25.1%. Central Switzerland on the contrary has only 18.11% foreign permanent residents. This is probably due to the many rural communities (Nguyen, 2017). As can be seen in Table 3.8, the participants had 14 different nationality combinations including Swiss. This is a reasonably representative sampling of the Central Switzerland population with 17.7% of the participants declaring foreign nationalities. There is a slight difference in the two sub-corpora with sub-corpus 1 having 80.7% Swiss and 19.3% foreigners and the sub-corpus 2 with 84% Swiss and 16% foreigners. The nationalities unique to sub-corpus 1 are marked below with light grey and the nationalities unique to sub-corpus 2 are marked with dark grey.

Table 3.8 Nationality of participants

Nationality					
Nationality	Frequency	Percent	Nationality	Frequency	Percent
Swiss	79	82.3	Iraqi	1	1.0
Turkish	3	3.1	Italian	1	1.0
Kosovar	2	2.1	Swiss & German	1	1.0
Swiss and Italian	2	2.1	Swiss & Dutch	1	1.0
Albanian	1	1.0	Swiss & Bulgarian	1	1.0
Bosnian	1	1.0	Swiss & Russian	1	1.0
British	1	1.0	Tibetan	1	1.0

3.5.4   Native languages

As can be seen in Table 3.9 below, the participants' native languages coincide with the nationalities given with a total of 13 different native languages reported.

Table 3.9 Native languages of participants

Native Languages					
Native Language	Frequency	Percent	Native Language	Frequency	Percent
Swiss German	75	78.1	Italian	1	1.0
Albanian	3	3.1	Kurdish	1	1.0
Turkish	3	3.1	Swiss GE & Arabic	1	1.0
High German	2	2.1	Swiss GE & Croatian	1	1.0
Croatian	2	2.1	Swiss GE & Dutch	1	1.0
Swiss GE & English	2	2.1	Swiss GE & Italian	1	1.0
Bosnian	1	1.0	Tibetan	1	1.0
English	1	1.0			

The participants' mothers had a total of 19 different mother tongues as can be seen below in Table 3.10. This reflects Swiss national statistics (Flaugergues, 2016, p. 10) which report that permanent residents with first- and second-generation immigration backgrounds are twice as likely to use two or more languages more than once a week than those without an immigration background.

Table 3.10 Mother's mother tongue

Mother's mother tongue					
Mother tongue	Frequency	Percent	Mother tongue	Frequency	Percent
Swiss German	64	66.7	Bosnian	1	1.0
Albanian	5	5.2	Bulgarian	1	1.0
Turkish	4	4.2	English & French	1	1.0
Croatian	3	3.1	Kurdish	1	1.0

English	3	3.1	Romanian	1	1.0
German	3	3.1	Russian	1	1.0
Italian	2	2.1	Serbian	1	1.0
Romansh	2	2.1	Spanish	1	1.0
Arabic & Tigrinya	1	1.0	Tibetan	1	1.0

3.5.5 Father's mother tongue

The participants reported their fathers as having the following 18 mother tongues. The high number of mother tongues other than the four Swiss national languages suggests that 5 to 15% of the participants who are Swiss could have an immigration background as well.

Table 3.11 Father's mother tongue

Father's mother tongue					
Mother tongue	Frequency	Percent	Mother tongue	Frequency	Percent
Swiss German	59	61.5	Bosnian	1	1
Albanian	6	6.3	Dutch	1	1
Italian	5	5.2	English & Welsh	1	1
Turkish	5	5.2	Romansh	1	1
Croatian	4	4.2	Serbian	1	1
German	3	3.1	Spanish & Italian	1	1
English	2	2.1	Swiss GE & French	1	1
Kurdish	2	2.1	Swiss GE & Italian	1	1
Arabic & Tigrinya	1	1	Tibetan	1	1

3.5.6 Languages spoken at home

A total of 19 languages were reported to be spoken at home. Both Russian and Tibetan which are both native languages of two of the participants' mothers were not reported as being languages spoken at home. In both cases, High German was given as the sole language spoken at home. Possible implications of this language disparity will be suggested in Chapter 4. Table

3.12 below illustrates the plurilingualism of the participants with 38.6% speaking languages at home that are not the language of the local community (Swiss or High German).

Table 3.12 Languages spoken at home

Languages spoken at home		
Languages	Frequency	Percent
Swiss German	56	58.3
Swiss German + 1 language	18	18.8
Swiss German + 2 languages	3	3.1
High German	3	3.1
High German + 1 language	10	10.4
High German + 2 languages	2	2.1
One language besides Swiss or High German	4	4.2

The above reported information about the participants' language profiles loosely corresponds with the Statistical data on the main languages spoken by Swiss residents[14] seen below.
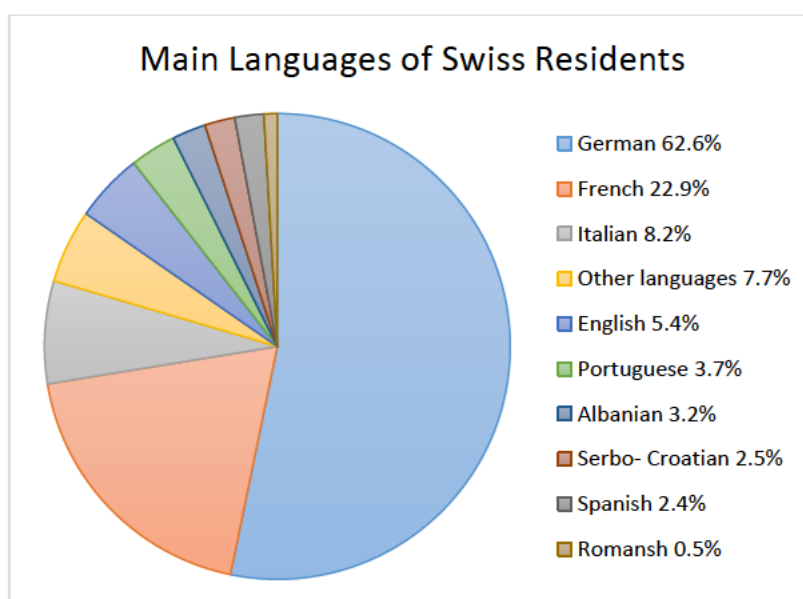
Figure 3-2 Main Languages of Swiss Residents

[14] Resident permanent population aged 15 or more and living in a private household. Several language choices possible Source: www.bfs.admin.ch Accessed 30.12.2019

3.5.7   Current studies

As stated in Section 1.3.3, at present the majority (over 60%) of Swiss youth choose to learn a profession by doing a vocational apprenticeship, while approximately 20% prepare for a Matura which leads directly to university studies. The remaining do transitional or general education paths (Bula and Segura, 2019). The participants in this study were chosen who were either doing an apprenticeship or studying for a Matura. At the time this study began other educational opportunities were less common as can be seen in Table 3.13. Therefore, this study of these participants deviates slightly from the current population.

Table 3.13 Completed education level of 20-year-olds in 2009

Completed level of education	Total	Men	Women
Apprenticeship	69%	74%	60%
Matura	20%	16%	23%
Obligatory school only (to 9th grade)	11%	7%	17%

(Translated from Bundesamt für Statistik 2010)

The implementation of the Bologna Process in Switzerland brought with it changes which have influenced the increase of tertiary education. One example is the reform concerning teacher education which made obtaining a bachelor to teach primary school obligatory. This has contributed to the significant increase in the tertiary education of Swiss residents ages 25-64 to rise from 10.5% in 2000 to 30.1% in 2021[15].

The distribution of participants' current studies when the interviews took place is illustrated in Table 3.14 below. It must be noted that there are differences in the compilation of the two sub-

---

[15]Source:https://www.bfs.admin.ch/bfs/de/home/statistiken/bildungwissenschaft/bildungsstand.assetdetail.16324 584.html accessed 10.05.2021

corpora. Sub-corpus 2 has less than half as many Matura participants as sub-corpus 1. This difference is partially compensated when interview length and spoken output are measured. The Matura students in sub-corpus 1 represent 29% of spoken words in sub-corpus 1 and 21% of spoken words in sub-corpus 2.

Table 3.14 Current studies of participants

Current studies Whole CSC					
Studies		Frequency			Percent
Apprenticeship		79			82.3
Matura		17			17.7
Current studies Sub-corpus 1			**Current studies Sub-corpus 2**		
Studies	Frequency	Percent	Studies	Frequency	Percent
Apprenticeship	40	**76.9**	Apprenticeship	39	**88.6**
Matura	12	**23.1**	Matura	5	**11.4**

3.5.8 Language of instruction at school

All the participants besides six reported being taught in a German speaking setting. Four reported being taught in German and English, one in German and French and the last in Arabic and German. The participants who were taught in languages other than German were divided equally between the sub-corpora.

3.5.9 French proficiency

Assuming the participants have completed primary school in Central Switzerland, all have been taught French as a foreign language from the fifth grade onwards. For the participants in sub-corpus 1 it was their first foreign language, whereas for the participants in sub-corpus 2, French was their second foreign language preceded by English taught in the third grade. The standardized learning objectives for both the first and second foreign language differ depending on the course of study. Foreign languages are rarely part of the curriculum for those who do an apprenticeship, Thus, for the majority of participants the minimum learning objective of

reaching CEFR level A2 at the end of obligatory school after grade nine was their last. Those studying for a Matura, on the other hand, continue studying at least two foreign languages until grade thirteen and are expected to reach a minimum of CEFR level B2. French proficiency declaration was a self-evaluation by the participants. The majority of the participants were quite familiar with the CEFR levels and could quickly assess their French ability. The interviewer assisted those that were unsure by asking questions about their French proficiency in line with the CEFR can do statements.

As can be seen in Table 3.15 below, there are major differences between the two sub-corpora. More than twice as many participants from sub-corpus 1 who had French before English claimed to have no French ability. However, the minimum objective of A2 was not met by 45.45% of the participants in sub-corpus 2 compared to 32.69% from sub-corpus 1. Furthermore, the increased number of Matura students in sub-corpus 1 was not reflected in the percentage of participants with a B2 level of French, but an increased number of B1 levels were reached instead. Investigating the proficiency levels of the Matura students, it was found that all 5 Matura participants in sub-corpus 2 reached the minimum B2 level, whereas the Matura students were divided almost evenly between A2, B1 and B2 levels in sub-corpus 1. From these statistics, it can only be speculated that learning English first might facilitate foreign language learning due to the recorded decrease of no ability in sub-corpus 2.

Table 3.15 French CEFR levels

French CEFR levels

CEFR level	Percent of sub-corpus 1	Percent of sub-corpus 2	CSC
No French	**25%**	**11.36%**	**18.75%**
French A1	**7.69%**	**34.09%**	**19.79%**
French A2	36.53%	31.82%	34.38%
French B1	21.15%	6.82%	14.58%
French B2	9.61%	15.91%	12.5%

The benefits of learning one language before the other is a topic of discussion in Switzerland where each canton can decide the order in which they are taught. Haenni-Hoti and Heinzmann have evaluated the situation in Central Switzerland and came to the conclusion that higher German and English competence in primary school have a positive correlation to French proficiency. They also found that children from multilingual homes had an advantage over monolingual children when learning French as a third or fourth language (Haenni-Hoti and Heinzmann, 2009; Heinzmann et al. 2010; Haenni-Hoti et al., 2011).

3.5.10 Years of English study

One of the main aims of this study was to discover the effects of early English on Spoken output. Although the participants were all residing in Central Switzerland, the autonomy of local school policies and variety of educational programmes can be seen in the range of how many years the participants have studied English at school. With early English starting in the third grade, a three to four-year range would be expected due to the age span, yet a six-to-eight-year span can be seen in Table 3.16 below. Hence, the four additional years of English expected between sub-corpora averaged only 3.4 years.

Table 3.16 Years of English studies

Years of English	Number of participants from Sub-corpus 1	Number of participants from Sub-corpus 2
2 years	1	
3 years	3	
4 years	3	
5 years	14	
6 years	21	4
7 years	9	2
8 years	0	3
9 years	1	18
10 years		15
11 years		2
Average years per person	5.6 years	9 years

3.5.11 English proficiency levels

English proficiency was established for each participant, as suggested by Carlsen (2012, p. 166), by evaluating the transcribed picture description part of each interview. The process was facilitated because each participant was given the same task and specialized vocabulary and sentence structure could be readily compared. Decisions were made considering range, accuracy, fluency, interaction, and coherence according to CEFR Table 3: Qualitative features of spoken language (Council of Europe, 2020, pp. 183-185). Further references including those from the Cambridge common mistakes series were consulted where necessary (Swan and Smith, 2001; Driscoll, 2005; 2007; Moore, 2005; Powell, 2005; Tayfoor, 2004).

The author is a trained University of Cambridge ESOL Examinations oral examiner in CEFR levels A2-C1 with 4 years' examination experience. Both herself and an additional English lecturer with over 20 years teaching experience which includes oral performance evaluation, individually rated all participants as being at a CEFR level A2, B1, B2, C1 or C2. Deviations were discussed and the final interrater reliability rate was 95.83% with a disagreement band of one level found with 4 participants. A trained University of Cambridge ESOL Examinations oral examiner in CEFR levels A2-C2 with 12 years' examination experience was commissioned as a third rater[16] to verify and assign final levels to the disputed 4 participants.

Carlsen (2012, p. 178) acknowledges that attention to proficiency-level assignment of learner corpus texts has been lacking in the research conducted to date. She suggests that a text-centred method where teachers' opinions, test scores and groups of raters is more reliable than learner-centred methods. Hulstijn (2010) also acknowledges the advantages of profiling proficiency levels but warns of the danger of circularity in establishing CEFR-related profiles when the

---

[16] Names and details of raters available upon request.

texts within a corpus are used for rating. He suggests avoiding the use of rating scales with reference to accuracy or linguistic forms (ibid p. 235). Although it is duly acknowledged that the danger of circularity exists, the decision to rate the proficiency levels using the corpus data was made because it was precisely the spoken sample of language that was collected during the controlled atmosphere that was being investigated for non-standard usages and not other samples of language where unknown circumstances could have existed. In addition, using two raters is similar to a University of Cambridge ESOL oral exam, especially with one of the raters having experience of having actually assigning marks in real exams. All Cambridge oral examiners must undergo yearly requalification which entails passing standardized rater assessment, thus a satisfactory degree of coherence was expected. It is believed that the additional use of a third rater to verify and decide on disputes added to the reliability. It is also suggested that 'a given learner's level of proficiency will tend to vary slightly from one day to the next, while the level of a given text will not.' (Carlsen, 2012, p. 168). By using the corpus data to determine proficiency, it can be argued that the language represents the proficiency level on the day of the interview.

Alternative text-centred methods of assessing proficiency levels include teachers' opinions and tests. However, access to the participants' language teachers was not available because at the time of the interviews most of the participants were not enrolled in language courses. Furthermore, the purpose of this study is to investigate spoken output so the use of a written exam to assess proficiency levels would not have been representative of spoken proficiency and a speaking test before or after the interviews would have minimized the desired collection of natural speech by establishing an exam atmosphere.

This approach of determining proficiency levels differs from other learner corpora. LINDSEI, for example, used a learner centred method that assumed that the university graduates would be advanced learners C1-C2. However, reviewing a random 10 % of the corpora revealed that the majority of the samples were in fact higher intermediate B2 and lower (Gilquin et al., 2010, p. 10). Knowing the CEFR level of each participant in the CSC corpus is advantageous because it adds greater depth to the analysis of non-standard usages.

An English CEFR level A2 is the minimum scholastic goal for Swiss ninth grade students in the formal learning objectives for students leaving obligatory school before doing an apprenticeship. For Matura students it is B2 (Bildungs- und Kulturdepartement des Kantons Luzern, 2016). An A2 level of English was the minimum accepted in order to be able to communicate during the interview. During the interviewing phase, two interviews had to be stopped because the participant was not able to converse in simple sentences. Thus, the CSC represents scholastically successful participants only. As can be seen below in Table 3.17, the majority have a B1 CEFR level with the remaining participants divided equally on either side and 7.3% achieving an expert level.

Table 3.17 CEFR levels of English in whole CSC

CEFR levels of English in whole CSC		
CEFR Level	Number	Percent
A2	25	26.04
B1	40	41.67
B2	24	25.00
C1	6	6.25
C2	1	1.04

The main difference between the two sub-corpora can be seen in the lower levels of A2 and B1 in the following Table 3.18. This suggests that the additional average of 3.4 years of English studies positively affected the participants' level of English.

Table 3.18 CEFR levels of English in sub-corpora

CEFR levels of English in Sub-corpora

no. of part. in sub-corpus 1	percentage of sub-corpus 1	CEFR Level	percentage of sub-corpus 2	no. of part. in sub-corpus 2
15	28.85%	A2	22.73%	10
19	36.54%	B1	47.73%	21
14	26.92%	B2	22.73%	10
3	5.77%	C1	6.82%	3
1	1.92%	C2	0	0

3.5.12 English Stays

One aspect of the affluent Swiss society is that travelling abroad on family holiday is not uncommon. In 2018 the Swiss average for all 8.6 million residents was 3.2 overnight stays per person per year with 67% of those stays spent abroad (Federal Statistical Office, 2019). The participants were asked to record how many weeks they had stayed in an English-Speaking country. Almost half or 47.9% had stayed in an English-speaking country from between one and eighty weeks. The amount and duration of stays was similar between the two sub-corpora. Most participants, n=29, reported staying in a European country (UK, Malta, Ireland), 24 reported staying in a North American country (USA, Canada), and 3 reported staying in Australia.

3.5.13 Familiarity with interviewer

Willing participants for this study were sought mainly in schools and institutions which the author was associated with in some way. As mentioned earlier, a dependent relationship with all participants was strongly avoided to ensure that the participants would speak freely without fear of being graded or evaluated. Despite this, the interviewer was familiar with 38.5% of the participants, vaguely familiar with 18.8% and unfamiliar with 42.7%. As can be seen in Table

3.19 below, the majority of participants in sub-corpus 1 knew the author and the majority of participants in sub-corpus did not know the author. Although great care was taken to conduct all interviews in the same fashion and with the same conditions, this needs to be taken into account.

Table 3.19 Familiarity with interviewer

	Familiar with interviewer		
	Total	sub-corpus 1	sub-corpus 2
Familiar	35.5%	61.5%	11.4%
Vaguely familiar	18.8%	23.1%	13.6%
Not familiar	42.7%	15.4%	75.0%

In addition, the method of obtaining spoken English through interviews is subject to observer's paradox (Labov, 1972a, pp. 209-210). It was predicted that the initial choice of topics in part one where participants tell of personal experiences would lessen the feeling of being observed as suggested by King and Horrocks (2010, p. 46). It must be noted that the corpus is practically void of swear words with only four instances and this does not reflect the everyday speech of Swiss 18- to 20-year-olds if they were speaking among themselves. However, it shows participants were able to use an appropriate genre when participating in an interview situation.

3.6 Transcriptions

The digital recordings were transferred to a personal computer and transcribed by the author, initially with SoundScriber[17] and later Dragon Professional[18]. Attempts to use the automatic transcription were discontinued because the pronunciation and use of code-switching between English and German made it only partially effective. Eventually, a combination of listening to

---

[17] SoundSriber was developed by Eric Breck while he was working for the University of Michigan on the MICASE project. It is available as freeware at http://micase.elicorpora.info/researchers/micase-statistics-and-transcription-conventions/our-transcription-tool-soundscriber.
[18] Nuance Dragon Professional version 15.30.000.006

the recordings in PotPlayer[19] with a 5-second reverse tab and speaking into Dragon resulted in more precise transcriptions. After the initial transcriptions, each interview was printed out and listened to on at least two more separate occasions to correct any inconsistencies before being coded to increase intra-rater reliability.

3.6.1   Annotation

A four-stage approach was used to annotate the transcriptions. The first stage involved assuring that the transcriptions reflected the spoken output of the participants. As mentioned above, utmost care was taken to assure all spoken output was transcribed as spoken by listening to and refining each transcription at least three times before non-standard uses (NSU) were coded.

Stage two: a list of possible NSU was created beginning with the potential eight ELF features named by Seidlhofer (2004, p. 220) and expanded during transcription. Each transcript was then initially coded. During that process, additional NSU features were discovered and NSU categories defined. For example, decisions were made on the definitions of the two broader categories of NSU-16 lexical choice and NSU-20 word order problems, redundant or omitted words. It was decided to keep those NSU categories broader to collect all examples which could then be investigated further during analysis.

Stage three: After 25 NSU categories were definitely defined, each transcript was recoded to include all NSU categories. This process was repeated after a two-month break, thus all transcripts were read and refined at least six times, three times during transcription and three times during coding.

---

[19] http://potplayer.daum.net/

Stage four: an independent researcher checked the coded transcripts for consistency and submitted the report in Table 3.20 below. Thus, it is assumed that the result of the four-stage annotation approach was increased intra-rater reliability.

Table 3.20 Report of random check for coding consistency

I hereby confirm that I have undertaken checks and readings of the transcripts collected by Susanne Oswald. The checks were done both randomly and at different intervals with the aim of highlighting consistency in the non-standard usages of English and the corresponding coding allocation.

The following areas and transcripts were checked in this order and consistency was found throughout:

NSU 1 Present Tenses, third person singular: 7, 14, 30, 143, 145, 101, 9, 23

NSU 10 Present and past tense inversion: 6, 12, 18, 137, 11, 21, 10, 15

NSU 15 Plural nous s – omission or insertion: 1,21, 8, 14, 127, 36, 134, 42

NSU 20 Word Order redundant or omitted words: 39, 58, 108, 63, 27, 4, 19, 55

NSU 16 Lexical choice (16.1) Influence of German language: 54, 35, 26, 58, 15, 22, 30, 21

Isabelle Kobau

English Lecturer

SHL Schweizerische Hotelfachschule Luzern, Switzerland.

3.6.2 Transcription and coding

Transcription guidelines for the CSC were adapted from the LINDSEI guidelines (Gilquin et al., 2010). Differences included the exclusion of marking phonetic features and length of pauses shorter than 3 seconds. These features were not a focus of this study. The following section

describes the mark-up and coding of non-standard usages. Examples are given from the corpus where possible.

Interview identification

Each interview transcript was preceded with a code encased in angled brackets which signified the sub-corpus and interview number.

<file sc"1"> Sub-corpus one, interview one

<file sc"2" > Sub-corpus two, interview one hundred and one consequently each interview was closed in the following way: </file >

Task identification

The beginning and end of the three interview tasks were marked with angled brackets as follows:

<S1>, <S2>, or <S3> begin set topic task one, two or three

</S1>, </S2>, or </S3> end set topic task one, two or three

<P> begin picture description task

</P> end picture description task

<E> begin questions about English use

</E> end questions about English use

Speaker turns

Two speakers participated in each interview.

The interviewer was identified as A and the participant was identified as B. The speaker and interview number were identified at the beginning and end of each speaker's turn as follows:

<A1> begin interviewer's turn interview one

</A1> end interviewer's turn interview one

<B1> begin participant's turn interview one

</B1> end participant's turn interview one

Punctuation

No punctuation marks were used to indicate sentence or clause boundaries.

Capitalisation

Proper nouns were capitalised.

Empty pauses

Empty pauses which were longer than 5 seconds were marked as …

Filled pauses and backchanneling

Filled pauses and backchanneling were identified such as *uhm* and were enclosed in rounded

brackets to ease their exclusion if necessary: (uhm), (uh), (uh huh) (huh) (em)

Unidentifiable words

Unidentifiable words were marked with an X for each unidentifiable word and encased in

angled brackets.

<X> one unidentifiable word

<XX> two unidentifiable words

<XXX> three or more unidentifiable words

Truncated words

Truncated or partial words were indicated by:

<B1> <trun>th</trun> there were five people </B1>

Contracted words

All contracted words were retained. *I'm, we're, they'll, she's* etc.

Non-standard forms of words

Non-standard pronunciations such as *cos, gonna, gotta, kinda, wanna* and *yeah* were retained.

German/foreign words code switching   Due to German being the second common

language of the interviewer and participants, code switching often occurred. German words

were enclosed in angled brackets <G> deutsch </G>. Swiss German spelling was used to retain the intended pronunciation.

Four types of code switching were marked.

1.   One-word replacements which did not interrupt the flow of speech;

      She saw a <G> Strassenmaler=1 </G>

2.   German discourse marker –*also, auch, oder, doch, ja, vielleicht mal*, used in place of *so, maybe* and *but*; no interruption in flow of speech;

      then <G> also=2 </G> *he went*

3.   Questions asking for clarification or translation;

       <G>Achterbahn wie sagt man=3 </G>

4.   Longer phrase or complete thought;

      <G> also Sie hant eppis gelert drüiss=4 </G>

Abbreviations

Abbreviations pronounced as a sequence of letters were transcribed as capital letters

Dates and numbers

All Figures were written out in words.

Anonymization

Names of schools or people were replaced with <name of school>, <name of teacher>

Overlapping speech

Overlapping speech is marked as at turn

<A1 > (uh huh) </A1 >

Voice quality

Laughing, whispering, cough [L], [W] [C]

Multiple complete words

<MW> then then </MW>

Contextual comments

<door opens young woman says sorry and closes door>

Self-correction

<SC> him (er) her </SC>

Non-standard usages

Initially, the eight features of English as a lingua franca determined by Seidlhofer (2004, p. 220) were a focus of this study. As the research developed, it was determined that analysing a wider range of non-standard usages would be more beneficial. Thus, as the corpus was being transcribed, twenty-five non-standard features were identified. The features were followed by <NSU> and the number of the corresponding feature, then closed with corresponding brackets. Each non-standard feature is described and examples from the corpus are given below.

<NSU-1> Omission or insertion of the third person present tense –*s*

This includes all instances of 3rd person diversion from Standard English norms: includes was/were and have/has NS usage

she would	<NSU-1>goes </NSU-1>	to the bank and so
I think if you	<NSU-1>talks </NSU-1>	in such
yeah and he	<NSU-1>like </NSU-1>	to paint in his free time

<NSU-2> Confusing the relative pronouns *who* and *which*

the other people	<NSU-2> which </NSU-2>	are standing around
for a boat	<NSU-2> who </NSU-2>	drives

<NSU-3> Omitting definite and indefinite articles *a, an, the* where they are obligatory in Standard English <NSU-3> 0 </NSU-3> The *0* indicates an omitted word.

I think and it was	<NSU-3> 0 </NSU-3>	really hard time but (uh) now
I (uhm) work in	<NSU-3> 0 </NSU-3>	airport
and she can't go into	<NSU-3> 0 </NSU-3>	lift because she needs

<NSU-4> Inserting definite and indefinite articles *a, an, the* where they do not occur in Standard English

where she's a top model with	<NSU-4> a </NSU-4>	beautiful hair now and
live here in Switzerland and they speak	<NSU-4> the </NSU-4>	Swiss German

<NSU-5> Failing to use correct forms in tag questions (e.g., *isn't it?* or *no?* instead of shouldn't they?)

a printer is something on the computer	<NSU-5> wasn't it </NSU-5>	(uhm) and it's

<NSU-6> Inserting redundant prepositions, as in "*We have to study about…*"

but (uhm) I like to to call	<NSU-6> with </NSU-6>	people and not write (uh) an

<NSU-7> Overusing certain verbs of high semantic generality, such as *do, have, make, put, take.*

how should I say the <G> Bericht=1 </G> and	<NSU-7> make </NSU-7>	analysis and I was also

During stage two of the annotation process oversimplification or overuse of high semantic verbs was found to belong to the category of lexical choice NSU-16.

<NSU-8> Replacing infinitive-constructions with that-clauses, as in *I want that*

but (uhm) my parents said you have to do	<NSU-8> that </NSU-8>	then I had to

<NSU-9> Overdoing explicitness (e.g. *black colour* rather than just black)

No instances were found; thus, no example can be given.

Learner Language

The following could be considered learner language inconsistencies with Standard English norms.

<NSU-10> present simple / past tense inversion

we went back to New York and	<NSU-10> fly </NSU-10>	back to Switzerland
was just a big hole and you didn't	<NSU-10> knew </NSU-10>	where to go
chair and posing and waiting (uhm) the artist	<NSU-10> drawed </NSU-10>	a picture and

<NSU-11> future tense *will/would* instead of *want*

the young woman I think she (uhm)	<NSU-11> will </NSU-11>	a (uh) painting for

<NSU-12> verb to *be* insertion or omission

If the verb to be was used incorrectly, it was categorised under the corresponding non-standard category. For example, confusion with *he is/ he are* would be under <NSU-1>and confusion with past and present tense would be under <NSU-10>.

it was very funny and I	<NSU-12> am </NSU-12>	met a girl she was
it's helpful when you	<NSU-12> 0 </NSU-12>	in another country

<NSU-13> Omission and substitution of prepositions

other people looked	<NSU-13> 0 </NSU-13>	the picture and
and in the future I can only profit	<NSU-13> of </NSU-13>	this language

<NSU-14> Full yes used as a discourse marker

the weather was very nice also sunny	<NSU-14> yes </NSU-14>
I learned a lot of (uhm) how do you say homework	<NSU-14> yes </NSU-14>

The tagging of this non-standard observation was abandoned after the second recoding in phase three of the annotation process because of the realisation by the researcher that interpreting how natural the intonation of a *yes* at the end of a sentence was highly subjective.

<NSU-15> plural –s omitted or inserted

and we slept two	<NSU-15> night </NSU-15>	and it's very big place
here in Lucerne when I meet some	<NSU-15> peoples </NSU-15>	from I don't know India
important for them and if they have	<NSU-15> this </NSU-15>	things then they are happy

<NSU-16> lexical choice

This category is quite broad because it includes all vocabulary inconsistencies such as: *to be/have, make/have, so lot/so many, talk/speak,* and making a non-standard word choice or fantasy word. Further information will be given in the analysis in Chapter 5.

not perfect I have years in school but my	<NSU-16> grammatic </NSU-16>	is not so good
it looks better with (uh)	<NSU-16> kirby </NSU-16>	hair [L] and so
very nice (uh uh) station and (uh) so we	<NSU-16> make </NSU-16>	(uh uh) a lot of things

<NSU-17> present, past perfect aspect and past simple inversion

who of course could see that the painter had	<NSU-17> did </NSU-17>	a very bad job [L] yeah
tourists are (uh) are from other countries that I	<NSU-17> see </NSU-17>	(uh) in my lifetime

<NSU-18> redundant use of continuous forms and *-ing*

when I am drunk I	<NSU-18> am speaking </NSU-18>	much better English
and I'm not here	<NSU-18> for learning </NSU-18>	English

<NSU-19> omission of continuous forms and *-ing*

they are not so good in	<NSU-19> build </NSU-19>	cars I think
at first I also had problems	<NSU-19> to sit </NSU-19>	the whole day in school

<NSU-20> word order, redundant or omitted words <NSU-20> 0 </NSU-20>,

and I started when I was eight seven years	<NSU-20> 0 (old) </NSU-20>	and (uh) all year I had one
just Swedish and some people can	<NSU-20> too German speak </NSU-20>	
she liked it and showed showed it to her	<NSU-20> friends girls </NSU-20>	

<NSU-21> word order inversion of adverbials of frequency and focus

Such as, *had also* instead of *also had*, *just have* instead of *have just*, *have always*

| I already know he's not dead because he | <NSU-21> also have already </NSU-21> | a movie |
| my (uhm) cousin | <NSU-21> works also </NSU-21> | there and he told me |

<NSU-22> omission of will future

Such as: *I search/ I'll search*

| if I go to another (uh) country I | <NSU-22> 0 </NSU-22> | just go travel independently |
| oh okay [L] (uhm) I | <NSU-22> 0 </NSU-22> | describe the film Aviator |

<NSU-23> prepositional possessive phrase

Such as: *the brother of my mother instead of my mother's brother*

| yeah without cows (uh uh) | <NSU-23> the brother of my father's girlfriend </NSU-23> |
| yes I think it's | <NSU-23> the cousin from my dad </NSU-23> |

<NSU-24> negation inconsistencies

Such as: *I don't can*

| because when I | <NSU-24> don't can </NSU-24> | speak the language |
| then yes also a great history but | <NSU-24> I don't really interest in </NSU-24> | this |

<NSU-25> *the* or *it* instead of personal pronouns

| put this on | <NSU-25> the </NSU-25> | (uh) body and then we played in the rain |

3.7   Limitations

This study can give insights into a specific group of 18–20-year-olds in Central Switzerland and is an example within the Swiss context where many demographic and educational variables are unique to each municipality or region. Consequently, the results reflect only this small group of participants. Some developmental trends and generalisations can be suggested, however transferring them to all of Switzerland would be negligent. On the other hand, results here can document the current state of English use and suggest further channels to investigate.

A second limitation to the results is reflected in the corpus being compiled using interviews and not natural occurring spontaneous speech. As Hunston (2002, p. 2), reminds us, the word *corpus* is used to describe 'a collection of naturally occurring examples of language'. It can be argued that the spoken language collected during these interviews would not fall in the category of 'naturally occurring'. However, the participants were observed as being very open and candid about their beliefs, they talked about happy and sad times, death, love, family and relationships. I feel I was given a glimpse into their personal world through conducting the interviews that I would not have had in a less structured natural conversation.

3.8   Analysis procedure

For analysis of the data, once the compilation and coding had been completed, a descriptive statistical method was chosen which would measure raw and normalized frequencies to calculate percent, per hundred words and ratios. Means and medians were employed to measure central tendencies in the CSC, whereas standard deviations were used as a measure of dispersion.

All data was first collected in Microsoft Excel version 2011 and prepared for analysis with SPSS 25. Initial analysis produced frequency lists and thus strategy decisions for further investigation were made. A combination of SPSS, and Excel sorting and analytical features were used to make all calculations. The use of AntConc (Anthony, 2019) facilitated concordance searches. The Someya Lemma List (no hyphens)[20] was used in conjunction with AntConc to conduct lemma searches. The 1994 BNC spoken corpus was used as a native reference corpus where needed. In conclusion, all research was conducted solely by the author.

---

[20] available at: https://www.laurenceanthony.net/software/antconc/ last accessed 30.01.2021

3.9   Conclusion

This chapter has outlined the structure of the CSC corpus. The methodology involved in constructing the corpus was explained, and the wealth of metadata collected was presented. Furthermore, detailed information on how the transcriptions were devised and annotated was provided to showcase how the corpus will facilitate answering research questions one and two. The following chapter re-states those research questions before presenting and discussing the analysis and findings of the CSC.

Chapter 4 **CORPUS ANALYSIS AND FINDINGS**

4.1   Introduction

Chapters 4 and 5 are devoted to presenting the analysis of the CSC corpus. After the introduction and some preliminary analysis, in-depth analysis will be presented. The in-depth analysis is divided up into two main parts. First, the mainly quantitative relationship between the number of non-standard usages and demographic, identity and educational variables will be explored to determine their impact on the participants' spoken output in Chapter 4. This will aid in answering the first and second research questions:

RQ 1: What is the relationship between the use of non-standard English features and Speakers' demographic, identity-related and educational variables?

The second area of in-depth quantitative and qualitative analysis in Chapter 5 will deal with the grammar and lexis of non-standard usages and their occurrence in each of the two sub-corpora which are divided into participants who had early English and those who did not. Each of the 25 non-standard usages which were identified and coded will be explored. This second area of analysis will reveal answers to the second research question:

RQ 2: What effect does early English, taught with a productive focused curriculum, have on the number and type of non-standard features used?

The remaining research question about the acceptance of non-standard usage will be dealt with in Chapter 6.

4.1.1 Word frequencies

To obtain a preliminary overview of the CSC composition, word frequencies were compared with five spoken corpora: a learner corpus, a general native speaker corpus, and three teenager /young adult corpora. By observing word frequency, we can hypothesize how language is used by a group and observe/predict changes in use. Tagliamonte (2016, p. 30) states that 'frequency is a red flag for linguistic change'. This comparison highlights possible areas of change that are in progress.

The CSC was compared to the following corpora. First, the German sub-corpus of LINDSEI (Gilquin et al., 2010), which has a similar structure as the CSC and was used to compare learners with German L1. Then, the native speaker CANCODE five-million-word spoken corpus (O'Keeffe, McCarthy and Carter 2007, p. 35), represents general native speakers of English. Next, the one-million-word TTC Toronto Teen Corpus comprised of 90 Canadian speakers aged 9-20 (Tagliamonte, 2016, p. 12) represents a broadly similar age group of native North American native English speakers. The Bergen Corpus of London Teenage Language COLT (Stenström et al., 2002) represents London teenagers aged 13-17, and the Havering Multicultural London English Corpus MLE (Cheshire et al., 2007-2010) represents 17-year-olds with a multicultural language background from outer London.

As expected, the majority (n=29) of the most frequent fifty words occurred in each of the six corpora. The CANCODE corpus was used as a base to understand the divergences from a general native corpus and the distribution of the ten most frequent words in that corpus are illustrated in light blue in Table 4.1 below. We also see an abundance of backchanneling or filled pauses in both learner corpora with the use of *uhm*, *uh*, *erm* and *mhm* in their top ten. This suggests parallels with previous research (Gilquin, 2008, p. 128; Götz, 2013, p. 110).

Table 4.1 Spoken Corpora Top 50 word frequencies

	Learner Corpora		General Native speaker	Teenager/young adult corpora		
	CSC	LINDSEI-GE	CANCODE	TTC	COLT	MLE Havering
1	i	a	the	i	you	i
2	and	i	i	like	i	and
3	the	and	and	and	the	you
4	it	the	you	you	and	Sue[21]
5	uhm	erm	it	the	it	like
6	uh	you	to	it	a	yeah
7	s	it	a	to	to	the
8	a	to	yeah	s	yeah	to
9	yeah	mhm	that	a	that	a
10	to	s	of	that	what	it
11	in	yeah	in	so	no	that
12	was	that	was	yeah	in	in
13	yes	in	it's	of	know	was
14	t	er	know	was	he	Just
15	so	of	is	t	of	of
16	you	t	mm	in	it's	so
17	like	laughs	Erm	but	oh	but
18	that	was	but	they	is	my
19	but	laughing	so	know	like	they
20	think	so	they	he	on	do
21	we	they	on	then	do	mm
22	is	but	have	do	was	what
23	she	is	we	my	don't	it's
24	because	like	oh	we	got	he
25	with	have	no	um	have	got
26	of	or	like	Just	Just	is
27	don	think	well	have	so	then
28	there	there	what	what	but	no
29	have	do	do	really	she	Unclear (tag)
30	he	she	right	don't	they	don't
30	not	starts	just	is	not	me
32	then	stops	he	no	me	on
33	or	really	for	she	this	all
34	they	well	erm	there	my	know
35	my	what	this	or	well	we
36	also	yes	be	for	one	go
37	English	not	all	not	I'm	she
38	can	don	there	oh	go	there
39	very	for	got	on	up	with
40	for	because	that's	all	erm	cos

41	no	know	not	one	get	have
42	know	very	don't	go	your	oh
43	this	this	if	me	all	not
44	really	we	think	with	we	get
45	when	just	one	this	are	really
46	just	he	with	at	be	for
47	okay	with	at	m	with	or
48	go	are	or	think	for	er
49	people	at	then	be	that's	name
50	speak	about	she	are	can	your

The learner corpora CSC and LINDSEI-GE show some additional distinctive features which are marked in yellow. All six corpora have *yeah* in the top 50, however only the two learner corpora also contain *yes*. The use of *yes* as a discourse marker or substitute for the German *ja* has been well documented in (Baumgarten and House 2010a; House and Lévey-Tödter, 2010). In addition, the two learner corpora contain the conjunction *because* and adverb *very* suggesting their overuse (see Andreu-Anrés et al., 2010; Lorenz, 1998; Flowerdew, 2000; Mendikoetxea et al., 2010). Uniquely, the CSC is missing the words *do* and *what* which are both in the top 50 words of the other corpora. This lapse is possibly due to the nature of the corpus where the participants asked relatively few questions.

When comparing the three teen corpora, *like* is in the top ten of two out of the three with *so* following. In the other three corpora *like* follows *so* below the top 10 mark. These are marked in purple in the table above. Tagliamonte documents the increased use of *like* by teenagers today, although its use has been recognised as colloquial and even vulgar as early as 1933 (Tagliamonte, 2016, p. 9). She also recognises that young people's use of *like* in quoting is now widespread throughout the world (ibid. p. 78). The CSC corpus can verify this with 8 out of 10 searches for *I'm like* resulting in *like* being used as a quotative verb. See examples below:

I get a bit uhm like kind of nervous when I speak because	I'm like	'can they really understand me' but uh no I don't
the stuff and in really like three weeks or so and	I'm like	'look uh listen get your stuff together and write it'
they say no I doesn't and no no	I'm like	'look they have given you a book so use it'
prob problems cuz I don't know this word and then then	I'm like	'oh no I don't know this word how can I'
calling me like 'come <name> you need to help us' and	I'm like	'oh hi yes' their very happy when they see me
with an SMS okay it's finished and	I'm like	'okay it's yeah'
want to learn English a bit better so speak English with me and	I'm like	'uh [L] you're welcome'
do we have to do or when do we have to go there	I'm like	'yeah yeah' then and then

The use of *really* (marked in orange) in the top 50 of all corpora besides the CANCODE and COLT might be explained by the age of the data because *really* has been named one of the four most popular intensifiers of the 21st century by Bordet (2017, p. 3) with her data coming from nine years of an American sit-com 20 years after the CANCODE and COLT corpora commenced data collection.

Three more words which associate the CSC with the teen corpora are *my*, *go* and *can* and are marked in green in the table above. *My*, for example, was used 234 out of 653 times in conjunction with a person (e.g., my sister) and *go* was used 298 out of 400 times to describe travel so it can be speculated that the storytelling nature of the CSC and descriptions of holidays contributed their increased use when compared to the general native corpus. An alternative hypothesis would be that the CSC is in line with Tagliamonte's observation that the nature of storytelling has been changing over the past 50 years with teenagers at the forefront of that change. Although further investigation at this time is out of the scope of this work, it would be noteworthy.

In conclusion, preliminary analysis of word frequencies has shown that the CSC is similar to the LINDSEI corpus in structure and content with the increased use of backchanneling, the use of *yes*, *very* and *really* and unique in its less often use of *do* and *what*. It is similar to the CANCODE general corpus as it reflects the basic words of the language. Lastly, it is similar to

the teenager corpora in its increased use of *because (cos)*, *can*, *go* and *my*. The result of this preliminary overview reiterates the fact that the CSC is a learner corpus of young adults with some unique features. In the following section, the variables which exemplify its uniqueness will be addressed.

4.2   Analysis of variables

4.2.1   Demographic variables

This section will investigate the relationship between the use of non-standard usages (NSU) of English and demographic variables. These include gender, age, nationality, native languages of participants and their parents as well as languages spoken at home. These variables were self-declared by the participants. Data was initially entered into SPSS 25 and thereafter calculated in Microsoft Excel version 2011.

In total 2,642 non-standard usages were found in the Central Swiss Corpus (CSC), making the overall mean frequency of 2.96 NSU per hundred words (phw) (SD=1.62). The interviewer conducted all interviews with the focus of creating an environment which gave each participant equal opportunities to express themselves as suggested by King et al. (2019, pp. 71-92). The comparison of each participant's NSU per 100 words allows for a unit of evaluation which is comparable with other corpora. The distribution of NSU per hundred words can be seen below in Figure 4-1. Furthermore, the second research question on the effect of how being taught early English has on the number and type of non-standard usages will begin in this section and be carried on in the following sections.

Total number of NSU per 100 words

0 to 6.35 NSU per 100 words and participant average 2.96
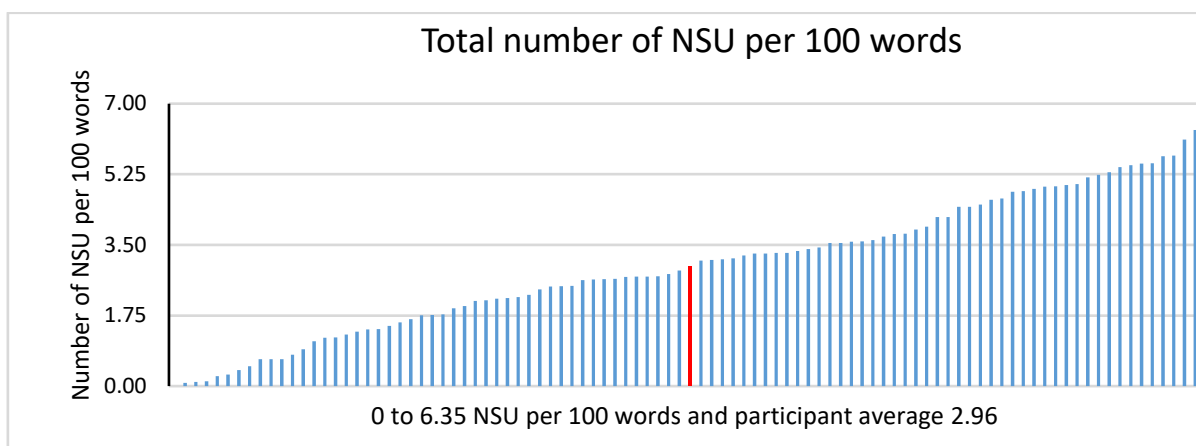
Figure 4-1 Total number of NSU per 100 words and participant

To begin, gender will be analysed in the following section.

4.2.1.1 Gender

In Switzerland, there is a general belief that the Swiss education system favours female students over males (Stadelmann, 2019) and that female students tend to be better in languages while male students tend to be better in maths. The 47 females were found to have made more NSU phw when the entire corpus was examined with 3.17 (SD=1.75) NSU phw compared to 2.76 (SD=1.47) NSU made by the 49 males.

When the sub-corpora were examined, the male participants were very similar with the 19 sub-corpus 1 males producing a phw mean of 2.78 (SD=1.16) NSU compared to the 25 sub-corpus 2 males with a phw mean of 2.72 (SD=1.75) NSU. However, the female participants showed greater differences. The 28 sub-corpus 1 females produced a phw mean of 3.75 (SD=1.64) NSU compared to the 19 female participants in sub-corpus 2 with a phw mean of 2.32 (SD=1.57) NSU. This equals a difference of 1.43 NSU phw between sub-corpora.

100

This difference in performance could reflect the Swiss education system and the time at which English is initially introduced in schools and could strengthen the argument of 'the younger the better' for foreign language acquisition. However, this would not explain why males in both sub-corpora performed similarly and females did not. On the issue of the best age to commence learning a foreign language, (Singleton, 2018, p. 56) concludes that by the end of secondary school early or later learners' performance is indistinguishable, whereas (Meyer, 2018, p. 78) argues that numerous variables and quality of teaching are more influential than age when instruction begins. Watts (2018, p. 62) even claims that no research validates 'the earlier the better' as the premise of a scientific argument, see (Pfenninger and Watts, 2019; Pichard, 2018; Lüdi, 2018; and Muñoz, 2014) for further discussion. The analysis of the combination of variables which affected performance of participants in this study will attempt to showcase tendencies to determine directions for further study.

4.2.1.2 Age

When the whole corpus was analysed, no significant correlation was found between age and number of NSU phw, (see Table 4.2 below). This was expected due to the minimal age span of 3 years. However, it is interesting to see that the youngest group used the lowest number of NSU phw as the younger participants might be expected to use more NSU due to less experience.

Table 4.2 Age and number of total NSU phw in CSC

Age	Number of participants	Mean NSU phw	Pearson correlation coefficient
18	57	2.86 (SD=1.50)	R value = 0.0734
19	22	3.10 (SD=1.67)	P value = .477265 Result is not significant at p < .05
20	17	3.13 (SD=2.00)	

Although investigating the sub-corpora revealed a slightly different picture in terms of the numbers of NSU phw within the sub-corpora, there remains no significant correlation between age and NSU phw. Here we can clearly see in Table 4.3 that in sub-corpus 1 there is some difference in the ages and amount of NSU phw with the 20-year-olds having a slightly higher frequency of NSU phw. Nevertheless, we see that the majority of participants in sub-corpus 2 used NSU less frequently than those in sub-corpus 1 with similar ranges of standard deviation. This clearly indicates that the participants in this study who began English in the third grade produced fewer NSU phw when compared with those who began English later.

Table 4.3 Age and number of NSU phw in sub-corpora

Sub-corpus	Age	Number of participants	Mean NSU phw	Pearson correlation coefficient
1	18	27	3.17 (SD=1.48)	R value = -0.1427
1	19	14	3.20 (SD=1.51)	P value =.312885 Result not significant
1	20	11	3.77 (SD=1.63)	at $p < .05$
Total 1		**52**	**3.30 (SD=1.51)**	
2	18	30	2.58 (SD=1.48)	R value = -0.0815
2	19	8	2.92 (SD=2.02)	P value =.601193 Result not significant
2	20	6	1.96 (SD=2.23)	at $p < .05$
Total 2		**44**	**2.56 (SD=1.67)**	

4.2.1.3 Nationality

The majority (82.3%) of the participants were Swiss nationals followed by 8 different foreign nationalities (see Table 3.8 p. 68) at 11.5% and 5 different dual nationalities Swiss/foreign at 6.2%. When the use of NSU phw of these 3 groups was compared in the entire CSC, the foreign participants used the most NSU phw at 3.94 (SD=1.70), followed by the Swiss participants at

2.90 (SD=1.59). The participants with dual nationality used the least NSU phw at 2.0 (SD=1.38).

These results could indicate that foreign nationals are initially at a disadvantage, whereas once they are assimilated into Swiss society and become Swiss citizens that their language knowledge of Swiss German and their native language could become an advantage.

Turning to the sub-corpora, the results show a slightly different picture. In Table 4.4 below we see that the frequency of NSU phw used by Swiss participants decreases in sub-corpus 2 with a similar standard deviation when compared with sub-corpus 1, whereas the use of NSU phw by foreign participants increased substantially in sub-corpus 2. This is likely due to the small number of participants in this group with a higher than average number of NSU phw. The dual national participants remain with the lowest NSU phw usage; however, they represent only 6.2% of the corpus and a greater number of participants would need to be evaluated to make tenable claims. Nonetheless, the results indicate that the Swiss participants in sub-corpus 2 used fewer NSU phw than their counterparts in sub-corpus 1 and foreigners tended to use more NSU phw. Although in general the participants in this study used fewer NSU phw in sub-corpus 2, here we see a greater fluctuation in the performance of foreign participants which signifies a possible area to investigate further to ensure that all students in the Swiss education system have equal learning opportunities.

Table 4.4 Nationality and number of NSU in the sub-corpora

Sub-corpus	Nationality	Number of participants	Mean NSU phw
Both	Swiss	79	2.90 (SD=1.59)
Both	Foreign	11	**3.94 (SD=1.70)**
Both	Dual	6	2.0 (SD=1.38)

1	Swiss	42	3.27 (SD=1.50)
1	Foreign	9	**3.59 (SD=1.67)**
1	Dual	1	2.21
2	Swiss	37	2.48 (SD=1.60)
2	Foreign	2	**5.53 (SD=0.82)**
2	Dual	5	2.0 (SD=1.38)

4.2.1.4 Native language

The complexity of language association in plurilingual Switzerland can be seen in the difference between nationality and declared native language. Thirteen participants recorded their native language differently than their nationality. Four participants stated having dual nationality, but only Swiss German as their native language, whereas five participants stated being only Swiss and having dual native languages. An additional two participants stated being only Swiss but having Croatian as their native language. Lastly, there were two participants with Swiss nationality and High German as their native language. Thus, when comparing native languages and NSU phw it can be assumed that over 10% of the participants are different from those who were included in the comparison of nationality and NSU phw.

When comparing the native languages of the participants, non-Swiss Germans averaged the most NSU phw 3.57 (SD=1.84) followed by Swiss Germans at 2.92 (SD=1.54) and the dual native language participants produced the least at 1.93 (SD=1.73), all with similar standard deviation. The sub-corpora follow a similar pattern as with nationalities except for the Swiss participants in sub-corpus 1 who used slightly more NSU phw.

Table 4.5 Native Language and number of NSU

Sub-corpus	Native Language	Number of participants	Mean NSU phw
Both	Swiss	75	2.92 (SD=1.54)
Both	Non-Swiss GE	15	**3.57 (SD=1.84)**
Both	Dual	6	1.93 (SD=1.73)
1	Swiss	39	**3.31 (SD=1.43)**
1	Non-Swiss GE	13	3.27 (SD=1.79)
1	Dual	0	-
2	Swiss	36	2.50 (SD=1.55)
2	Non-Swiss GE	2	**5.53 (SD=0.82)**
2	Dual	6	1.93 (SD=1.73)

4.2.1.5 Mother's mother tongue

The participants' mothers had a total of 19 different mother tongues including Swiss German. The majority, 66.6%, had Swiss German as their mother tongue and 31.3% had one non-Swiss German mother tongue, whereas two participants' mothers had dual non-Swiss German mother tongues. Gradually a trend can be seen in the relationship between exposure to diverse languages and cultures and amount of NSU phw. Participants with non-Swiss German speaking mothers (n=30) averaged 3.07 (SD=1.83) NSU phw, followed closely by 2.97 (SD=1.50) for participants' mothers with a Swiss German mother tongue (n=64), and the two participants with mothers who had dual non-Swiss German mother tongues had the lowest number of NSU phw at an average 0.93 (SD=1.17).

Turning to the sub-corpora, both participants with Swiss German mothers and non-Swiss German mothers used fewer NSU phw in the sub-corpus 2. The 2 participants with dual non-Swiss German mothers were only found in sub-corpus 2. Table 4.6 below shows the mean

number of reductions of NSU phw between the sub-corpora. This indicates that both groups fairly equally reduced the number of NSU phw when taught from an early age.

Table 4.6 Mean number of fewer NSU phw in sup-corpus 2 per participant

Mother's L1	Reduced number of NSU phw
Swiss German	-0.52
Non-Swiss	-0.96
Dual non-Swiss GE L1	same

4.2.1.6 Father's mother tongue

The trend continues when we examine the participants' father's mother tongue. The majority (n=59) or 61.5% of the participants' fathers had Swiss German as their mother tongue and they used the most NSU phw at a mean 3.04 (SD=1.53) when the whole corpus was evaluated. A third (n=32) of the participants' fathers had one of 16 non-Swiss German mother tongues and their children used a mean 2.89 (SD=1.77) NSU phw. Five participants' fathers had dual mother tongues with the lowest mean of NSU phw out of all three groups at 2.48 (SD=1.81).

When the sub-corpora were divided into three groups and examined, there was a reduction of NSU phw in all groups except for the multiple L1 group.

Table 4.7 Mean number of fewer NSU phw in sup-corpus 2 per participant

Father's L1	Reduced number of NSU phw
Swiss German	-0.67
Non-Swiss	-1.01
Dual/multiple L1	**+0.09**

The results reinforce the trend that fewer NSU phw were produced by the majority of the participants in sub-corpus 2. When considering the father's first language, the children from those with foreign languages benefited most from early English when quantity of NSU are examined. Although there was a slight increase in the number of NSU phw in the dual/multiple group, they had the lowest mean to begin 2.48 (SD=1.81) which was 0.48 below the CSC corpus mean of 2.96.

4.2.1.7 Languages spoken at home

The first or second language of the participants' parents are an indication of cultural influence on the participants' language background. Likewise, when we examine the use of those languages at home in relation to the amount of NSU phw, the pattern is reinforced. An initial look at the amount of NSU phw used by participants who spoke only German at home (Swiss German and High German) shows an average of 3.08 (SD=1.46) NSU phw and all others who spoke another language (or combination of languages) at home used an average of 2.77 (SD=1.85) NSU phw.

As we break the data down further, the effects of regularly speaking more than one language can be seen to influence the amount of NSU per participant. The 28 participants who recorded speaking German and one other language at home had an average 2.87 (SD=1.73) NSU phw. A further 5 participants recorded speaking German and 2 or more other languages at home and had an average of 2.26 (SD=2.47) NSU phw. The remaining 4 participants recorded speaking only one or more non-German languages at home and had an average of 2.67 (SD=2.32) NSU phw.

As the lexis and syntax of Swiss German differs considerably from High German, its regular use could affect L2 language use, therefore the data concerning languages used at home was once again broken down to investigate this. When only Swiss German was spoken at home an average of 3.09 (SD=1.49) NSU phw were used. The combination of Swiss German and one other language lowered the average to 1.99 (SD=1.62) NSU phw. The addition of another language, Swiss German and two other languages lowered the average to 1.47 (SD=2.34). The same gradual decrease is not observed when the use of High German is concerned. Speaking only High German at home led to an average use of NSU phw of 2.96 (SD=0.61) but increased to 4.23 (SD=0.77) when one further language was spoken at home and decreased again to 3.05 (SD=3.36) NSU phw when High German and two non-German languages were spoken at home. This discrepancy might be explained by the use of High German as a lingua franca within a family to communicate when no other common language is available. It would be considered unnatural for a Swiss German speaking person to voluntarily speak High German in a family setting because speaking Swiss German carries with it a strong power of identity (Weinreich, 2011, p. 93).

The immigrant background of some of the participants is from English speaking countries. The justification for their inclusion in this study is that according to Swiss statistics (see Figure 3-2 p. 71) 5.4% of Swiss residents claim English as their main language. Therefore, English speakers are considered a part of Swiss society and should be included. The CSC includes 5 participants with English backgrounds. One participant is British and speaks only English at home with their British family. Two participants have English speaking fathers and speak 25% English at home and two participants have English speaking mothers and speak 50% English at home. Thus, English speakers are actually under-represented with only 3 participants or 3.1% claiming to have English as a main language.

In Table 4.8 below, an overview of the languages spoken at home is given. Here we see that the number of NSU phw tend to gradually decrease as the number of languages used at home increases. This shows the generally positive influence of plurilingualism on the spoken output of this group of learners. There are some exceptions such as the slightly higher use of NSU phw in participants who speak one or more non-Swiss or High German language and the higher use of NSU phw among High German dual language speakers.

Table 4.8 Overview of languages spoken at home and corresponding mean NSU

Languages Spoken at home	No of participants	Mean NSU phw
Swiss German and High German only	59	**3.08** (SD=1.46)
All other language combinations	37	2.77 (SD=1.85)
Swiss German and High German only	59	**3.08** (SD=1.46)
Swiss German and High German plus 1 language	28	2.87 (SD=1.73)
Swiss German and High German plus 2 languages	5	2.26 (SD=2.47)
One or more language other than Swiss or High German	4	**2.67** (SD=2.32)
Swiss German	56	3.09 (SD=1.49)
Swiss German plus 1 other language	17	1.99 (SD=1.62)
Swiss German plus 2 other languages	3	1.47 (SD=2.34)
High German	3	2.96 (SD=0.61)
High German plus 1 other language	10	**4.23** (SD=0.77)
High German plus 2 other languages	2	3.05 (SD=3.36)
One or more language other than Swiss or High German	4	2.67 (SD=2.32)

4.2.1.8 Conclusion

The following Figure 4-3 illustrates the demographic variables and non-standard usages per one hundred words. Firstly, non-Swiss nationals and non-Swiss German L1 used the most NSU phw. The green highlighted bars represent Swiss participants who comprise the majority of the corpus. We see that they hover around the corpus average. Furthermore, some variables above the average of 2.96 indicate a monolingual environment, whereas those below the average generally indicate a plurilingual environment.

Demographic variables and number of NSU per 100 words

Non Swiss nationality		
Non Swiss German L1		
Female		
20 year olds		
19 year olds		
Swiss German spoken at home		
Swiss German or German only spoken at home		
Mother L1 non Swiss German		
Father L1 Swiss German		
Mother L1 Swiss German		
Corpus average		
Swiss German L1		
Swiss Nationality		
Father L1 non Swiss German		
Swiss German or German plus one language spoken at home		
18 year olds		
one non-Swiss German language spoken at home		
Swiss /German plus one or two languages spoken at home		
Male		
Language(s) other than Swiss German or German spoken at home		
Father dual L1 non Swiss German		
Swiss German or German plus 2 languages spoken at home		
Dual nationalities		
Father dual L1 non Swiss German and Swiss German		
Dual L1 Swiss German and non Swiss German		
Mother dual L1 non Swiss German		

Figure 4-2 Demographic variables and number of NSU phw

To sum up, this section examined seven variables. First, gender was analysed with females found to use 0.41 more NSU phw than males. Whereas, when sub-corpora were compared, it was found that females used 22% fewer NSU per hundred words in sub-corpus 2 when compared to the females in sub-corpus 1. Age was examined next with no substantial findings in the whole corpus. When the sub-corpora were analysed, the first indications of an overall reduction of NSU in sub-corpus 2 became apparent.

110

Continuing with nationality, it was found that foreign nationals (non-Swiss) used the most NSU phw followed by Swiss nationals and participants with dual nationalities who used the least. Next, native languages were analysed and a comparable pattern to nationalities was found. Non-Swiss Germans used the most NSU phw, in the whole corpus. In addition, participants with dual native languages used the least. The most noteworthy result of examining the sub-corpora was that the dominant Swiss German L1 group showed a 24.47% reduction of NSU phw in sub-corpus 2. This indicates that they benefitted from early English.

With some minor exceptions, the number of NSU phw of participants with Swiss German mother tongued parents and non-Swiss German parents were similar and those with dual mother tongues used the least. When the sub-corpora were examined again with minor exceptions, most of the groups examined used fewer NSU phw in sub-corpus 2. Lastly, languages used at home were analysed. Participants with monolingual Swiss German homes used the most NSU phw and those with a combination of Swiss German and other languages at home used decreasingly fewer NSU phw the more languages were spoken at home. The exception was when High German was spoken with a combination of other languages the pattern was not similar.

The detail of demographic information gathered gave insight into the correlation between participants' demographic variables and the number of non-standard usages. It was shown that the exposure to and use of more than one language was related to less frequent use of non-standard usages in this study. The results reflect the language use of this particular group. Therefore, further investigation on a greater scale is required to make definite conclusions.

4.2.2   Identity-related variables

This second section explores the relationship between the use of non-standard usages (NSU) of English and variables that are related to the identity of the participants. These include:

- the effects of the familiarity with the interviewer

- the participants' views on the importance of speech correctness versus message comprehension

- their views on whom they will speak English with in the future

- how they feel when speaking English and if they consider themselves learners or users of English.

The information was gathered during the third part of each interview where standardized questions were asked. As in the first section, data was initially entered into SPSS 25 and thereafter calculated in Microsoft Excel 2011.

4.2.2.1 Familiarity with interviewer

There were no major differences found in the amount of NSU and familiarity with the interviewer. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' familiarity with the interviewer and use of NSU. There was not a significant correlation between the two variables found, $r(96) = -.11$, $p = .286$. Those that knew the interviewer averaged 2.54 (SD=1.83) NSU phw compared to those who vaguely knew her at 2.50 (SD=1.05) NSU phw and those who did not know her at 2.65 (SD=1.59) NSU phw. When the sub-corpora were investigated, sub-corpus 1 followed a similar pattern as the whole corpus with familiar participants having a high mean NSU phw. However, sub-corpus 2 was just the opposite. In order to have fully investigated any correlation between interviewer familiarity and NSU phw in the sub-corpora, an even amount of familiar and non-familiar participants in each sub-corpora would have had to have been reached.

Table 4.9 Interviewer familiarity and number of NSU phw

Sub-corpus	Familiarity	Number of participants	Mean NSU phw
1	familiar	**32**	3.44 (SD=1.59)
2	familiar	5	1.64 (SD=2.06)
1	vaguely familiar	12	3.52 (SD=1.07)
2	vaguely familiar	6	1.48 (SD=1.03)
1	unfamiliar	8	2.41 (SD=1.58)
2	unfamiliar	**33**	2.90 (SD=1.61)

4.2.2.2 Importance of grammatical correctness vs understanding

During the interview, the participants were asked which they thought was more important; being grammatically correct or being understood. The hypothesis was that the participants who viewed grammatical correctness as more valuable would speak more precisely because of their heightened awareness. The answers ranged from believing that grammatical correctness was a necessity to differentiating when grammar is needed and the idea that grammar was not important as the three representative quotes below demonstrate.

'when you don't speak with grammar the others don't uh understand you and it is very hard to to learn or to speak when you don't have uh grammar' (participant 131)

'it depends where you uhm talking because uhm for example here in the lesson you look on the the grammar that is right because we have to do this on the tests and exams but if I go uhm to another country where they speak English also abroad on holidays it's really important that the people understand me I need to buy food or ask people for the right directions or I have or if I want to visit

something at church or something then it's important that people understand me and they won't correct me if do grammar faults yeah' (participant 103)

'whenever I speaked English uh forget about the grammar it's about the understanding yeah' (participant 10).

These viewpoints reiterate the conflicting standpoints in Swiss education policy where in the functionality of oral communication, being understood outweighs grammatical correctness in primary school (Werlen, 2006, p. 11) and the increased use of international language diplomas as vocational education exit exams (SBBK, 2021). As children begin to learn English as a foreign language in primary school, production is emphasised with didactic recommendations similar to the ideological foundations of ELF. Somewhere between primary and secondary school students are then expected to perform to nativelike standards. Pfenninger and Lendl (2017) highlight the problems of this language disparity and suggest the need for greater cooperation and alignment of curriculum between primary and secondary schools.

Another point of disparity is that Matura students are primed for the language proficiency needed for university entrance. The majority who attend secondary and vocational school are left to educate themselves if they want to cross the educational divide into university because foreign language education is not part of the curriculum in most vocational study paths. Although vocational schools offer elective language courses in the evening, apprenticeship employer permission is needed. Thus, equal learning opportunities are strived for but not guaranteed.

Coming back to the results, the nine participants who judged the importance of grammatical correctness and understanding to be equal used the least number of NSU phw at a mean of 2.35

(SD=1.50) phw. The majority of participants (n=76) believed that being understood was most important and they averaged 2.95 (SD=1.55) NSU phw. Surprisingly, the 11 participants who believed that grammatical correctness was most important used the most NSU phw at a mean 3.57 (SD=2.07). The same pattern was observed in both sub-corpora with the distribution of the three groups being equal. The only exception was the number of NSU which were higher in sub-corpus 1. This discrepancy of higher NSU use in sub-corpus 1, however, has already been established in section 4.2.1.

To understand why the 11 participants who thought grammatical correctness was most important used the most NSU phw, First the NSU phw range was checked to rule out several participants distorting the average and it was found that actually two of the participants had the lowest number of NSU in the entire corpus and the rest were varied as can be seen in Table 4.10 below.

Table 4.10 View of grammatical correctness and NSU

Sub-corpus	Number of NSU phw	Sub-corpus	Number of NSU phw
1	0.12	2	0.11
1	5.70	2	3.54
1	2.16	2	3.30
1	3.62	2	4.65
1	4.81	2	4.95
1	6.35		

Evaluation of the type of NSU revealed that the 9 participants who used over 20 NSU all used the most common 4 NSU. These included non-standard usages of lexical choices, present/past tense, word order and prepositions, which eliminated peculiar inconsistencies. Next, proficiency levels were examined and this revealed that the participants had a range of proficiency levels; A2 (n=4), B1 (n=4), B2 (n=1), and C1 (n=2). This indicated that the correlation was not proficiency bound. Lastly, their opinion of who they were going to speak

English with in the future revealed that 7 of the 11 believe that they will speak mostly with native English speakers in the future and this could be the reason why they value correctness. Perhaps this indicates traditional learner expectations when learning a foreign language; to be able to converse and fit into a native-speaking community.

4.2.2.3 Who they think they will speak English with

Participants were asked who they thought they would speak English with in the future, native or non-native speakers. This question was asked to assess the participants' understanding of how English is presently used in Switzerland and their individual aspirations of its future use. In total 30.2% believed they would be speaking English with native speakers in the future, 38.54% believed they would be speaking with non-native speakers and 31.25% believed they would be speaking with both. When we look at the sub-corpora, the distribution is similar to that of the whole corpus with the majority of the participants believing that they will speak English mainly with non-native speakers. This awareness is certainly heightened in quad-lingual Switzerland where it is not uncommon for compatriots with different first languages to speak English as a lingua franca for communication. (Durham, 2016, p. 16; Rosenberger, 2009, p. 129).

Those who said they would be speaking with native speakers n=29 used the highest number NSU phw at mean 3.22 (SD=1.86), while those who said they would be speaking with both native and non-native n=30 were slightly lower at 2.86 (SD=1.49). The majority n=37 who said they would be speaking with non-native speakers used the least at 2.84 (SD=1.54). As with the participants in the section above who used the most NSU phw and feel grammatical correctness is important, those who believe they will be speaking with native speakers were found to be an evenly mixed group without obvious markers why they produced more NSU phw than those

116

with other beliefs. Language learning anxiety and striving for perfection have been linked to oral performance (Gregersen and Horwitz, 2002; Woodrow, 2006). Further investigation into the correlation of anxiety, expectations and performance might clarify why CSC participants who value correctness and expect to converse with native speakers performed below average.

4.2.2.4 Feel or act differently

Participants were asked the open question: Do you feel or act differently when you are speaking English? Most of the participants (n=49) said they did not feel or act differently when speaking English and the majority of those said that it was just another language. This indicates that speaking multiple languages is normal for them. Of those that said they felt or acted differently, 10 participants mentioned their voice being different and only 3 mentioned speaking English as very positive while 34 participants associated speaking English as negative. The comments ranged from feeling a bit uncomfortable to hating English because it is difficult. All comments were evaluated and divided into 4 categories; **positive difference** 'I feel good, cool, proud, smile, have fun', **neutral** or no difference 'it's just like German, it's a language', **negative difference** 'nervous, unsure, different, I can't say what I think' and **very negative difference** 'ashamed, very uncomfortable, very insecure'.

A Pearson correlation coefficient was computed to assess the linear relationship between positive feelings of speaking English and fewer NSU phw. There was a positive correlation between the two variables, $r(95) = .24$, $p = .015$. In Table 4.11 below, we clearly see a difference how the participants of the two sub-corpora feel when using English. In sub-corpus 1, there is a range of feelings with negative and neutral feelings dominating the field and 5 participants with very negative feelings. Those participants had the highest number of NSU phw. Furthermore, when the number of words spoken is calculated it shows an average of only 520

words per person. This is half of the corpus average. Turning to sub-corpus 2, we see that the

majority of the participants did not feel different when speaking English. These results indicate

that the attitudes towards speaking English of the participants in this study were positively

affected by starting to learn English at an earlier age.

Table 4.11 Feelings towards speaking English

Sub corpus	No. of participants	No. of NSU phw	Sub corpus	No. of participants	No. of NSU phw
Positive difference					
1	7	3.04 (SD=0.88)	2	5	**1.51** (SD=0.75)
Neutral or no difference					
1	18	2.88 (SD=1.66)	2	**31**	2.58 (SD=1.72)
Negative difference					
1	**21**	3.63 (SD=1.58)	2	8	3.12 (SD=1.74)
Very negative difference					
1	**5**	3.55 (SD=1.25)	2	-	-

As mentioned above, exploring the effects of anxiety could prove to be fruitful in further studies

to determine how the degree of positive or negative feelings affects production. In hindsight,

perhaps questions such as 'Do you feel positive or negative towards speaking English?' and

'Do you feel your attitude towards English affects your ability to use it?' would have given

more insight.

4.2.2.5 Learner, user or both

The last question of the interviews was 'Do you consider yourself a learner or user of English

or both?'. The rationale for this question was to determine how the participants perceived

themselves as English speakers. Participation in some native or ELF spoken corpora allows

participants to self-declare if they are in fact English users and deem themselves proficient.

According to the participants' answers, the whole CSC has more self-proclaimed learners than

users and almost as many participants who consider themselves both learners and users as can

be seen in Table 4.12 below. Yet, a third of the participants consider themselves users and could have hypothetically contributed to other high stakes corpora. Nonetheless, the CSC corpus should still be considered a learner corpus with insight into a cross-section of Swiss youth. When we look at the difference between the two sub-corpora, we can conclude that the effects of early English were positive on the L2 self-image of the participants with 16.7% fewer claiming to be learners and 19.4% more claiming to be both in sub-corpus 2.

Table 4.12 Corpus construct; Learner or user of English or both

	Percentage in CSC	Percentage in Sub-corpus 1	Percentage in Sub-corpus 2
Learner	38.5%	46.2%	29.5%
User	33.3%	34.6%	31.8%
Both	28.1%	19.2%	38.6%

Turning to the relationship between L2 self and NSU phw, Table 4.13 shows the statistics.

Table 4.13 L2 self and number of NSU

Sub-corpus	L2 self	Number of participants	Mean NSU phw
1	learner	24	3.59 (SD=1.33)
1	user	18	2.89 (SD=1.70)
1	both	10	3.36 (SD=1.54)
2	learner	13	3.31 (SD=1.83)
2	user	14	2.47 (SD=1.65)
2	both	**17**	**2.06** (SD=1.43)

As we see, there is a difference between the total of NSU phw and identifying oneself as either a learner or user of English. This is counter to small scale previous work of the author where it was discovered that *learners* used fewer ELF features than *users* when only features of English as a lingua franca were investigated (Oswald, 2010). Therefore the eight ELF features described on page 48 were examined and two features were found where the learners had fewer NSU;

confusing the relative pronouns *who* and *which* and omitting definite and indefinite articles *a, an, the* where they are obligatory in Standard English. After further investigation, the difference was considered too slight and not pursued.

4.2.2.6 Conclusion

In summary, the variables classified as identity variables were thought to portray the participants' ideologies and feelings. As can be seen in Figure 4-3 below when the NSU phw are compared, those who felt both grammar and understanding are important used the least NSU phw. Furthermore, those who believed grammatical correctness was of greater importance used the most NSU phw. When the variables above the corpus average are compared with NSU phw in Figure 4-3, all variables except familiarity with the interviewer express variables with a negative aspect or association with possible higher anxiety levels.

Directly following, the NSU phw of the sub-corpora are illustrated in Figures 4-4 and 4-5. Here we see that in both sub-corpora that negative attitudes towards speaking English are linked to increased NSU phw. Furthermore, we see the difference in variables below or above the corpus average.

Figure 4-3 Identity variables and number of NSU phw
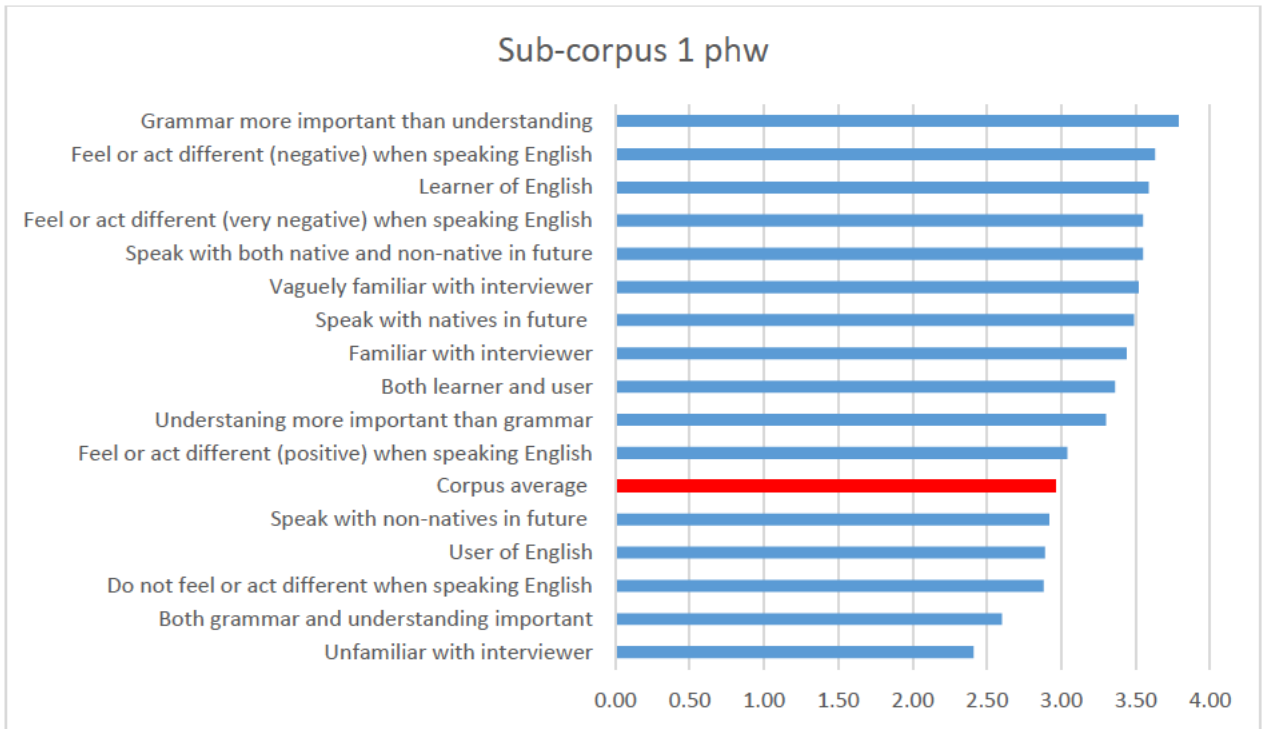
Figure 4-4 Identity-related variables sub-corpus 1 phw
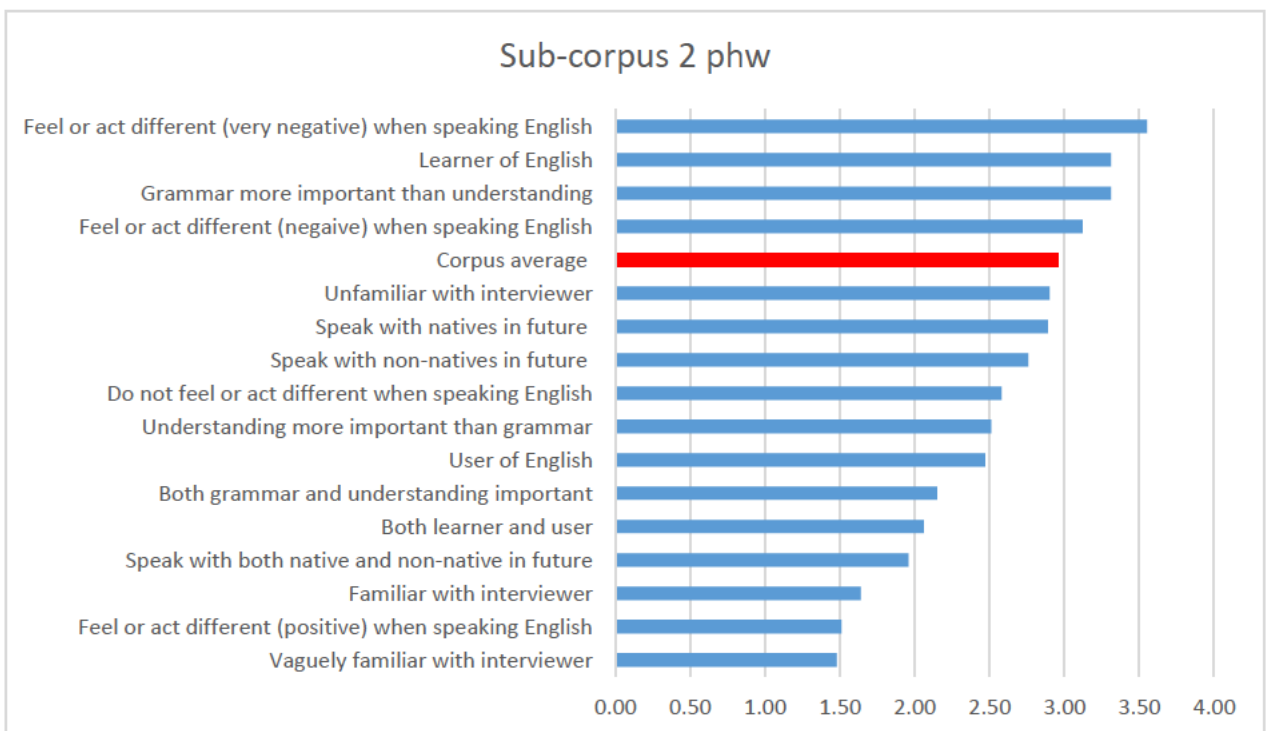
Figure 4-5 Identity-related variables sub-corpus 2 phw

4.2.3 Educational variables

Demographic and identity variables are generally given, and very individual. Educational variables, on the other hand, can be affected by policy changes. Therefore, the outcome of the educational variables can give insight to which if any areas might benefit from further development or reform.

4.2.3.1 Current studies

The CSC participants either followed an apprenticeship or Matura path of study. The 79 apprentices averaged 3.25 NSU phw compared to the 17 Matura participants who used a mean 1.60 NSU phw. The main reason for the large discrepancy between the two groups can be explained by the curriculum of the two educational paths. Apprentices are trained to perform a profession and, from the age of sixteen, foreign languages are no longer part of the curriculum for most vocational paths. The apprentices in the CSC are the exception because 36 were doing a business apprenticeship where English is included in the curriculum and Cambridge English Qualifications at B1 are taken as a final exam. The remainder of the apprentices (n=40) were being trained as either electricians, electrical planners, carpenters, gardeners, draughtsmen, nurses or medical assistants with all but three participants enrolled in bilingual courses where English is used as the language of instruction for at least one lesson per week during their 3-4-year apprenticeship. This approach of Content and Language Integrated Learning (CLIL) is becoming a popular addition throughout vocational schools in Switzerland where formal language instruction is not included in the curriculum. Thus, the apprentice participants in this study are likely to have a slightly better command of English than the general population of similar youth. The Matura participants on the other hand are prepared for university entrance and have at least one other foreign language besides English in their curriculum throughout their secondary school years and are expected to reach a CEFR level of B2/C1 in each.

Turning to the differences between the sub-corpora, we see in Table 4.14 below that there is a mean difference of almost one NSU per hundred words between the apprentices, with fewer in sub-corpus 2. The Matura participants in sub-corpus 2 followed with producing a mean of 1.21 NSU phw fewer than sub-corpus 1. These results show that the participants in this study who started English earlier produced considerably fewer NSU when compared with other participants who followed the same study path. Furthermore, participants who studied for a Matura produced almost half as many NSU phw as those who did an apprenticeship. One could speculate that this might be due to the increased number of years of study, but as the following section shows, this is not necessarily the case. It is much more likely that the more demanding curriculum and general scholastic expectations of Matura students are higher than those at the vocational level.

Table 4.14 Study path and NSU

Sub-corpus	Current studies	Number of participants	Mean NSU phw
1 and 2	Apprenticeship	79	3.25 (SD=1.56)
1 and 2	Matura	17	1.60 (SD=1.17)
1	Apprenticeship	40	**3.71** (SD=1.35)
2	Apprenticeship	39	**2.79** (SD=1.63)
1	Matura	12	**1.96** (SD=1.21)
2	Matura	5	**0.75** (SD=0.44)

4.2.3.2 Years of English study

The 96 participants in the study ranged from having 2 to 11 years of English studies. One might assume that there would be a gradual decrease of the number of NSU phw the longer one studied. As can be seen in the Figure 4-6 below, this is not always the case.

The volatility of the NSU phw and years of study can be explained by the structure of the corpus and the two sub-corpora. Interestingly, the majority at 28% of the participants had 6 years of English. This group is comprised of 18 apprentices and 9 Matura participants. Although they received the same number of years of study, there is a substantial difference in the amount of NSU phw. The apprentices who had 6 years of English studies averaged 3.53 (SD=1.43) NSU phw, compared to the Matura participants who also had 6 years of English 1.89 (0.97) NSU phw. This indicates that there are other factors involved in the number of NSU phw and years of study such as the differing learning objectives or teaching approaches as mentioned above.

Figure 4-6 also shows a gradual decrease of NSU phw over the span with a zig-zag effect which indicates the transition from teaching English in secondary to primary schools. An example is that roughly half of the participants (n=46) had either 6 or 9 years of study and averaged similar amounts of NSU (approximately 3 phw), with those with 9 years averaging even slightly more NSU than those with 6 years of study. As stated above, those with 6 years of study are a mix of apprentice and Matura participants, whereas those with 9 years are all apprentices and all but one belongs to sub-corpus 2 which means they started English lessons in the third grade.
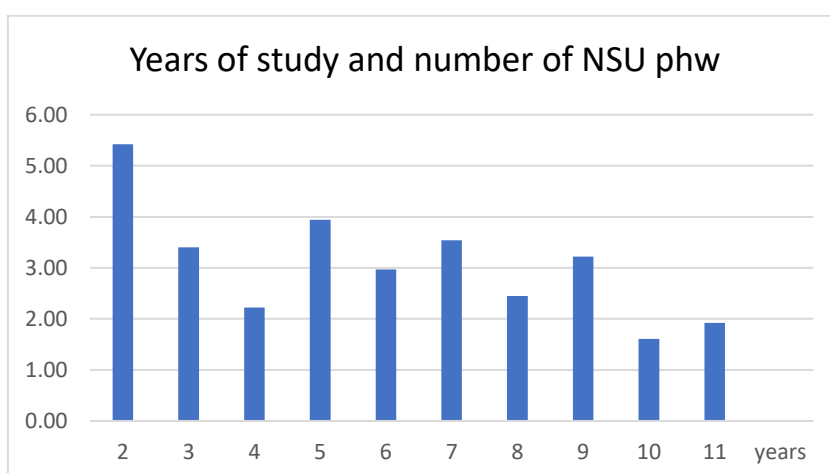
Figure 4-6 Years of study and NSU phw

Analysing the number of NSU and years of study indicates how often the spoken output deviated from standard varieties. However, it cannot be used as an indicator of language proficiency. This is apparent when the years of study and proficiency levels are compared.

To reach an **A2** level, 1 participant had only 2 years of English study, 7 participants had 5 years, 9 participants had 6 years, 1 participant had 7 years, 6 had 9 years and one had 11 years.

To reach a **B1** level, 2 participants had 3 years of English study, 3 participants had 4 years, 4 participants had 5 years, 8 participants had 6 years, 3 participants had 7 years, 1 participant had 8 years, 10 participants had 9 years, 8 participants had 10 years and one had 11 years.

To reach a **B2** level, 1 participant had 3 years of English study, 1 participant had 5 years, 8 participants had 6 years, 4 participants had 7 years, 2 participants had 8 years, 2 participants had 9 years and 6 participants had 10 years.

Six participants had a **C1** level, and it took them respectively each 4, 5, 6, 7, 9 and 10 years to reach it. Lastly, the one participant with a C2 level had 6 years of English study.

4.2.3.3 Proficiency levels

Turning to proficiency levels and number of NSU phw comparison. A Pearson correlation coefficient was computed to assess the linear relationship between CEFR proficiency levels and number of NSU phw. There was a significant correlation found between the two variables, $r(96)$ = -.97, p = .005. Table 4.15 below shows that in the whole CSC as well as in the sub-corpora, the number of NSU phw gradually decreases as the proficiency levels increase. This was expected as accuracy is considered one of the three components of proficiency by many L2

practitioners and SLA researchers (Housen et al., 2012, pp. 1-2). The perception that proficiency consists of the interlinked components of complexity, accuracy and fluency was introduced by Skehan (1996). This study focuses on accuracy; however, the corpus could be used at a later date to investigate complexity and fluency.

Table 4.15 CEFR proficiency levels and number of NSU

Sub-corpus	CEFR level	Number of participants	Mean NSU phw	Pearson correlation coefficient
Both	A2	25	4.38 (SD=1.18)	R value = -0.972
Both	B1	40	3.01 (SD=1.19)	P value = .005601 Significant at
Both	B2	24	1.84 (SD=1.21)	p < .05
Both	C1	6	0.55 (SD=0.92)	
Both	C2	1	0.49	
Sub-corpus 1	A2	15	4.21 (SD=1.31)	R value = -0.990
Sub-corpus 1	B1	19	3.66 (SD=1.18)	P value = .00107 Significant at
Sub-corpus 1	B2	14	2.39 (SD=1.25)	p < .05
Sub-corpus 1	C1	3	1.12 (SD=1.47)	
Sub-corpus 1	C2	1	0.49	
Sub-corpus 2	A2	10	4.64 (SD=0.97)	R value = -0.986
Sub-corpus 2	B1	21	2.62 (SD=1.13)	P value = .0138 Significant at
Sub-corpus 2	B2	10	1.07 (SD=0.61)	p < .05
Sub-corpus 2	C1	3	0.17 (SD=0.21)	
Sub-corpus 2	C2	0	-	

It would seem fundamental that accuracy corresponds with proficiency levels. However, it remains questionable how significant accuracy is during initial phases of learning in the primary school classroom when as mentioned on page 16, teacher guidelines and learning objectives state that the communicative intention always comes before formal correctness. Furthermore, Davies (1976, p. 441) distinguished three major stages of foreign language knowledge which

described how productive skills are developed after receptive skills. The Swiss Conference of Cantonal Ministers of Education report agrees and states: 'in language acquisition: usually a higher level is reached in the area of receptive skills than in the area of productive skills' (Swiss Conference of Cantonal Ministers of Education, 2007, p. 22). This has subsequently led to the Swiss CEFR proficiency goals being higher in receptive skills than productive skills for all foreign languages and all levels of mandatory education (ibid. p. 23). Thus, it is possible that the CSC participants' receptive English skills are higher than recorded because only spoken production was accessed.

4.2.3.4 Native or non-native English teachers

Across all levels of education, 88% of schools in Switzerland are public (bfs.admin[22]). Although there are exceptions, it is the norm that English teachers in public schools are Swiss or non-native English speakers. The proficiency levels of primary school English teachers can also vary greatly and some variation of NSU phw was expected.

As anticipated, the majority or 64 of the participants learned English from non-native teachers, 22 were taught by both native and non-native teachers, 4 were taught by native English speakers and 6 participants did not know if their teachers had been native or not. The NSU phw were calculated and contrary to expectations, no correlations to teachers' language background were found. Each of the 4 groups produced similar results when NSU phw were calculated with a nominal difference of 0.51.

---

[22] https://www.bfs.admin.ch/bfs/en/home/statistics/education-science/educational-institutions/school-status.html

4.2.3.5 Language of instruction

Only two participants were instructed in primary school in a language other than German: one in Arabic and the other in English. Those two participants were at opposite ends of the NSU and proficiency spectrums. It was concluded that the data was insufficient for comparison.

4.2.3.6 French proficiency

As mentioned in Section 3.5.9, the majority (66.66%) of the CSC participants' self-evaluated French proficiency levels reached the minimum goal of A2 as stated in the curriculum guidelines for Central Switzerland (Bildungs- und Kulturdepartement des Kantons Luzern, 2020). Whereas when the sub-corpora were investigated, 73.06% reached the minimum goal when French was taught first compared to only 59.08% when it was taught second. This does not correspond with the expectations after research conducted in Central Switzerland in 2010 stated that all students reached the goal of A1 in French listening after completion of the sixth grade. (Heinzmann et al., 2010, p. 1).

As mentioned in Section 1.3.3, a minimum of two foreign languages are taught during the mandatory school years in Switzerland. The sequence and age in which foreign languages are taught changed between 2005-2008 which means that the 52 participants in sub-corpus 1 were taught French in the fifth grade and English in the seventh grade. The 44 participants in sub-corpus 2 on the other hand were taught English in the third grade and then French in the fifth. Although all of the participants were taught French as a foreign language commencing in the fifth grade for a duration of at least 4-5 years, 15.62% claimed to have no French language skills.

In the comparison of French and English CEFR levels in Table 4.16 below, we see that the majority of participants in this study reached a B1 proficiency level or higher of English. However, these results might be better than the Swiss average because A2 was the minimum requirement to participate in this study. With 33.32% of the participants not reaching an A2 level in French, we can also speculate that a similar amount of the population does not have A2 level English skills and that this study is compiled of learners who have successfully reached the minimum scholastic expectations in English. This is supported by the NFP Project 56 summary (Heinzmann et al., 2010, p. 1) which concluded that 77% of pupils reached the minimum English learning objectives at the end of the sixth grade. Coming back to Table 4.16, we clearly see that the majority of participants have one CEFR level lower French than English proficiency and a lower percentage of sub-corpus 2 participants reached B1 and B2 levels.

Table 4.16 Percentage of English and French CEFR proficiency levels

CEFR level	English total CSC	English Sub-corpus 1	English Sub-corpus 2	French total CSC	French Sub-corpus 1	French Sub-corpus 2
none	-	-	-	15.62%	19.23%	11.36%
A1	-	-	-	17.70%	7.69%	29.54%
A2	26.04%	28.89%	22.72%	**37.50%**	**40.38%**	**34.09%**
B1	**41.66%**	**36.53%**	**47.72%**	15.62%	21.15%	9.09%
B2	25%	26.92%	22.72%	13.54%	11.53%	15.90%
C1	6.25%	5.76%	6.81%	-	-	-
C2	1.04%	1.92%	-	-	-	-

It was expected that higher levels of self-proclaimed French proficiency would be correlated with higher levels of English proficiency and thus fewer NSU phw. A Pearson correlation coefficient was computed to assess the linear relationship between French CEFR proficiency levels and number of NSU phw. There was not a significant correlation found between the two variables, $r(96) = -.68$, $p = 2.0$. When we examine the corpus as a whole in Table 4.17 below, it appears that a possible positive relationship between the level of French proficiency and fewer

NSU phw is only seen in those with a B2 level. Reviewing the sub-corpora, the data is inconclusive with slight variations of NSU phw within French levels in the sub-corpora and again only participants with B2 level French using considerably fewer NSU phw.

Table 4.17 French proficiency and number of NSU phw

French proficiency and number of NSU			
French CEFR level	Number of participants	Mean NSU phw	Pearson correlation coefficient
Total CSC corpus			
No French	15	3.29 (SD=1.37)	R value = -0.687
French CEFR level A1	17	2.96 (SD=1.64)	P value = 2.0004 Not significant at
French CEFR level A2	36	3.17 (SD=1.65)	p < .05
French CEFR level B1	15	3.24 (SD=1.65)	
French CEFR level B2	13	1.68 (SD=1.31)	
Sub-corpus 1			
No French	10	2.85 (SD=1.37)	R value = -0.336
French CEFR level A1	4	3.96 (SD=1.14)	P value = .58038 Not significant at
French CEFR level A2	21	3.58 (SD=1.56)	p < .05
French CEFR level B1	11	3.37 (SD=1.64)	
French CEFR level B2	6	2.54 (SD=1.49)	
Sub-corpus 2			
No French	5	4.18 (SD=0.93)	R value = -0.856
French CEFR level A1	13	2.65 (SD=1.68)	P value = .06416 Not significant at
French CEFR level A2	15	2.61 (SD=1.66)	p < .05
French CEFR level B1	4	2.88 (SD=1.89)	
French CEFR level B2	7	0.95 (SD=0.49)	

4.2.3.7 Stays in English speaking country

It is a general belief that a stay in a foreign country will improve foreign language skills whether classroom instruction is involved or not. However, Tanaka and Ellis (2003, p. 67) report that

the main findings of Freed (1998) and Coleman (1997) conclude that study abroad programmes tend to increase fluency but have minimal effect on accuracy and complexity. Götz and Mukherjee (2018) recorded a strong correlation between months spent abroad and fluency in their study of the German section of the LINDSEI learner corpus. To determine if the CSC show similar results, the speech rate was calculated and compared to the NSU phw. The total number of participants' spoken words were divided by the recording length to produce words per minute. The recording length includes the interviewers' questions. However, the interviewer's speed and length of talking time is considered to be consistent throughout the interview process due to the predetermined constraints.

Table 4.18 below summarises the results by showing that there was no measurable difference in the words per minute between those who had a stay abroad and those who did not. However, we do see a slight increase from those with between 5 and 14 weeks and after 40 weeks. What is interesting is that there is a substantial decrease in the overall amount of NSU phw used by those who stayed in an English-speaking country.

Table 4.18 Number of words per minute and NSU

Number of weeks in English speaking country			
Number of weeks	Number of participants	Mean number of words per minute	Mean number of NSU phw
No stay abroad	50	89.43 (SD 40.12)	3.46 (SD 1.64)
Stay abroad	46	90.54 (SD 37.89)	**2.42** (SD 1.24)
1-2 weeks	14	72.92 (SD 28.42)	2.88 (SD 1.32)
3-4 weeks	15	88.63 (SD 31.04)	2.40 (SD 1.16)
5-9 weeks	5	119.51 (SD 65.17)	2.88 (SD=2.01)
10-14 weeks	4	111.69 (SD 28.33)	1.53 (SD=0.95)
20-24 weeks	5	**73.44** (SD 25.08)	2.80 (SD=1.43)
40-80 weeks	3	134.38 (SD 23.37)	0.19 (SD=0.18)

Now turning to accuracy, the participants who had stayed in English speaking countries averaged 2.42 NSU phw compared to 3.46 NSU phw from those who had never been to an

English-speaking country. To determine if the length of stay had an effect on NSU phw the number of weeks and NSU phw use were computed using a Pearson correlation coefficient to assess the linear relationship between length of stay and number of NSU phw. There was not a significant correlation found between the two variables, $r(45) = -.28$, $p = .057$[23]. CEFR levels were also investigated and as can be seen in Table 4.19 below, stays between 1 and 24 weeks made little difference except for the 4 participants who spent 10-14 weeks. As we can see, the three participants who stayed between 40-80 weeks all had a C1 level, so it is difficult to conclude whether the stay or the proficiency level resulted in the reduced number of NSU phw. Interestingly, two of the three who can be considered the most fluent according to the word per minute, recorded having only an A2 level of French proficiency.

Table 4.19 Number of weeks in English speaking country

Number of weeks in English speaking country			
Number of weeks	Number of participants	mean number of NSU phw	CEFR levels
None	50	3.46 (SD=1.64)	19-A2, 23-B1, 5-B2, 2-C1, 1-C2
1-2 weeks	14	2.88 (SD=1.32)	5-A2, 3-B1, 6-B2
3-4 weeks	15	2.40 (SD=1.16)	6-B1, 9-B2
5-9 weeks	5	2.88 (SD=2.01)	1-A2, 3-B1, 1-C1
10-14 weeks	4	**1.53** (SD=0.95)	2-B1, 2-B2
20-24 weeks	5	2.80 (SD=1.43)	3-B1, 2-B2
40-80 weeks	3	**0.19** (SD=0.18)	3-C1

4.2.3.8 Conclusion

To sum up the effect of educational variables on non-standard usages, the participants who studied a Matura used half as many NSU phw as those doing an apprenticeship. As stated above this is most likely due to the increased time in the classroom and higher scholastic goals. Then, years of study were examined, and it was determined that the NSU phw tended to decrease as years of study increased; however, the numbers remained volatile until over 10 years of English

---

[23] This calculation excluded the data from the one participant who spent 80 weeks abroad so as not to construe the findings.

study which were clearly the Matura students because they were the only group who could have had 10 years of English. The fluctuation can be attributed to policy change with half of the participants falling within 6 and 9 years of English studies.

When the whole corpus and its sub-corpora are evaluated, there is a clear correlation between higher CEFR English proficiency level and fewer NSU. A significant correlation between higher French proficiency and fewer NSU was not found. However, it was suggested as being possible after a French proficiency level of B2 had been reached. Furthermore, a higher percentage of the participants reached the minimum scholastic targets in French when it was taught first. Due to the fact that all participants needed an A2 level to participate in the interviews, it cannot be determined if the same is true when English is taught first or second. The data pertaining to differing languages of instruction was too small to evaluate and there was no difference found between native and non-native teachers.

Lastly, stays abroad were evaluated and it was found that the participants who stayed in an English-speaking country used 1.04 NSU fewer per hundred words 2.42 phw as opposed to 3.46 phw from those who had not stayed in an English-speaking country. When the length of the stay was investigated it was found that there was not a significant correlation between the length of stay and amount of NSU phw. Only the three participants who stayed between 40-80 weeks showed a dramatic drop in NSU phw and they all have a C1 proficiency level. Figure 4-7 illustrates the preceding summary of educational variables.
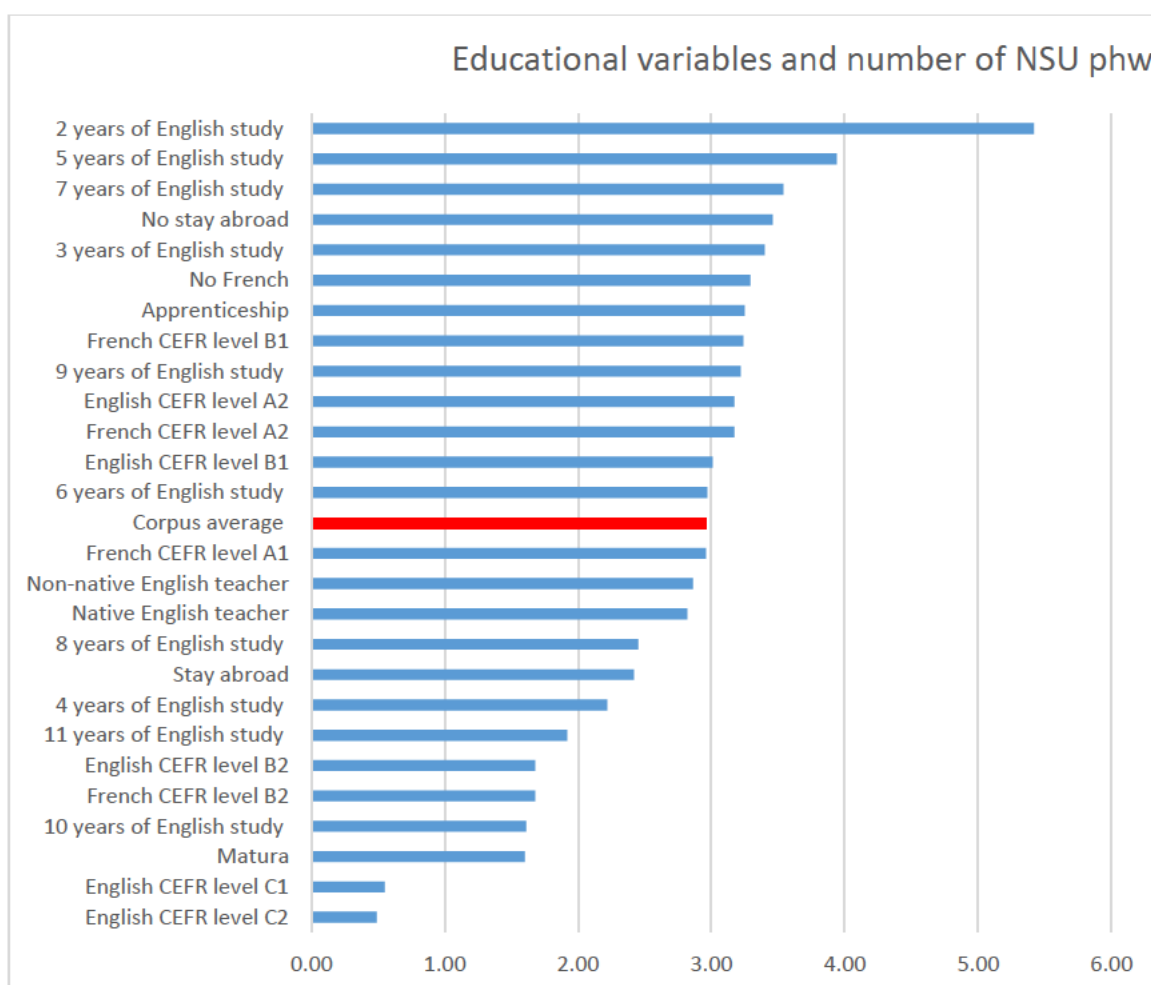
Figure 4-7 Educational variables and number of NSU phw

4.2.4 Miscellaneous Variables

There was further data gathered that could aid in obtaining a better understanding of the participants' spoken output, most of which will aid in exploring fluency. As this study is focused primarily on the aspect of accuracy, exploring those variables in detail is not currently possible in the scope of this thesis. It is planned to explore the data further for information about fluency and complexity for future research projects. Data still to be fully explored includes free speech topic chosen, number and length of participants' turns, word types and tokens, interview length, speed of speech, truncated words, filled pauses and back channelling, number of laughs, self-correction, repeats. An especially interesting topic is the use of German code-switching. Of the 850 instances, four types of code-switching were found: one-word switch, discourse

markers, questions or clarifications and longer phrases. Figure 4-8 below shows an initial attempt to understand this aspect of the CSC corpus and will be explored fully at a later date.
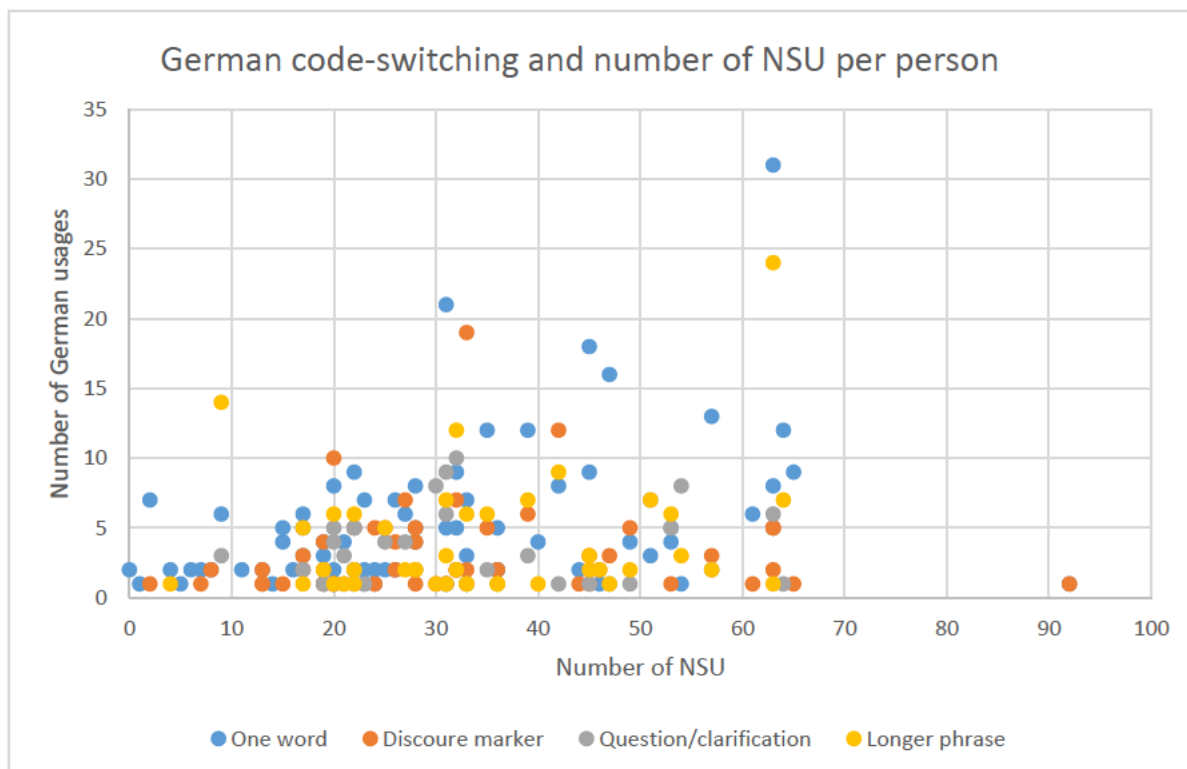
Figure 4-8 German code-switching and number of NSU

4.3 Summary of variables and non-standard usages findings

Section 4.2 dealt primarily with answering research question one which sought possible relationships between NSU and demographic, identity-related and educational variables. The main objective of posing this question was to understand the influence of these variables on spoken language and determine if further investigation is merited or pedagogical recommendations warranted.

While investigating the demographic variables it quickly became apparent that in general the early English starters of sub-corpus 2 used fewer NSU phw which indicated that they were more

accurate than the late English starters of sub-corpus 1. Nevertheless, further evidence provided valuable information about gender, age, nationality, native languages of participants and their parents. Beginning with gender and age, there were no significant correlations found with non-standard uses. A relationship between nationality and NSU phw, however, was found. Non-Swiss were found to use the most NSU followed by Swiss and then dual nationals using the least. When native languages were analysed, Swiss German L1 participants used slightly less NSU phw and non-Swiss German L1 participants used the most NSU phw with dual native language participants using overall the least. The advantage of being a Swiss national decreased when the L1 of parents and languages spoken at home were evaluated against NSU phw. Participants of parents with Swiss German L1 used similarly high numbers of NSU phw with those with non-Swiss German L1 parents and those with dual native languages other than Swiss German who used the least. A similar pattern followed with participants who lived in monolingual Swiss German homes using the most NSU phw and those who lived in multilingual homes using decreasingly fewer NSU phw the more languages that were spoken at home.

Turning to identity-related variables, no conclusive evidence was found correlating familiarity with the interviewer and NSU phw. Views on the importance of grammatical correctness and understanding, however, revealed interesting results. Participants who viewed grammatical correctness most important used the highest number of NSU phw, those who viewed understanding most important used fewer and the participants who viewed grammatical correctness and understanding as equal, used the least NSU phw. The majority of participants who highly valued correctness, also believed that they would be speaking primarily with native speakers in the future and again the participants who anticipated native standards used the greatest number of NSU phw when compared with those who believe they would converse with

non-native speakers or both. Positive and negative feelings about using English also affected performance with those who felt positively towards speaking English using the least NSU phw with increased use of NSU phw with increased negativity and presumed anxiety. Lastly, the question of L2 self-perception was asked to discover if the participants' NSU phw use was related to their perception of being a learner or user of English or both. Minimal difference was found when the whole CSC was investigated. However, a closer look at the sub-corpora revealed that more participants with early English perceived themselves as users or both user and learner when compared to the sub-corpora of late starters.

Educational variables included current studies of the participants with apprentices and Matura students representing the Swiss dual-track system. Due to the nature of the Swiss educational system, it was not surprising that the Matura students outperformed the apprentices, but it is noteworthy that the evidence suggests that both Matura and apprentices benefited from Early English with a substantial reduction of NSU phw. Next, it was verified that higher CEFR proficiency levels corresponded with fewer NSU phw in all areas. The hypothesis that initial language instruction with generalist as opposed to specialist language teachers would affect spoken accuracy was not established with only nominal differences found. The investigation into possible effects of primary language of instruction was abandoned when it was found that only two percent of the participants were not instructed in a typical Swiss German classroom. Interesting correlations were found with French proficiency and English NSU phw in that there was variable accuracy observed until French level B2 was achieved and those participants used considerably fewer NSU phw in English. Lastly, the effect on stays in English speaking counties and NSU phw were investigated and the evidence shows that participants who had stayed in an English speaking country used one NSU phw less than those who had not. However, there were no significant correlations found between length of stay and number of NSU. It can be argued

that this is due to the shorter time frames investigated in weeks instead of months as in previous studies. From a pedagogical stance, the results above have brought forth areas for discussion and further investigation which will be covered in greater detail in Chapter 7.

Chapter 5 ANALYSIS OF THE GRAMMAR AND LEXIS OF NON-STANDARD USAGES

The previous section provided an understanding of the correlation between the variables and frequencies of non-standard usages. In this section, the corpus will be investigated to explore the grammar and lexis used in a non-standard way. Through this investigation, a greater understanding of the exact language that was used in a non-standard way will come to light which allows for conclusions to be drawn about language use and pedagogical implications. Keeping research question two in mind, the sub-corpora will be compared throughout.

RQ 2: What effect does early English, taught with a productive focused curriculum, have on the number and type of non-standard features used?

The non-standard usages found in the CSC are listed in Table 4.20 below. Although sub-corpus 2 consists of slightly more words (n=1,401) than sub-corpus 1, the difference accounts for less than one percent and therefore, for simplicity, total NSU numbers are discussed throughout this section. However, since there were an uneven number of participants (52:44) in the sub-corpora, when reporting on participants' NSU use and proficiency levels from the sub-sections, percentages have been included to aid comparison.

Table 5.1 below gives an overview of the 25 NSU by stating instances and their sub-corpora distribution as well as number of participants who used them. The light green shaded areas indicate which sub-corpus has the greater number of NSU. The overall greater amount of NSU (n=364) in sub-corpus 1 is herewith reflected. The percentage of participants' usage of 11 of the 25 NSU were found to have at least a 5% variance between the sub-corpora and the corresponding table cells are marked in orange. As the corpus was analysed, a small amount of

classification inconsistencies, mainly with the verb *to be,* were found and reallocated to finalise

the NSU classification process.

Table 5.1 Distribution of non-standard usages (NSU)

Type of non-standard usage (NSU)	No of NSU in CSC	No of part. in CSC	% of CSC part.	NSU in Sub-corpus 1	No of part. sub-corpus 1	% of Sup-corpus 1 part.	NSU in Sub-corpus 2	No of part. sub-corpus 2	% of Sup-corpus 2 part.
1. Omission or insertion of the third person present tense –s	170	61	63.54	107	34	65.38	63	27	61.36
2. Confusing the relative pronouns *who* and *which*	19	15	15.63	8	7	13.46	11	8	18.18
3. Omission of definite and indefinite articles *a, an, the*	49	29	30.21	38	20	38.46	11	9	20.45
4. Insertion of definite and indefinite articles *a, an, the*	108	50	52.08	78	36	69.23	30	15	34.09
5. Failing to use correct forms in tag questions	1	1	1.04	1	1	1.92	-	-	-
6. Inserting redundant prepositions	28	22	22.92	18	13	25.00	10	10	22.73
7. Overusing certain verbs of high semantic generality	-	-	-	-	-	-	-	-	-
8. Replacing infinitive-constructions with that-clauses	3	2	2.08	3	2	3.85	-	-	-
9. Overdoing explicitness	-	-	-	-	-	-	-	-	-
10. present / past tense inversion	454	79	82.29	249	43	82.69	205	36	81.82
11. future tense *will/would* instead of *want*	8	7	7.29	7	6	11.54	1	1	2.27
12. Verb *to be*	62	39	40.63	37	24	46.15	25	15	34.09
13. Choice of preposition	224	79	82.29	123	47	90.38	101	32	72.73
14. full yes used as a discourse marker	34	-	-	-	-	-	-	-	-
15. Plural –s omitted	160	60	62.50	101	38	73.08	59	22	50.00
16. Lexical choice	807	93	96.88	424	50	96.15	383	43	97.73
17. Present perfect and simple inversion	24	15	15.63	5	4	7.69	19	11	25.00
18. Continuous form instead of infinitive	69	36	37.50	47	24	46.15	22	12	27.27
19. Infinitive instead of continuous	21	17	17.71	9	9	17.31	12	8	18.18
20. Word order omitted or redundant words	265	78	81.25	161	43	82.69	104	35	79.55
21. Adverbial position with auxiliary verbs	45	28	29.17	20	12	23.08	25	16	36.36
22. Omission of will future	16	14	14.58	10	8	15.38	6	6	13.64
23. Prepositional possessive phrase	17	12	12.50	10	5	9.62	7	7	15.91
24. Negation inconsistencies	53	31	32.29	27	18	34.62	26	13	29.55
25. *The* or *it* instead of personal pronoun	5	5	5.21	3	2	3.85	2	1	2.27
Total	2642	773		1486	446		1122	327	

As a reminder, the sub-corpora CEFR distribution Table 3.18 has been reproduced from Section 3.5.11 below. Although the CEFR distribution is similar, the table clearly shows that sub-corpus 2 has a higher percentage of B1 participants and thus proportionally lower levels of A2 and B2 level participants, while sub-corpus 1 has higher percentages of A2 and B2 participants.

Reproduced Table 3.18 from CEFR levels of English in sub-corpora

CEFR levels of English in sub-corpora

no. of part. in sub-corpus 1	percentage of cub-corpus 1	CEFR Level	percentage of cub-corpus 2	no. of part. in sub-corpus 2
15	28.85%	A2	22.73%	10
19	36.54%	B1	47.73%	21
14	26.92%	B2	22.73%	10
3	5.77%	C1	6.82%	3
1	1.92%	C2	0	0

Each of the 25 NSU categories coded within the corpus will be explored in depth in this section and where possible follow the pattern of first describing the usage, whereupon Swan (2016) was used as a comprehensive guide because his pedagogic grammar descriptions reflect the current use of English in Swiss classrooms. The description will be followed by the presumed didactical introduction of how or when the item was taught from grade three to nine. Then if available, pertinent research on the NSU category will be introduced. The findings will then be illustrated with examples. Uncoded versions of the examples have been used to ease reading. The examples are followed with a number in brackets e.g., <1>, <101>. Numbers ranging from 1-63 refer to sub-corpus 1 and numbers ranging from 101-145 refer to sub-corpus 2. After the examples, any differences between the sub-corpora will be highlighted. Each section is subsequently concluded with a brief summary of the findings.

As mentioned in Section 3.8, the general methodology used involved the initial analysis of investigating raw frequencies with AntConc 3.5.8 (Anthony, 2019). Next, the data was transferred to Microsoft Excel version 2011 for evaluation and classification. Proficiency levels were then checked to determine distribution between levels and sub-corpora. Where beneficial, a lemma list[24] was used to aid in the identification of the correctly used words in the corpus and ratios calculated or participant metadata was consulted to draw conclusions from the corpus data. The 1994 BNC spoken corpus was used as a reference corpus where needed.

To estimate when the participants were introduced to each of the features, syllabi were consulted to determine which material was used in the classroom. It was assumed that for the most part, the participants attended compulsory education in Central Switzerland. This implies that the participants in sub-corpus 1 were taught using typical L2 coursebooks from the seventh grade that include grammar explanations and exercises from the beginning of formal English instruction which would indicate a deductive style of teaching grammar and vocabulary. *New Inspirations* student books 1-3[25] (Garton-Sprenger and Prowse, 2011) were used as the obligatory English coursebook for grades seven to nine in Canton Lucerne from 2011[26]. Therefore, it is assumed throughout the analysis that they were used to teach the majority of the participants in sub-corpus 1. The scope and sequence of each book is available online and were consulted to determine when each type of NSU was taught.

---

[24] Someya Lemma List (no hyphens) was used, available at: https://www.laurenceanthony net/software/antconc/
[25] Books 1 and 2 were used over a three-year period before early English commenced. Presently several communities are using books 2 and 3 over a three-year period as the sequel to Young World book 4.
[26] Lehrmittelverzeichnis 2015/2016 Obligatorische und fakultative Lehrmittel an den Luzerner Volksschulen https://volksschulbildung.lu.ch,
https://volksschulbildung.lu.ch/unterricht_organisation/uo_faecher_lehrmittel/uo_fl_lehrmittel

Since the introduction of early English in the third grade, the obligatory course material in Central Switzerland has been for the most part the *Young World* 1-4 series (Arnet-Clark et al., 2005) for grades three to six and *Open World* 1-3 series (Fischer et al., 2011) for grades seven to nine. The series was especially written for Swiss primary and secondary schools. In the philosophy of Lehrplan 21, students are almost exclusively taught using an inductive method. This means grammar rules are acquired through a task-based approach and grammatical errors are not explicitly corrected in L2 instruction (Bildungs- und Kulturdepartement des Kantons Luzern, 2016, p. 9). Since the Young World books have been in use since the beginning of early English from 2005, it is assumed from this point on that the majority of the participants in sub-corpus 2 were taught using these books and teaching methods, although there can be exceptions. The author is aware that it is difficult to assure if and how the above-mentioned classroom material was used. Nevertheless, it is felt that any findings related to the material could indicate areas of interest which require further research.

Students who attended a Matura school would have had the same books and teaching methods in primary school and then depending on the individual schools most likely deductive teaching styles with typical L2 coursebooks supplemented with Cambridge exam preparation at B2/C1 level and an array of English literature (Kollegium St. Fidelis, 2020, pp. 60-68).

5.1.1 NSU-1 Omission or insertion of the third person present tense –*s*

Description

NSU-1 investigates omitting or inserting the third person singular -s. The third person singular present tense is formed by adding -s or -es to the base form of a singular verb. This can be challenging for learners of English because in all other persons the unchanged base form of the

verb is used. The use of the verb to be *is/are* and *was/were* as well as *have/has* and *do/does* used as full verbs were included in this category and considered irregular third person uses.

Classroom introduction

For sub-corpus 1, present simple was taught in *New Inspirations* book 1 in the third of eight chapters. For sub-corpus 2, the third person singular was first randomly used throughout the *Young World* books in text and stories, then taught in the second half of grade five with repetition exercises in year seven in *Open World*. As with the coursebooks used in sub-corpus 1, the online scope and sequences were consulted[27].

Previous research

In studies of developmental sequences in first and second language acquisition, grammatical morphemes, known as the smallest grammatical units, have been found to be mastered in certain sequences with the understanding that once a certain morpheme has been mastered those above it in the sequence have also been mastered (see Brown, 1973; de Villiers and de Villiers, 1973; Krashen, 1977). The third person singular is considered to be one of the later morphemes to be mastered in both first and second language acquisition and non-standard usages can still be found, however, less often in advanced learners (Mohammadkhani et al., 2011). Seidlhofer (2004) alleviates the pressure for learners by claiming that non-standard use of the third person singular is non-impeding and a feature of ELF. This stance ignores inaccuracy for the sake of understanding, which might or might not be favourable depending on the purpose of the communication.

---

[27] https://www.klett.ch

Findings

While analysing the data, 170 instances of a non-standard use of third person –s were found. A total of 39 words were used in a non-standard way. The words were evaluated to determine their relevance. Table 5.2 below illustrates the number of NSU in the whole corpus as well as both sub-corpora. The light blue shading in the first column indicates words used only once and the darker shading in the next two columns shows the words which were only found in one of the two sub-corpora. To evaluate the regularity of each word's non-standard usage, percentages of standard and non-standard usages were determined by calculating the amount of infinitive and third person singular -s uses of each of the 39 words. By determining the percentages, we can gauge the accuracy of the language produced, as can be seen in columns 5 and 6 in the table below. Here we see that 9 of the words were used in a non-standard way over 5% of the time and 5 words over 12% of the time.

Table 5.2 Words used in non-standard way NSU-1

Total No. of uses	Sub-corpus 1	Sub-corpus 2	Word	No. of Imperative and 3rd person -s	Percentage of NSU 1
1	1	0	speak/speaks	408	0.25%
3	2	1	are/is	1161	0.26%
4	4	0	think/thinks	964	0.41%
7	5	2	like/likes	1181	0.59%
2	2	0	do /does	272	0.74%
1	1	0	talk/talks	135	0.74%
13	5	8	was/were	1616	0.80%
1	0	1	use/uses	116	0.86%
1	1	0	work/works	98	1.02%
1	1	0	take/takes	77	1.30%
1	1	0	mean/means	72	1.39%
1	1	0	draw/draws	61	1.64%
3	1	2	see/sees	181	1.66%
8	5	3	go/goes	446	1.79%
3	3	0	understand/understands	161	1.86%

19	10	9	don't/doesn't	845	2.25%
1	1	0	read/reads	44	2.27%
6	3	3	say/says	259	2.32%
19	14	5	have/has	818	2.32%
2	2	0	come/comes	79	2.53%
1	0	1	end/ends	36	2.78%
1	1	0	tell/tells	33	3.03%
1	1	0	study/studies	32	3.13%
1	1	0	find/finds	28	3.57%
2	1	1	drive/drives	55	3.64%
1	0	1	buy/buys	25	4.00%
8	8	0	look/looks	185	4.32%
1	1	0	hear/hears	23	4.35%
12	6	6	want/wants	258	4.65%
8	4	4	make/makes	168	4.76%
4	3	1	live/lives	59	6.78%
6	2	4	show/shows	83	7.23%
2	1	1	ask/asks	26	7.69%
1	1	0	dance/dances	12	8.33%
6	6	0	try/tries	47	**12.77%**
13	4	9	paint/paints	72	**18.06%**
1	1	0	leave/leaves	5	**20.00%**
3	2	1	present/presents (verb only)	9	**33.33%**
1	1	0	fall/falls	2	**50.00%**
170	107	63	**Total**	10152	1.67%

Starting the investigation with the 17 words that were used once, we can see that 14 of them were used by participants in sub-corpus 1 and the majority are high frequency verbs. The corpus revealed that there were only 2 instances where participants made more than 1 of these singularly used NSU-1. This eliminates these NSU-1 being attributed to only a few participants. To determine if these NSU-1 were due to lower proficiency levels in sub-corpus 1, the proficiency levels of the 14 participants were examined. Sub-corpus 1 had 2 participants at A2 level, (n=4) at B1 level and (n=5) at B2 level. Sub-corpus 2 had (n=2) A2 level participants and (n=1) B2 level participant. Surprisingly, only 4 of the singular NSU-1 were used by participants with an A2 level, (n=4) with a B1 level and (n=9) with a B2 level. This preliminary finding

might suggest that the participants in sub-corpus 1 who were presumed to have been taught using a deductive teaching method used the third person singular less accurately. To be able to draw more precise conclusions on how the participants of the sub-corpus differ, the proficiency levels of all of the participants who used NSU-1 were examined. The results of sub-corpus 1 revealed that the distribution of NSU-1 was fairly evenly distributed between the A1, B1 and B2 levels with nearly 70% using at least one NSU-1. The NSU-1 in sub-corpus 2 on the other hand, were evenly distributed between levels A2 and B1, but at a higher ratio, see Table 5.3 below. These results indicate that the participants in both sub-corpora continue to struggle with third person singular -s accuracy at A2 and B1 CEFR levels including those with a B2 level in sub-corpus 1.

Furthermore, a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-1. There was a significant correlation between the two variables found, $r(5) = -.95$, $p = .010$.

Table 5.3 CEFR levels and use of NSU-1

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-1	Mean NSU-1 per participant
1	11	73.33	A2	36	3.27
1	13	68.42	B1	35	2.69
1	9	64.29	B2	35	3.88
1	1	33.33	C1	1	1.00
2	10	100.00	A2	32	3.20
2	19	90.48	B1	29	1.52
2	2	20.00	B2	2	1.00
Total	**61**			**170**	**2.78**

Next, the four most frequently used words were investigated; *don't/doesn't*, *has/have*, *paint/paints* and *was/were*. A total of 64 NSU-1 were used with these words. This is 38% of all NSU-1 used. As mentioned in Section 3.4, the picture description was used to encourage the participants to use similar words and structures which aided proficiency assessment. This is

evident because 55% of the NSU used with these four words were found in the picture description section. Examples of the most common uses are seen below.

(1)   painting her uh after he's finished uh she **don't** like it uh because the hair she **don't** like the hair <120>

(2)   the picture that he paint of her but she **don't** like it she thinks she looks not like that in real <55>

(3)   done it again and it takes really long because he **have** to make it better and better she was a <56>

(4)   because he uh **have** problems at home and so he <136>

(5)   a painter **paint** this girl on a Leinwand (canvas) yes and I think <137>

(6)   and yeah ah and the the party **were** very good it was on the beach <40>

(7)   the airplane from Zurich to London yeah we **was** in the hotel so uhm the name of the hotel was uhm oh uhm I think it was <39>

Of these examples, *she don't like* stuck out as being used 8 times by 7 different participants which would indicate that this expression was problematic. An internet search revealed that *she don't like* is being used in numerous popular music lyrics in a vernacular way including *She don't like the lights* (Bieber, 2012) where *she don't* is used 42 times within 4 minutes. All other instances of NSU-1 with *don't/doesn't* were investigated. Of the 19 instances, 17 were used with either he/she/it- *don't like, show, understand, look* or *matter*. Each of those combinations were similarly found in popular music lyrics. Answering the question if learning the standard form of *she doesn't* was negatively influenced by popular music is out of the scope of this thesis but would be an interesting investigative future path to follow.

As the third person singular was being investigated, an interesting pattern arose. Some words appeared not to be associated with other NSU when used in the infinitive or third person singular. For example, the words *dance, end, fall, find, leave* and *study* appeared in the corpus between 2 and 36 times and no other NSU were found beside the one non-standard usage of dropping or adding the third person singular –s. These six words were used by 6 different participants. Therefore, the initial assumption was that these words were well learnt or practised. To investigate further, the past tense and continuous form of these words were analysed and only one instance of *fall* being used in another non-standard way was found. Thus, the words *dance, end, find, leave* and *study* were only used once in a non-standard way and that was in the third singular form. Further patterns between the association among certain words and more than one non-standard usage were found.

The majority or twenty-six of the remaining thirty-four words also had non-standard usages in the past tense (NSU-10) as can be seen in Table 5.4 below, where the amount of past tense NSU-10 are highlighted. A non-standard ratio of over 5 % was found with 24 words and 13 words had a ratio of over 10% non-standard usages when total uses in the corpus were calculated. This indicates that these words are prone to be challenging for the participants in this study. In Table 5.4, the words are ranked from lowest percentage of NSU per number of words to the highest. The lowest percentages are marked in green and those over 10% are marked in orange. Words with irregular tenses are marked darker. Here we see that the distribution of irregular verbs is similar in the two groups.

151

Table 5.4 NSU-1 words used with other NSU

Total No of uses	Sub-corpus 1	Sub-corpus 2	Word	NN 7	NN 10	NN 11	NN 12	NN 15	NN 16	NN 17	NN 18	NN 19	NN 20	NN 21	NN 22	NN 24	total no of NSU	Total no words in corpus	NSU ratio to all uses in corpus
7	5	2	Like/likes						1					1			9	1208	0 7%
4	4	0	think/thinks		3			1						2			10	982	1 0%
1	1	0	Work/works											1			2	106	1 9%
1	0	1	Use/uses											1		1	3	138	2 2%
1	0	1	End/ends														1	40	2 5%
1	1	0	Study/studies														1	37	2 7%
3	2	1	Are/is		14		12		6		6	1	10				52	1823	2 85%
1	1	0	Take/takes			1									1		3	99	3 0%
3	3	0	Understand/understands		5							1					9	273	3 3%
1	1	0	Speak/speaks		4				1		1		6	3			16	463	3 5%
1	1	0	Find/finds														1	28	3 6%
2	2	0	Do /does	2	3				3				3		1		14	366	3 83%
13	5	8	Was/were		34		1		6	1	8		2	1			66	1616	4 1%
1	1	0	Mean/means		1								1				3	72	4 2%
1	1	0	Read/reads										1				2	44	4 5%
3	1	2	See/sees		6				2	3							14	254	5 5%
2	2	0	Come/comes		3				2				1				8	125	6 40%
4	3	1	Live/lives		1												5	71	7 0%
1	1	0	Talk/talks		5				5						1		12	169	7 1%
1	1	0	Dance/dances														1	14	7 1%
19	10	9	Don't/doesn't		19		3		1	1						18	61	850	7 18%
6	3	3	Say/says		14				6								26	340	7 6%
1	1	0	Fall/falls														1	12	8 3%
2	1	1	Ask/asks		3												5	59	8 47%
1	1	0	Hear/hears		1				1								3	35	8 6%
12	6	6	Want/wants		16									1			29	302	9 6%
1	1	0	Leave/leaves														1	10	10 0%
1	1	0	Tell/tells						5			1					7	70	10 0%
6	6	0	Try/tries								1						7	69	10 1%
19	14	5	Have/has		40		1		9	4			5	10		6	94	917	10 3%
8	8	0	Look/looks		7				5				1			2	23	222	10 4%
2	1	1	Drive/drives		7												9	68	13 2%
1	0	1	Buy/buys		4				1								6	45	13 33%
8	5	3	Go/goes		67				7			1	3	3			89	650	13 7%
6	2	4	Show/shows		5			1	3								15	109	13 8%
1	1	0	Draw/draws		12				1			1	1				16	73	21 9%
13	4	9	Paint/paints		5				3			3					24	89	27 0%
8	4	4	Make/makes	1	12				35				1				57	210	27 1%
3	2	1	Present/presents		1				1								5	12	41 7%
170			Total	3	292	1	17	2	104	9	15	7	36	24	3	27	710	12070	5.9%

The words *buy, draw, paint* and *present* all have high rates of NSU. Due to the nature of the interviews, many participants were inspired to use these words. The word *buy* was often used in the context of describing travel by the 67% of participants who choose topic two in part one of the interviews. Next, one of the two words *draw* and *paint* were necessary to describe the picture descriptions. *Paint* was used slightly more than *draw* to describe the pictures. The high rate of NSU at 21.9% and 27% might be an indicator of the collective abilities of the participants. In the Mohammadkhani et al. (2011) study of third person singular learner

152

accuracy, they found that intermediate learners had an average accuracy rate of 75%. This corresponds roughly with the average proficiency level of the CSC corpus participants. Lastly, the word *present* had a 41.7% ratio of NSU. It was used as a verb 9 out of the 12 times it was used. *Present* is very close to the German word *präsentieren* which might account for its use instead of the more common verb *show.*

Conclusion

In conclusion, all of the words used with NSU-1 are basic words that would be learned in an elementary coursebook at an A1/A2 level, but some are problematic, others not. Although these 39 words make up only 12.25% of the corpus, 26.87% of the total NSU are attributed to them. It is recommended that more attention should be given in the Swiss classroom to practice the third person singular with basic verbs if the goal is to advance beyond an intermediate level. It is not possible to decipher if Swiss students have trouble forming the third person singular in general or if the specific verbs needed to describe the picture sequence in this study were especially difficult. To find out it would be necessary to conduct further studies by either compiling a specialized corpus with similar tasks describing a picture sequence that would channel students to use certain verbs or compiling a larger general spoken corpus.

5.1.2 NSU-2 Confusing the relative pronouns who and which

Description

'Relative clauses are often used to modify nouns and some pronouns – to identify people and things or to give more information about them' (Swan, 2016, p. 233). When *who* and *which* are used in a relative clause, they are called relative pronouns. *Who* is used for animate objects such as people and *which* is used for inanimate objects such as things.

Classroom introduction

The sub-corpus 1 coursebooks cover the relative pronoun *who/that* at the beginning of grade eight then other relative pronouns including *which* towards the end of grade eight. The sub-corpus 2 coursebooks also taught relative pronouns at the beginning of grade eight. However, the participants would have had an additional four years of English instruction beforehand.

Previous research

Relative pronouns can cause trouble for L1 German speakers because there is no distinction between *who* and *which* in German (Swan and Smith, 2001, p. 47) for example:

English   I know the woman **who** lives next door.

German   Ich kenne die Frau, **die** nebenan wohnt.

English   The card, **which** was on my desk, is missing.

German   Die Karte, **die** auf meinem Schreibtisch lag, fehlt.

Confusing *who* and *which* is also on the list of ELF features that Seidlhofer (2004, p. 16) suggests is common among non-native speakers and does not hinder understanding. In search of verifying this commonality, Mollin found fewer than 1.2% of instances of *who* and *which* being used interchangeably in her 400,000-word Euro-English corpus (Mollin, 2006, p. 128). Rosenberger, on the other hand, found that 4.43% of *who* and *which* in his mainly spoken Swiss corpus were used interchangeably but after investigating came to the conclusion that they were errors, not possible features of Swiss English (Rosenberger, 2009, p. 147).

Findings

Fifteen participants, with various proficiency levels, used nineteen instances confusing the relative pronouns *who* and *which.* See Table 5.5 below.

Table 5.5 CEFR levels and use of NSU-2

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-2	Mean NSU-2 per participant
1	4	21.05	B1	5	1.25
1	3	21.43	B2	3	1.00
2	2	20.00	A2	3	1.50
2	5	23.81	B1	6	1.20
2	1	10.00	B2	2	2.00
Total	**15**			**19**	**1.26**

Seventeen were straightforward inversion of using *which* for people and *who* for things. Interestingly, of the nine instances where *which* was used instead of *who*, only one instance was used by a sub-corpus 1 participant and the remaining were used by sub-corpus 2 participants. The use of *who* instead of *which* was more evenly distributed with sub-corpus 1 participants using 5 instances and sub-corpus 2 participants using 3 instances. The remaining 2 instances from sub-corpus 1 participants were one instance where *with* was used instead of *which*, and *they* instead of *who*. The following are examples from the corpus:

(8)   we bought a ticket for a boat **who** drives uh down through the waterfalls yeah <139>

(9)   because uh English is uh is a language **who** most of the people uh talk a little bit so uh <37>

(10)   because you have the influencers **which** are here like for five years <145>

155

(11) and then the other people **which** are standing around the picture uh recognise this <118>

Conclusion

To sum up, there was a difference found between the sub-corpora with sub-corpus 1 participants showing greater command of the use of the relative pronoun *who.* Due to the small sample size, definite conclusions cannot be drawn. However, it would be interesting to investigate if the indicated separate teaching of the relative pronoun *who/that*, as specified above was a determining factor in the future accuracy of the relative pronoun *who*. Calculations revealed that 12% of all *who* and *which* uses as a relative pronoun were non-standard with the participants being slightly less accurate with *which*. This number merits further investigation. If over 10% of usages are confirmed to be non-standard, and it is known that relative pronouns are difficult for German speakers to master it is recommended that a more deductive approach earlier on would be beneficial.

5.1.3 NSU-3 Omitting definite and indefinite articles *a, an, the* where they are obligatory in Standard English

Description

Articles belong to determiners and are small words that are used at the beginning of noun phrases and help to show if something is known or unknown to the parties communicating, thus they are known as *definite* (the) or *indefinite* (a/an) (Swan, 2016, p. 133). Although article use is considered one of the most difficult points in English grammar (ibid.), Swan agrees with Seidlhofer (2004) and Cogo and Dewey (2006) that it does not impede understanding. Indeed, situational use such as a conversation versus a writing task might affect use. German also has

articles, but they are used differently. For example, they do not use zero article for general statements, nor are articles used with professions (Swan and Smith, 2001, p. 45).

Classroom introduction

There was no mention of articles being specifically taught in the scope and sequence of either sub-corpus coursebooks. Therefore, it is assumed that they were taught in context from the beginning. This could mean that the pupil's awareness of articles laid heavily on the individual teacher and their teaching methods.

Previous research

Omission of definite and indefinite articles has been well established in learner and second language English corpora. Research from the Swiss English Database by Rosenberger (2009, p. 198) and Dröschel (2011, p. 171) revealed that 63.5% of non-standard article use contributed to omitted articles with 2.29% of all articles by German L1 speakers in the corpora omitted with slightly more indefinite articles being omitted than definite articles. Furthermore, in a small-scale study of L2 university students, Dewey reports the omission of *the* in conjunction with *same* in 32% of all instances where he describes the omission as 'efficiency in ELF communication' (Dewey, 2009, pp. 64-65).

Findings

There were 49 NSU-3 instances found in the corpus which were made by 29 participants of various proficiency levels. Eighteen participants used only one NSU-3, whereas 11 used multiple. The majority (n=34) were indefinite articles omitted *a* (n=29) and *an* (n=5) the remaining 15 omitted were definite articles *the*. In total 38 NSU-3 were used by sub-corpus 1

participants and 11 were used by sub-corpus 2 participants. The following Tables 5.6 and 5.7 illustrate the proficiency levels and the number of NSU-3 used in indefinite and definite articles.

Table 5.6 CEFR levels and use of NSU-3 indefinite articles

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-3	Mean NSU-3 per participant
1	6	40.00	A2	11	1.83
1	5	26.32	B1	7	1.40
1	4	28.57	B2	4	1.00
1	1	33.33	C1	2	2.00
2	3	30.00	A2	4	1.33
2	4	19.05	B1	5	1.25
2	1	10.00	B2	1	1.00
Total	**24**			**34**	**1.41**

Table 5.7 CEFR and use of NSU-3 definite articles

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-3	Mean NSU-3 per participant
1	3	20.00	A2	4	1.33
1	2	10.53	B1	5	2.50
1	2	14.29	B2	3	1.50
1	2	66.67	C1	2	1.00
2	1	4.76	B1	1	1.00
Total	**10**			**15**	**1.50**
Total all NSU-3	**34**			**49**	**1.44**

It is interesting to see that participants with lower proficiency levels used more NSU-3 with indefinite articles, and NSU-3 with definite articles were used less often and were not concentrated at the lower proficiency levels. This suggests a correlation between proficiency levels and omitted definite and indefinite articles. To verify, a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-3. There was a significant correlation between the two variables found,

$r(5) = -.98$, $p = .002$. It is also worth mentioning that none of the NSU-3 were used by Matura students who presumably received the most classroom instruction.

Half of the omitted indefinite articles (n=17) were omitted before nouns and the other half (n=17) were omitted before adjectives as the following examples show:

Before proper nouns

    (12)   to tell them where I have learned to drive (**a**) car where I uh where I <131>

    (13)   he's pa he's (**a**) painter and my colleague Nadia <20>

Before adjectives

    (14)   like in ten seconds I was yeah its (**a**) real really nice family they had a <14>

    (15)   Cape Town and it was (**an**) interesting country yeah <120>

Of the 15 NSU-3 with the definite article seven were omitted before proper nouns, five before a measurement of time and three before directional areas (south).

Before proper nouns

    (16)   uhm yes in New York in (**the**) Bronx <6>

    (17)   with a vampire boy and eh in (**the**) film the characters are really smart <2>

    (18)   drew a second time and uhm (**the**) painter does this and <21>

Measurements of time

    (19)   a learner and uhm I hope in (**the**) future I will be <25>

Directional area

> (20)   no no I was in uhm France in (**the)** south of France in Cannes uhm because <24>

Conclusion

Indefinite articles were found to be omitted more often than definite articles with A2 level participants omitting indefinite articles more often. In general, sub-corpus 2 omitted fewer articles than sub-corpus 1, especially definite articles with only one omission. Furthermore, there were no omissions of *the* in connection with *same* found in the 77 instances of *the same* in the CSC. Thus, there was no evidence to support Dewey's (2009) observations. When all article use in the CSC corpus was calculated only 0.87% were found to have article omissions. Both the Swiss English Database and Dewey's corpus mentioned above were compiled from proficient speakers. However, the Swiss youth in the CSC appear to have been more accurate.

> 5.1.4   NSU-4 Inserting definite and indefinite articles *a, an, the* where they do not occur in Standard English

Previous research

The definition and timing of when articles are taught is the same as in the Section 5.1.3 above. Regarding the research on articles in Section 5.1.3, Dröschel (2011) and Rosenberg (2009) claimed that most non-standard article usages were omitted articles. The opposite was the case with the CSC where a total of 108 NSU-4 were found, which account for 69.8% of the non-standard usages of articles. Proficiency has been found to have an effect on the insertion of the definite article *the*. Lui and Gleason (2002) found that the insertion of *the* was less common in low-intermediate learners, increased at intermediate level then became less common again in proficient learners. Lui and Gleason (ibid.) claim that their findings support earlier research on

the acquisition of *the* and 'inflectional morphemes in L2 in general' (Liu and Gleason, 2002, pp. 15-18). This could possibly explain the abundance of definite article insertion in the CSC which is comprised of mainly intermediate learners.

Findings

In total 108 NSU-4 were found with the insertion of 11 indefinite and 97 definite articles were found. The indefinite article *a* was found inserted 11 times in a non-standard way in plurals, comparisons and when a zero infinitive would be expected for various reasons, examples follow:

Plurals

(21)   where she's a top model with **a** beautiful hair now and beautiful smile <138>

Comparative

(22)   yeah because the film is uhm **a** shorter than than uhm the book and the book you can you can see what the feelings are <2>

Zero

(23)   the city because it's always uhm [L] **a** friendly weather and the sea <55>

(24)   money uhm and uhm then I started **a** business school at <name of school> <1>

When we look at the proficiency levels of the participants who inserted an indefinite NSU-4 in Table 5.8 below, we see that most of the NSU-4 were made in sub-corpus 1 by 8 different participants which indicates 15.38% of sub-corpus 1 and 4.5% of sub-corpus 2 still have difficulty mastering indefinite article use. A Pearson correlation coefficient was computed to

161

assess the linear relationship between the participants' CERF proficiency levels and use of inserting the indefinite article *a* (NSU-4). There was not a significant correlation between the two variables found, r(5) = -.75, p = .139.

Table 5.8 CEFR levels and the insertion of the indefinite article *a*

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-4	Mean NSU-4 per participant
1	4	26.67	A2	4	1.00
1	1	5.26	B1	1	1.00
1	3	21.43	B2	4	1.33
2	1	4.76	B1	1	1.00
2	1	10.00	B2	1	1.00
Total	**10**			**11**	**1.10**

The definite article was inserted 97 times in a non-standard way as can be seen in the following examples.

Places 16 times

    (25)   that when you're in the uhm not in **the** Switzerland and you will have informations <27>

    (26)   so we went to pubs and uhm visited **the** Stonehenge and everything you you do <14>

Institutions/schools 15 times

    (27)   uhm because I must learn this in **the** school <27>

    (28)   and I was walking it was very different to **the** school then you sit down in school and came home in the train and in the evening <1>

Collective nouns people 12 times

(29)   I can write a text good but when I speak with uh with **the** people then I uh forget

lots of words <129>

(30)   then is it more important you can talk with **the** people and I think <1>

Times and numbers 11 times

(31)   then we talked also English but after **the** time she learned Swiss German <1>

(32)   [L] yeah maybe at **the** first I will be a learner <23>

Superlatives 10 times

(33)   but with uhm German subtitles but **the** most is in German also <18>

(34)   so but in Switzerland also it's like **the** most people speak English <32>

Life 6 times

(35)   very good and cheap yeah I think **the** life there would be cheaper <23>

(36)   another yeah perspective to **the** life and yeah <136>

With a language 5 times

(37)   live here in Switzerland and they speak **the** Swiss German I think <127>

(38)   uhm you can learn more by using **the** English yes I think for me <5>

Miscellaneous 22 times

(39)   and chicken noodles [L] yes **the** normal food [L] yes <17>

(40)   yeah so **the** I don't think I go <39>

Table 5.9 below illustrates the CEFR levels and NSU-4. Here it is apparent that the CSC results do not correspond with Lui and Gleason's (2002) findings nor the results from the Swiss English Database. On the contrary, there is not only an increase in the intermediate participants but throughout the sub-corpora with 63% of the sub-corpora 1 participants and 30 % of the sub-corpus 2 participants inserting the definite article *the* at least once. The results of insertion of definite articles were, however, influenced by 2 participants. One participant in sub-corpus 1 used 9 NSU-4 and 1 participant in sub-corpus 2 used 11. Both of these participants had a high number of total NSU. Respectively, 5.71 phw and 6.11 phw. However, they did not use excess NSU with indefinite articles only the insertion of *the*. Furthermore, a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and insertion of the definite article *the* (NSU-4). There was a significant correlation between the two variables found, $r(5) = -.93$, $p = .019$.

Table 5.9 CEFR levels and insertion of the definite article *the*

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-4	Mean NSU-4 per participant
1	9	60.00	A2	16	1.77
1	11	57.89	B1	30	2.72
1	11	78.57	B2	20	1.81
1	1	33.33	C1	1	1.00
1	1	100.00	C2	1	1.00
2	6	60.00	A2	19	3.16
2	5	23.81	B1	8	1.60
2	1	10.00	B2	1	1.00
2	1	33.33	C1	1	1.00
Total	**46**			**97**	**2.10**
Total NSU-4	**56**			**108**	**1.92**

Conclusion

To sum up the omission and insertion of articles, the CSC corpus had over twice as many insertions than omissions. This is contrary to results from research of proficient users where the

opposite was found (Rosenberger, 2009; Dröschel, 2011; and Dewey 2009). The findings in the CSC corpus do not correlate with research regarding the acquisition of article use in intermediate learners (Liu and Gleason, 2002). It was found that sub-corpus 1 participants omitted and inserted more articles than sub-corpus 2. Although there was no indication in the scope and reference of the coursebooks when or how articles were taught, we know that a deductive approach was used which would be in line with earlier research (Pica, 1986) that showed that 'instructed learners tended to overuse grammatical morphemes more than naturalistic and mixed learners' (Liu and Gleason, 2002, p. 17). Thus, the increased insertion of articles in sub-corpus 1 might be attributed to the deductive teaching approach. In comparison with all the instances of articles in the CSC corpus, 1.91% were non-standard insertions.

5.1.5 NSU-5 Failing to use correct forms in tag questions (e.g., *isn't it?* or *no?* instead of *shouldn't they?*)

Description

Tag questions or question tags are small questions that come at the end of sentences. They are found mostly in spoken informal discourse and are used to ask for agreement or confirm truth. They fall under the umbrella category of discourse markers and are formed by adding the opposite tag to an affirmative or negative sentence (Swan, 2016, p. 305). Two categories of question tags are suggested by Baker (2015 p. 315), canonical question tags as above and invariant tags. The latter are formed by adding single words such as, *right*? *okay*? *yeah*? or even *huh*? and *innit*? to the end of an utterance. The same words can be used regardless of the components of the main sentence.

Classroom introduction

Sub-corpus 1 participants were taught question tags with the verb *to be* in the middle of grade seven and then general question tags at the end of grade nine. Sub-corpus 2 participants, on the other hand, were only taught question tags at the beginning of the eighth grade or in their sixth year of English study.

Previous research

It has been observed by Tottie and Hoffman (2006, p. 304) that canonical question tags are used far less by under 20-year-olds than older people. However, their study did not include non-native speakers. Baker (2015, p. 315) suggests that younger and L2 English speakers use more invariant tags and the use of local language tags with English is on the rise. Both will be verified below. Firstly, although they were taught, the CSC corpus is void of canonical question tags except for one example (see below) and it was used in a non-standard way.

>   (41)   or what do you say yeah is it a printer by a by a uhm printer [L] no printer is something on the computer **wasn't it** <5>

Findings

Although canonical question tags are scarce, the CSC does have a plethora of invariant tags. For example, 194 of the 502 *yeah* in the corpus are at the end of a full sentence turn and are potential invariant tags. The corpus was searched for *yeah*. Of the 1,858 instances, 502 were found at the end of a turn. Of those, 271 were single- or double-word answers to a question. The remaining 231 turns were evaluated, and it was determined that 37 turns were direct answers to a question which left 194 turns with yeah as the closing word. For example:

(42)   for example sausages and chips fish and chips such things **yeah** <103>

(43)   I think she's proud of it I think she looks beautiful on it **yeah** <102>

This type of invariant tag was used by 68 of the 96 participants which equates to 70% and was divided equally between the sub-corpora. A similar pattern was found when *yes* was investigated at the end of a turn. Of the 468 instances of *yes* at the end of a turn, 338 were found to be one- or two-word answers to a question and a further 43 were considered longer answers to a question which left 87 instances as possible invariant tags. Only 43 of the 96 participants or 45% used these; however, it is considerable and strengthens the statement that these are broadly used by young adults.

At a closer investigation of an exchange, we see that *yeah* is used as a discourse marker to open two phrases in Example 1, then again in 3 and 5. However, when *yes* and *yeah* are used at the end of 3 and 5, they act as an invariant tag or final affirmation of information given.

(44)   <B11> for two weeks with my father and this this was very impressive and I was in Las Vegas and in Los Angeles and San Francisco yeah and it's very different to here uhm they all drive big cars and and a lot of McDonalds and so and **yeah** I found it very interesting there **yeah** I liked it </B11>

(45)   <A11> uh huh would you like to go back </A11>

(46)   <B11> **yeah** I think so **yeah** maybe for a year or something like that I I'd like to go there again **yes** </B11>

(47)   <A11> uh huh which part of America would you like to go to then </A11>

(48)   <B11> **yeah** California I like very **yeah** </B11>

An interesting observation was made that 8 participants used yeah and 2 participants used yes to finish a turn after they had code-switched into German. See below:

(49)   I'm very <G> wie seit mir gfallt es=3 </G> yeah <6>

(50)   I I cannot uhm build uh <G> ä Satz oder Frage=4 </G> yeah <29>

(51)   things uhm a little <G> es Dorf Dorf mir falts nit ih=4 </G> yeah <17>

(52)   uhm <G> ich muss überlege=4 </G> yeah <21>

(53)   and the electronic and all the things <G> het sich entwicklet=4 </G> yes <39>

There were also cases of three German invariant tags found with multiple meanings: *oder* (or, right, innit), *also* (so, well, thus) and *aber* (but, however, though). Out of the 6 *oder*, 5 were used at the end of a turn. *Aber* was used only once, but at the end of a turn and 5 of the 143 uses of *also* in German were used at the end of a turn. This strengthens the argument that L2 invariant tags are being used when speaking English and shows the use of the English *yeah* mixed with L1 code-switching.

(54)   an office job <G> aber </G> <3>

(55)   cuz he is he was uh so a woodman <G> oder </G> <133>

(56)   yes it's a kind of architect but something above down <G> also </G> <116>

Conclusion

Although *yeah* and *yes* dominate the CSC corpus, 3 instances of *okay* and 1 instance of *right* used as invariant tags were also found. It can be concluded that the participants in the CSC corpus effectively used invariant tags to signal presence of the listener in line with Baker's observation that invariant tags 'have no pragmatic function beyond simply acknowledging the presence of the listener' (Baker, 2015 p. 316). However, the CSC data could imply that invariant tags are also being used to draw the listener back into the conversation and give final verification of an affirmative answer as in the dialogue example above. It is noteworthy that 12.16% of all *yeah* and *yes* in the corpus were potentially used as invariant tags.

5.1.6 NSU-6 Inserting redundant prepositions, as in "*We have to study about…*"

Description

Prepositions are once again an aspect of Standard English which is difficult to learn. Firstly, most prepositions have several functions, and they may or may not correspond to foreign languages. This is also true of fixed expressions which cannot be guessed and must be learned (Swan, 2016, p. 209). Most German prepositions are similar to English; however, problems occur when fixed expressions and multiple meanings exist (Swan and Smith, 2001, p. 47).

Classroom introduction

In the sub-corpus 1 coursebooks, prepositions of place were taught first at the beginning of grade seven and prepositions of time several months later. The sub-corpus 2 coursebooks introduce prepositions of place in the middle of grade three, then four years later at the beginning of grade seven prepositions of direction and location are taught again.

Previous research

NSU-6 is one of Seidlhofer's (2004, p. 220) original ELF features which she observed. Although the Oxford English Dictionary (OED, 2021) definition of *redundant* is: superfluous, excessive; surplus and unnecessary, ELF advocates Cogo and Dewey (2012, pp. 52-61) seem to concentrate on describing a shift or transition away from native speaker use of prepositions with emphasis on the omission not insertion of prepositions. Furthermore, they claim that redundancy can occur when a pattern which is standard is duplicated in a non-standard way (Cogo and Dewey, 2012, p. 56). Önen (2014) explored both and examples of omission will be given in Section 5.1.14.

Findings

Initial coding of redundant prepositions was less fruitful than expected. However, during the NSU finalisation process, more redundant prepositions were found to total 28 instances. Redundant prepositions found included five times *at*, five times *for*, seven times *in*, five times *of,* two times *on*, three times *to* and one time *with*. Most of the instances exemplify typical learner problems either because of the difficulty of learning all the exceptions and particular expressions with prepositions, L1 interference or just misspeaking. When all instances of *at, for, in, of, on, to* and *with* in the CSC were calculated against NSU-6 only 0.43% that were used in a non-standard way were found to be redundant.

The following examples from the corpus are of common learner problems where prepositions are not used in Standard English.

> (57) she looked at the picture… and then she uhm took the picture **at** home and showed it to her friends <59>
>
> (58) I would say five thousand for more dollar **in** a month <3>
>
> (59) we and the ate just one uhm one time **on** a day <40>

In the following phrases L1 interference is prominent. In the first phrase 'what for a beautiful picture' is a direct translation of the German 'Was für ein schönes Bild'. In phrase two the participant has used both the German and its English equivalent *vor* or *für* and *ago*. The third phrase appears to be a phonetic error 'sprechen auf Englisch'.

> (60) so he's lucky and can tell her friends what **for** a beautiful picture she has <135>
>
> (61) my ancestors are from Albania but **for** 300 years ago <127>
>
> (62) because uhm here in Switzerland you don't speak much **of of** English of course I have uh some English <33>

Lastly, misspeaking is a common occurrence in all spoken discourse, such as the examples below.

(63)   I could (uhm) put **on** my fingers on the koala [L] that's really cute <56>

(64)   I like it very much the culture because of the nature it's very dry **in** it doesn't rain <119>

As can be seen below in Table 5.10, thirteen participants from sub-corpus 1 used NSU-6 eighteen times and sub-corpus 2 participants used almost half with ten NSU-6 used by ten participants. The percentage of participants who used NSU-6 was evenly distributed at slightly above 30% except for the B2 levels and A2 levels in sub-corpus 2. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-6. There was not a significant correlation between the two variables found, $r(5) = -.67$, $p = .211$.

Table 5.10 CEFR and use of NSU-6

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-6	Mean NSU-6 per participant
1	5	33.33	A2	5	1.00
1	6	31.58	B1	8	1.33
1	1	7.14	B2	3	3.00
1	1	33.33	C1	2	2.00
2	2	20.00	A2	2	1.00
2	7	33.33	B1	7	1.00
2	1	10.00	B2	1	1.00
Total	**23**			**28**	**1.21**

Conclusion

The insertion of redundant prepositions in the CSC were concluded to be common learner errors and German L1 interference with fewer instances found in sub-corpora 2. Instead of an

abundance of the insertion of redundant prepositions as initially suggested by Seidlhofer (2004, p. 220) when analysing ELF discourse, the CSC participants had a total of 224 instances of non-standard choice of prepositions. This will be discussed in Section 5.1.14.

5.1.7   NSU-7 Overusing certain verbs of high semantic generality, such as *do, have, make, put*, *take*

Description

First described as general purpose verbs by Clark (1978), the verbs *do*, *go*, *make*, *get*, *put* and *take* can be used in place of more specific verbs usually associated with a higher level of proficiency. These are the verbs that both native speaker children and L2 learners first produce and use until their lexis expands.

Classroom introduction

Both coursebook series for sub-corpus 1 and 2 publish wordlists with the German translation, pronunciation, and the new vocabulary word in an example sentence. Tallying up these wordlists gives an estimate of the number of words that the participants were exposed to in the classroom. The sub-corpus 1 coursebooks 1-2 expose learners to 2,368 words over a three-year period. The sub-corpus 2 coursebooks, on the other hand, expose pupils in primary school to 2,210 words and an additional 1,350 words in secondary school to total 3,560 words by the end of ninth grade or after seven years of study.

Findings

Only three instances of overusing high semantic words were coded during transcription. All three of these could also have been tagged as NSU-16 lexical choices. It was therefore decided to investigate the overuse of high semantic verbs within the broader area of lexical choices

because it was deemed more meaningful to investigate their lexical diversity than overuse. Table 5.11 gives an overview of the total amount of NSU which involve high semantic verbs. The only possible interesting word to investigate as being overused would be *make*. However, it was found to have been used in a non-standard way instead of the other high semantic verbs *do*, *have*, *want*, *learn*, *visit*, *take*, *see* and *do*. See Section 5.1.17 for further information about non-standard lexical choices.

Table 5.11 Non-standard use of high semantic verbs

High semantic verb	Number in Corpus	Total number of NSU
go	650	90
make	231	60
do	1121	69
put	22	2
take	99	6

5.1.8   NSU-8 Replacing infinitive constructions with *that* clauses, as in *I want that*

Previous research

Replacing infinitive constructions with *that* clauses is a feature which Seidlhofer (2004, p. 220) presented as a lexicogrammatical ELF feature. Although much research has been done, little knowledge has been gained about this hypothesis. Cogo and Dewey re-evaluated the hypotheses made by Seidlhofer and reformulated the lexicogrammatical features as language choices instead of deficits (Cogo and Dewey, 2006).

- Use of 3rd person singular zero
- Extension of relative which to include functions previously served only by who
- Shift in the use of articles, (among other patterns this involves preference for zero article where L1 article use is largely idiomatic, and preference for definite article to attach extra importance to a referent in a stretch of discourse)
- Invariant question tags (and use of other similar universal forms, such as this for this and these)

- Shift in patterns of preposition use, e.g., we have to study about
- Extension to the collocational field of words with high semantic generality, e.g., take an operation
- Increased explicitness, e.g., how long time in place of how long
  (Cogo and Dewey, 2006, p. 75)

Findings

The following 3 cases of possible NSU-8 replacing infinitive constructions with that clauses were tagged during transcribing but are not considered relevant. A further concordance line search of the 1,131 cases of *that* did not conclude in finding any other cases of NSU-8.

    (65)   but (uhm) my parents said you have to do **that** then I had to <1>

    (66)   then I had to do **that** and (uhm) yes I did it <1>

    (67)   book much better than the film still **that** I like it <2>

Conclusion

The EFL feature of replacing infinitive constructions with *that* clauses was not found to be a feature of the Swiss English represented in this study.

5.1.9 NSU-9 Overdoing explicitness (e.g., black colour instead of just black)

Previous research

This is the last of Seidlhofer's (2004, p. 220) original ELF lexicogrammatical features. There were no instances found when transcribing the interviews. It is possible that there were some instances that were classified as a non-standard lexical choice (NSU-16) or non-standard word order (NSU-20). However, overdoing explicitness has not developed into a prominent research area within the EFL research community and was therefore not perused further in this study.

Conclusion

An attempt was made to explore the original nine lexicogrammatical features of ELF as hypothesised by Seidlhofer (2004, p. 220), but it was not fully possible due to the rarity of some of the features in the CSC corpus. At present 'knowledge of grammatical variability in ELF remains superficial' (Laitinen, 2020, p. 427). Until more in-depth grammatical analysis is conducted as Laitinen (ibid.) suggests, comparison with learner data remains challenging.

5.1.10 Summary of original lexicogrammatical features of ELF findings

To sum up thus far, of the originally hypothesized ELF features NSU-1 to NSU-9, the zero third person singular (NSU-1) stands out as being the most widely used by the CSC participants. Three-quarters of the verbs used with NSU-1 were also found to be correlated with the use of NSU-10 (simple and past tense inversion). Furthermore, a third of the verbs used with NSU-1 were found to be used at a non-standard rate over 10% and it was indicated that this feature is widespread throughout the corpus. If increasing the accuracy of third person singular is a pedagogical objective, the words indicated as problematic could offer directional guidance.

The next widely used feature was the insertion and omission of indefinite and definite articles, NSU-3 and 4. Starting with the omissions (NSU-3), it was found that the omissions of indefinite articles were more numerous and were proficiency bound at the lower proficiency levels, whereas definite article omission was less frequent and not proficiency bound. Turning to the insertion of articles (NSU-4), indefinite articles were inserted considerably less frequent than definite articles. Overall, however, twice as many insertions of articles over omissions were found which contradicts previous research by Dröschel (2011), Rosenberg (2009) and Liu and

175

Gleason (2002). The German L1 of the majority of the CSC likely contributed somewhat to this outcome.

Less extensively used, the insertion of redundant prepositions (NSU-6) and confusion of relative pronouns (NSU-2) can generally be attributed to common learner errors or L1 German interference. With only one example of a canonical question tag in the CSC, NSU-5 was initially considered not relevant. However, an abundance of invariant tags was found with *yeah* and *yes* which indicated that invariant tags were being used to draw the listener back into conversations and give final verification of an affirmative answer.

The last three NSU associated with ELF features were found to be less relevant. First, the overuse of high frequency verbs (NSU-7), was not observed as a feature because substitution of one high frequency verb for another and lexical variation was found more common. Therefore NSU-7 related instances were handled under lexical choice in NSU-16. Lastly, *that* clauses (NSU-8) and overdoing explicitness (NSU-9), were almost non-existent and considered irrelevant.

In terms of sub-corpora differences between the original lexicogrammatical features of ELF, in general, fewer NSU were found in sub-corpus 2 and NSU use in sub-corpus 1 was spread over a wider range of CEFR levels than in sub-corpus 1. The exception was that sub-corpus 2 used more instances of NSU-2 and *which* was used instead of *who* almost solely by sub-corpus 2 participants. This suggests the early English starter participants from sub-corpus 2 were overall more accurate.

The following fifteen NSU sections could be considered learner language inconsistencies with Standard English norms. They follow the same structure as the previous NSU sections with a

brief description of the feature followed by when it was introduced in the classroom, previous research if available and corpus findings. A brief conclusion can be found at the end of each NSU section and a further summary in Section 5.2.

5.1.11 NSU-10 Present simple / past tense inversion

Description

'Verbs are the central element in sentences. Most of the things are built around them' (Swan, 2016, p. 1). Verbs have a base form, third person singular, past tense, past participle and *-ing* form. The present simple is used to express permanent situations such as facts and routines, whereas the simple past expresses past events which are finished.

Classroom introduction

In the sub-corpus 1 coursebooks, present simple is the first item covered and is followed by present continuous towards the end of grade seven. The simple past is not covered until the beginning of grade eight and then followed by repetition of the present simple, continuous and past simple again before the perfect tense is introduced towards the end of grade nine.

The sub-corpus 2 coursebooks seem to take a scaffolding approach where the tenses are introduced early and then reviewed repeatedly until the various aspects have all been covered. In grade three present simple and present continuous are introduced, in grade four present continuous is repeated and in grade five present simple is repeated before the past tense and its irregular verbs are introduced. Finally, past tense regular verbs are covered in grade six. In secondary school in grade seven and eight, all of the tenses are repeated several times again before the passive voice is introduced in grade nine.

Previous research

Learners are generally taught present tense before the past tense with its many irregular forms that need to be memorised or internalized before they are mastered. Thus, when speaking, learners might revert back to what they first learned, apply the past simple *-ed* rule or be at various stages of developmental sequences. The past tense belongs to the last section of second language grammatical morpheme acquisition sequence according to Krashen (1977). Research by Bardovi-Harlig (2000) about past tense found that learners are more likely to use past tense with verbs that express a tangible achievement or accomplishment as opposed to a state. Hence it is believed that the *lexical aspect* of stative and dynamic verbs influences the verb tense choices of learners of English (Lightbown and Spada, 2006, p. 92).

In Swiss research, Rosenburg (2004, p. 162) dismissed past/present tense inversion in the Swiss English Database as too infrequent to investigate, whereas Dröschel (2011, p. 210) argued that the use of zero past tense marking, although infrequent, could be an indicator of new English varieties.

Findings

A total of 454 instances of a non-standard use of present simple and past simple inversion were found which were made by 79 of the 96 participants with almost equal percentage of the sub-corpora participants 82.69% from sub-corpus 1 and 81.81% from sub-corpus 2. In total 78 verbs were used in a non-standard way. The 36 regular verbs used accounted for 90 NSU-10 with sub-corpus 1 participants using 60 instances and sub-corpus 2 using only 30. The remaining 359 NSU-10 were used with 42 irregular verbs and more evenly distributed between the sub-corpora with 185 NSU-10 in sub-corpus 1 and 174 NSU-10 in sub-corpus 2.

Looking closer at the regular verbs, in all but three instances the present simple was used instead of the past simple and 21 verbs were used only once. To investigate the influence of lexical aspect, stative verbs were isolated. The five stative verbs *want* (n=16), *look* (n=7), *need* (n=2), *like* (n=1) and *wish* (n=1) were found and accounted for 27 or almost one third of the NSU-10 uses with only six uses made by sub-corpus 2 participants. To determine if the CSC participants were hesitant to use stative verbs in the past tense, all uses of the five verbs above in the CSC corpus were investigated and the percentage of present and past tense use calculated. Then they were evaluated and compared with the 1994 BNC spoken corpus. Table 5.12 below shows that there was no hesitation to use stative verbs in the past tense when the use of the verbs in question were compared to a native speaker corpus. On the contrary, the similarities prove that the participants' percentage of past simple tense use closely resembles standard use, although the use of *wish* is disproportionate because of the low frequency of five instances.

Table 5.12 Percentages of regular stative verbs used in past tense in CSC and BNC

Verb	Percentage of past tense in CSC	Percentage of past tense in BNC
want	15.03%	15.05%
look	18.8%	13.41%
need	7.89%	8.25%
like	2.17%	1.38%
wish	**20%**	8.2%

Next, all NSU-10 verbs which were used more than three times were investigated further to analyse their potential importance. The frequency of all present simple and past instances of the words were calculated with the amount of NSU-10 to discover the rate of non-standard use. As can be seen in Table 5.13 below, the percentage of all of the verbs calculated were above the corpus average rate of 2.68% with six out of ten higher than 5%. This indicates that the participants, particularly those from sub-corpus 1, were less accurate with these verbs.

Interestingly all 4 NSU-10 made with *we rent a car* were from different participants when recounting past travel experiences. Thus, this chunk has possibly been learned incorrectly because there is only one instance of *rent* in the past tense in the entire corpus.

Table 5.13 Percentage of non-standard use of past in regular verbs

Frequency	Amount of NSU-10	Word	Percentage of NSU-10
207	7	look	3.38%
126	5	talk	3.96%
387	16	want	4.13%
107	5	show	4.67%
56	3	ask	5.35%
54	3	start	5.55%
89	5	paint	5.61%
43	3	change	6.97%
51	6	visit	11.76%
9	4	rent	44.44%

Now turning to the irregular verbs, 42 verbs were used. There were 17 examples of the use of past instead of present and 340 where present was used when a past tense would be used in Standard English. There were seven instances of the regular *-ed* being used instead of the irregular as can be seen below in the corpus examples:

in the air (uhm) some also we	**builded**	and all (uhm) after	sub-corpus 2
to write back (uh) Iron Man	**catched**	the shield because his reflexes	sub-corpus 2
to sit nine hours straight and I was never	**flyed**	before that long	sub-corpus 2
and yes it's wonderful there I	**goed**	there every year	sub-corpus 1
he's from Thailand or something like this and he	**speaked**	really funny English	sub-corpus 1
everybody understands me whenever I	**speaked**	English (uh) forget about the	sub-corpus 2
one day we (uhm)	**waked**	up early in the morning to	sub-corpus 2
no no I already (uh)	**writed**	the story in my	sub-corpus 2

It is quite interesting to see that all but two of NSU-10 where the *-ed* rule was used on irregular verbs were used by 4 different sub-corpus 2 participants. This is a small data sample; however, it clearly indicates that 9.09% of sub-corpus 2 participants still produce the error of applying the *-ed* rule with irregular past tense verbs compared to only 3.84% from sub-corpus 1.

Irregular stative verbs were also investigated, and the following nine were found: verb to be including *is, am, are, was* and *were* (n=52), *know* (n=8), *mean* (n=1), *feel* (n=1), hear (n=1), *have* (n=40), *see* (n=10), *think* (n=4), and *understand* (n=9). These stative verbs constituted 37.05% of the NSU-10 with irregular verbs. In Table 5.14 below we can see that the use of irregular stative verbs does not align with native speaker use as the regular verbs did. Differences between the corpora were found in the use of *see* and *understand* because they were used interchangeably in the 1994 BNC spoken corpus to convey understanding such as in 'I see' and in other ways such as 'let's see' and 'see how it goes'. In the CSC, however, *see* was used mainly to covey the meaning of sight.

Table 5.14 Percentages of irregular stative past verbs in CSC and BNC

Verb	Percentage of past tense in CSC	Percentage of past tense in BNC
to be	**56.82%**	51.21%
know	**4.72%**	2.92%
mean	0%	**4.39%**
feel	18.36%	**22.74%**
hear	34.28%	**52.25%**
have	22.31%	**24.94%**
see	**24.89%**	8.45%
think	6.12%	**18.09%**
understand	**42.58%**	10.27%

The remaining 34 verbs were dynamic. While the irregular verbs occurred more often, only 13 occurred just once. Again, the frequency was calculated of the verbs occurring more than three times to determine which verbs posed the most difficulty to produce accurately.

Table 5.15 Frequency of dynamic verbs and NSU-10

Frequency	Amount of NSU-10	Word	Percentage of NSU
1013	4	think	0.39%
555	6	speak	1.08%
675	8	can	1.19%

181

529	8	know	1.51%
2777	52	verb to be: is, am, are, was, were	1.87%
854	19	don't	2.22%
263	10	understand	3.80%
1053	40	have	3.80%
131	5	get	3.82%
233	10	see	4.29%
323	14	say	4.33%
151	8	can't	5.30%
180	12	make	6.67%
60	7	drive	**11.67%**
545	68	go	**12.48%**
61	8	eat	**13.11%**
42	6	buy	**14.28%**
18	4	sit	**22.22%**
16	4	build	**25.00%**
17	5	fly	**29.41%**
51	19	draw	**37.25%**

Table 5.15 above illustrates that the majority of verbs used in a non-standard way were above the corpus average of 2.68% with the high frequency verbs *think, speak, can, know* and the *verb to be* being used quite accurately as opposed to *drive, go, eat, buy, sit, build, fly* and *draw* which are basic verbs learned very early on.

In Table 5.16 below, observations of the total number of NSU-10 in both sub-corpora show that the distribution is similar, with sub-corpus 2 CEFR level B2 participants using 10% fewer NSU-10 as their counterparts in sub-corpus 1 and 6% more in level B1. Furthermore, a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-10. There was a significant correlation between the two variables found, $r(5) = -.91$, $p = .031$.

Table 5.16 CEFR and use of NSU-10

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-10	Mean NSU-10 per participant
1	15	100.00	A2	92	6.13
1	16	84.21	B1	106	6.62
1	10	71.43	B2	49	4.90
1	2	66.67	C1	2	1.00
2	10	100.00	A2	83	8.30
2	19	90.48	B1	101	5.31
2	6	60.00	B2	20	3.33
2	1	33.33	C1	1	1.00
Total	**79**			**454**	**5.74**

Conclusion

Over 80% of the corpus participants used NSU-10. This indicates that past simple verb tense accuracy has not been reached by the majority of the participants and that the highlighted verbs should be taken into consideration by language teachers when teaching and testing. The data shows that an inductive teaching approach had a positive effect on the accuracy of regular past tenses with sub-corpus 1 producing twice as many as sub-corpora 2, but not on the accuracy of irregular verbs. The collection of extensive metadata allowed for further investigation of the language background of the 17 participants who did not use any NSU-10.

Table 5.17 Language spoken at home and CEFR levels of participants with no NSU-10

Sub-corpus 1		Sub-corpus 2	
Languages at home	CEFR level	Languages at home	CEFR level
Swiss German only	B1	2 languages	B1
Swiss German only	B1	3 languages	B1
3 languages	B1	Swiss German only	B2
Swiss German only	B2	Swiss German only	B2
Swiss German only	B2	Swiss German only	B2
2 languages	B2	3 languages	B2
2 languages	B2	2 languages	C1
2 languages (English)	C1	2 languages (English)	C1
3 languages (English)	C2		

As a reminder, 58.3% of participants recorded speaking only Swiss German at home, 30.2% 2 languages and 5.3% 3 languages. Here we see in Table 5.17 above that ten of the participants

with no NSU-10 speak multiple languages at home and seven speak only Swiss German. This implies a correlation between past tense accuracy and regular use of multiple languages at home. The absence of A2 level participants reinforces the general understanding that lower level CEFR levels have lower levels of accuracy with the past simple.

5.1.12 NSU-11 Future tense *will/would* instead of *want*

Description

In Standard English, the future can be projected in several ways; with the simple present, present continuous, *going to* and *will/shall*. The latter can cause confusion in German L1 speakers because *will* in German literally translates to *want*.

Classroom introduction

The sub-corpus 1 coursebooks first cover *going to* future in the middle of grade eight and repeat it at the beginning of grade nine right before covering *will* future and the *present continuous* future in the next unit. The sub-corpus 2 coursebooks do not cover future until secondary school in the middle of grade seven where *will* future is taught and then *going to* future is taught at the end of grade seven. In grade eight, future with present continuous is introduced right before *will* and *going to* are repeated. At the end of grade nine, advanced points of *going to* future complete this tense/aspect.

Findings

There were eight instances found of using *will* or *would* in a non-standard way. The first seven are clearly a case of L1 interference where the German *will,* which translates into *want* in English (Swan and Smith, 2001, p. 43) was substituted for *want* or *want to*. The last example appears to be a redundant insertion of *will*.

picture and she find it not so good she she uhm	**will**	that painter he drew a second time and uhm
when they are the best they can win 6,000 euros and she	**will**	win this because she will have new breasts [L]
can win 6,000 euros and she will win this because she	**will**	have new breasts [L] for a man yes and this is very lustig
maybe and then we goes uhm to animals we	**will**	shoot some animals there and uhm I also I love this
he draws uhm the … the young woman I think she uhm	**will**	a uh painting for her from her head so
then it's okay yeah at the moment yes but I	**would**	learn it yeah irgend ein mal when you have a job
so now I am here and I think uhm that's a good job I	**will**	I learned a lot and I uhm I think I will like it
always from America to Switzerland and you	**will**	fly ten hours or twelve hours and that's a little bit much

These are classic learner errors, which often occur with Swiss beginners. When we examine the CEFR levels in Table 5.18 below, the sub-corpus 2 participant is a beginner. However, four of the five sub-corpus 1 participants have an intermediate level which does not follow intuitive expectations. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-11. It was confirmed that there was not a significant correlation between the two variables found, $r(5) = .72$, $p = .161$. If we then examine the languages at home, all participants and their parents' first language was Swiss German with one participant stating they additionally spoke 5% English and 5% French at home. The use of typical L1 interferences can be expected when the statistical information of NSU and Swiss German use are referred to in Section 4.2.1.

Table 5.18 CEFR and use of NSU-11

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-11	Mean NSU-11 per participant
1	1	6.67	A2	1	1.00
1	3	15.79	B1	3	1.00
1	2	14.29	B2	3	1.50
2	1	10.00	A2	1	1.00
Total	**7**			**8**	**1.14**

Another variable factor to consider in the increased number of NSU-11 in sub-corpus 1 is the order that the future was introduced and length and mode of instruction. Sub-corpus 1

participants were taught *going to* before *will* future, whereas sub-corpus 2 participants were taught the opposite order with more repetition.

Conclusion

To sum up, the non-standard use of *will/would* is considered to be due to L1 interference. All participants who used NSU-11 were from Swiss German backgrounds with the all but one NSU-11 from sub-corpus 1. It is possible that the order in which the future was taught might have an influence on its accuracy but other factors including the prolonged years of instruction cannot be excluded. With eight of the 183 instances non-standard the NSU rate for *will* is 4.37%.

5.1.13 NSU-12 Verb to be insertion or omission

Description

NSU-12 deals with the insertion or omission of the verb *to be*. The non-standard use of verb *to be* tenses was handled in the corresponding sections. The verb *to be* has many functions. It can be used as an auxiliary verb, linking verb, to form questions, to form negatives, continuous tenses and passives (Swan, 2016, p. 17).

Classroom introduction

The verb *to be* is taught in both coursebooks early on. In the sub-corpus 1 coursebooks, it is the very first item covered, whereas in the sub-corpus 2 coursebooks it is used throughout in texts from the beginning.

Findings

There were 62 instances of NSU-12 found. The verb *to be* was inserted 23 times and omitted 39 times. While investigating the 23 NSU-12 instances of insertion, there were three main categories of insertion found; false starts or mid-utterance redirection, attempted continuous construction and non-standard insertion of *be*. In the examples of false starts or mid-utterance redirection below we see that the participants appear to have redirected their statements. There were 13 instances found in total, three instances with *am* and ten with *is*.

(68)   spoke with us and how are you [L] it was very funny and I **am** met a girl she was English <30>

(69)   I was skiing down and I didn't know it so I' **am** just drove down and there <13>

(70)   the yellow band you can have alcohol and because it **is** means you are eighteen <17>

(71)   but if I am just uhm or on holiday or it **is** doesn't matter <14>

There were three instances of attempted use of present continuous, where the verb *to be* was used but not the *-ing* verb form. As seen in the two examples below either the present continuous or present simple could have been used to imply a slightly different meaning. The insertion of *are* complicates understanding. Possible intention of *ing* is added in brackets.

(72)   to talk to talk with him and yes both **are** profit (ing) from this and <36>

(73)   cuz they **are** live (ing) here in Switzerland <127>

The third category, the non-standard insertion of the verb *to be* is more difficult to decipher the speakers' intentions as seen in the examples below.

(74)   the woman doesn't like the picture and the artist must **be** draw a new picture

         <121>

(75)   [L] and my French it wasn't **be** better but my English [L] <24>

Most of the NSU-12 instances of insertion were used by sub-corpora 1 participants as Table 5.19 below shows. Here we see that in sub-corpus 1 the majority of participants had B1 and B2 proficiency levels, whereas the sub-corpus 2 participants were mainly at A2 level. This suggests that sub-corpus 1 participants might have more dysfluency in speech flow as they exhibited more false starts or mid-utterance redirection which were used by more participants. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of insertion of NSU-12. There was not a significant correlation between the two variables found, $r(5) = -.86$, $p = .060$.

Table 5.19 NSU-12 insertion and CEFR levels

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-12	Mean NSU-12 per participant
1	3	20.00	A2	3	1.00
1	7	36.84	B1	7	1.00
1	7	50.00	B2	7	1.00
2	3	30.00	A2	5	1.66
2	1	10.00	B2	1	1.00
Total	**21**			**23**	**1.09**

Several patterns were found with the omission of the verb *to be*. Ten instances occurred before a continuous verb, four times before *understood*, twice before *gunna*, twice before *grammatically*, twice before *just* and remaining instances were random. The intended use of the verb *to be* is illustrated in brackets in the examples below.

(76)   I don't use English I (**am**) learning English <8>

(77)   okay uhm to (**be**) understood <10>

(78)   was like when you want to come back we (**are**) gunna book a flight no worries <63>

(79)   yeah eh grammatically uhm when I uhm (**am**) grammatically correct and I try to explain something <19>

(80)   speak grammatically correct but if I (**am**) just uhm or on holiday <14>

(81)   talk with native people but and it could also (**be**) people from other languages <2>

(82)   I went there alone with the train and I (**was**) afraid a little because it was the first time <24>

Table 5.20 below shows that sub-corpus 2 had a higher rate of verb *to be* omission with greater percentages of participants at all proficiency levels except B1 where they were slightly lower. In addition, a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and omission of NSU-12. There was not a significant correlation between the two variables found, $r(5) = -.85$, $p = .067$.

Table 5.20 NSU-12 omission and CEFR levels

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-12	Mean NSU-12 per participant
1	5	33.33	A2	5	1.00
1	5	26.31	B1	6	1.20
1	5	35.71	B2	9	1.80
2	5	50.00	A2	7	1.40
2	5	23.80	B1	7	1.40
2	4	40.00	B2	4	1.00
2	1	33.33	C1	1	1.00
Total	**30**			**39**	**1.30**
Total NSU-12	**51**			**62**	**1.21**

Conclusion

In conclusion, the non-standard insertion of the verb *to be* was found to have 3 categories with the majority from sub-corpus 1 which could indicate a higher level of disfluency. The omission of the verb *to be* was found to be more frequent in sub-corpus 2 in all but CEFR level B1. It can be speculated that one reason could be the attempt to use continuous forms. Of the 39 participants who used NSU-12, only 6 used both insertion and omission of the verb *to be*. This suggests that insertion and omission should be investigated separately.

5.1.14 NSU-13 Omission and substitution of prepositions

Description and Classroom introduction

As mentioned in 5.1.6, many prepositions are similar in English and German; however, they remain difficult to learn and internalise. As described in Section 5.1.6, both sub-corpus coursebooks offer opportunities to recognise and acquire standard usages.

Previous research

As mentioned in Section 5.1.6, Both Cogo and Dewey (2012) and Önen (2014) suggest that not only the insertion of redundant prepositions but the omission and substitution of alternative prepositions are a feature of ELF. For example, Önen states that in her IST-Erasmus corpus that in 8 out of 11 concordance lines the *to* was omitted in conjunction with *listen* and *listening* (Önen, 2014, p. 102). The CSC shows the opposite results with 8 out of 10 uses of *listen* or *listening to* used in a standard way and only 2 instances of omitting the *to*. This exemplifies the vast variance possible when exploring small, specialized corpora. Since the CSC is comprised of German L1 learners in which the German *hören* (listen) is often used with the preposition *zu* (to), it is not unexpected that the learners would have a good command of its use. The IST-Erasmus corpus on the other hand is comprised of 79 participants with 24 different first

languages from around the globe (Önen, 2014, p. 59), where prepositional compliments might not be the norm.

Findings

There were 224 instances of NSU-13 found with 175 non-standard preposition choices and 49 omissions of prepositions. An overview of the NSU-13 instances can be seen in Table 5.21 below. The three prepositions *from*, *on* and *at* stand out as being used in a non-standard way at a substantially high percentage rate when compared to the frequency in the whole CSC.

Table 5.21 NSU-13 preposition frequency

Preposition	Corpus frequency	No of NSU-13 omission	No of NSU-13 substitution	Total number of NSU-13	Percentage of NSU-13
With	773	1	6	7	0.91%
To	1841	18	12	30	1.63%
In	1637	2	35	37	2.26%
by	70		2	2	2.86%
For	536	6	10	16	2.99%
Of	768	10	17	27	3.52%
About	253	2	8	10	3.95%
From	344	2	17	19	**5.52%**
On	341	3	34	37	**10.85%**
At	324	5	34	39	**12.04%**

Staring with *from*, 13 of the 17 substitutions were variations on the example below where *from* was substituted for *of* when talking about *a picture **of** her* or similar situations relating to a picture. An additional 3 instances were also substitutions of *from* for *of*, see Example 84.

> (83)   there was a painter who painted a portrait **from** (of) a woman and he painted it
>   as she as he saw her <106>

> (84)   and also Hobbit he was the voice actor **from** (of) the Dragon <138>

These can probably all be attributed to German L1 interference where *from* has been used instead of *von*.

| German | ein Bild *von* ihr |
| English | a picture *of* her |

Next, over 10% of all occurrences of *on* in the CSC were found to be non-standard and the 37 instances can be grouped into 4 categories. The first is *on* substituted for *in* as a preposition of place or time as in Example 83. There were 8 such instances. In addition, *in* was substituted as in which language is spoken with 11 occurrences (Examples 86 and 87). The intended prepositions are in brackets.

(85)   uhm it's **on** (in) summer or winter <16>

(86)   I don't think so no oh perhaps uhm when I watch the series perhaps Scrubs or Malcom in the Middle so **on** (in) English <12>

(87)   they also talk to us **on** (in) German that's a little bit sad but just for my father <56>

It is probable that Examples 4 and 5 are again L1 interference as the following sentences show where *on* has been used instead of the German *auf*, although the use of *in* is also possible in High German. As another option, the avoidance of prepositional use in the Swiss German example where *with* (mit) refers to the dative mir (me). Thus, the Swiss German speakers have several interlanguage examples to refer to when choosing a suitable preposition.

German	Er sprach *auf* Englisch mit mir.
	Er sprach mit mir *in* Englisch.
English	He spoke to me *in* English.
Swiss German	Er hat englisch geredet *mit* mir.

The third category is the substitution of *on* for *at* with 11 occurrences. Eight of them were prepositions of place; *at work, at a train station, at bars* etc., two involved *at Christmas* and only one occurred in combination with the verb *look*.

(88)   uhm for example here in the lesson you look **on** (at) the the grammar that is right because we have to do this on the tests <103>

The last category with 3 occurrences of *on* substituted for *to*. All three were directional as exemplified below.

(89)   we needed to go because of an event and he wanted to go **on** (to) the event without spoilers <138>

The preposition with the highest rate of non-standard use is *at*, which occurred as a substitution 34 times. Four categories of substitution were found with *at* substituted for *in* the highest with 18 occurrences. The following examples illustrate how *at* was used to express place and time. In these examples the *at* would be replaced with *in*, *im,* and *am* in German. Thus, the complexity of interchanging similar prepositions becomes apparent.

(90)   I was **at** (in) Albania I am from Kosovo but Albania is an other country <10>

(91)   okay uh I was in Mexico **at** (in) 2006 <30>

(92)   it was uhm thirty uhm degrees and then **at** (in) the evening it was uhm every

         evening was an event <29>

The next group of 8 substitutions were made with *at* instead of *on* while expressing place or

position in a broader sense. As we can see, the learners chose *at* as a substitute for *im*, *am*, *an*

and *auf*, demonstrating once again the complexity of preposition choice.

at (on) holiday <135>	im Urlaub
at (on) the seventh day <29>	am siebten Tag
at (on) the beach <12>	am Strand
at (on) a normal day <19>	an einem normalen Tag
at (on) the woman <117>	auf der Frau
at (on) the cruise ship <124>	auf dem Kreuzfahrtschiff
at (on) the chair <10>	auf dem Stuhl

This continues with the last two categories with *at* substitution for *to* (n=6) and *for* (n=3). In

this group *at* was used instead of the German *zu*, *zum* and *nach*.

(93)   now she's **at** (for) the fifth time by Pete's to make a picture <34>

(94)   we was uh going **at** (to) the <X> and Mayas <30>

(95)   yeah and we went first **at** (to) New York <139>

The 49 instances of the omission encompassed 9 prepositions. The following examples

illustrate the most common.

(96)   my girlfriend is uh uh from Turkey and because **(of)** that uh we go uh

         sometimes there it's nice and beautiful <129>

(97)   when me and my family we uhm all year we went we go **(to)** Italy for about I

         think two weeks or so <12>

194

(98)   but uh my father cannot speak very well English and my mother absolutely no

so I have to translate **(for)** them <120>

(99)   I see a man on the picture he is painting a woman and then the woman is

looking **(at)** the picture and she find it not so good <21>

(100)   I I don't go much **(on)** holidays outside of Switzerland <137>

(101)   and they have big cars and things like that for them I think that to have

something is very important and yeah **(in)** Switzerland the people are a maybe

a little they don't want a little they want a little bit less than in America <11>

Turning to the overall omission and substitution of prepositions, Table 5.22 below shows that 79/96 participants or 82.29% of all participants used NSU-13 at least once. In sub-corpora 1 it was 90.38% and 82.29% in sub-corpus 2. This implies that prepositions continue to be difficult to master at all CEFR levels, as is evident in the heightened use of NSU-13 by 2 C1 level participants from sub-corpus 1. Although fewer sub-corpus 2 participants used NSU-13, they averaged 3.15 NSU-13 per participant compared to 2.61 NSU-13 per participant from sub-corpus 1. Therefore, the outcome is similar. Furthermore, a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-13. There was a significant correlation between the two variables found, $r(5) = -.91$, $p = .029$.

Table 5.22 CEFR and use of NSU-13

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-13	Mean NSU-13 per participant
1	15	100.00	A2	38	2.53
1	19	100.00	B1	48	2.52
1	11	78.57	B2	26	2.36
1	2	66.66	C1	11	5.50
2	9	90.00	A2	33	3.66
2	14	66.66	B1	40	2.85
2	7	70.00	B2	26	3.71
2	2	66.66	C1	2	1.00
Total	**79**			**224**	**2.83**

Conclusion

The NSU-13 accounted for 8.47% of the total number of NSU in the CSC corpus with 224 instances. Ten prepositions were involved in a non-standard substitution and 9 prepositions were omitted 49 times. In this section, it was illustrated how many of the non-standard usages can be considered L1 interference because of the similarities and differences between Standard English, German and regional Swiss German. These assumptions have been made assuming that German was the L1 interferer. However, with many of the CSC participants having multilingual interlanguage, preposition choice could have stemmed from an array of other languages. If heightened accuracy in prepositions is a scholastic goal, the methods currently being used to teach and practice them will need to be evaluated and improved.

5.1.15 NSU-14 Full *yes* used as a discourse marker

While transcribing the interviews, it was observed that there were many instances of a speaking turn ending with a full *yes* and this was originally hypothesized to be used not only to back-channel and agree but as affirming the statement to be true. Baumgarten and House also found this phenomenon during oral exams with German university students and labelled it *turn/utterance-final* (Baumgarten and House 2010b, p. 1). Halfway through transcription and

coding, NSU-14 was abandoned because of the subjectivity involved in deciding the intention of the speaker. Nonetheless, during analysis, the use of *yes* and *yeah* was recognised as invariant tags and was fully investigated in 5.1.5.

5.1.16 NSU-15 Plural –s omitted or inserted

Description

Plurals belong to the grammatical category of noun agreement. English differentiates between countable and uncountable nouns. Countable nouns, as their name suggests, can be counted, for example, *papers* or *boxes* and are generally formed by adding –s or –es. Additionally, there are irregular and special plurals which must be learned. Uncountable nouns are not perceived as being separable items that can be counted, but singular (Swan, 2016, p. 116). This can cause confusion for learners with certain nouns that are considered countable in their L1. For example, in German *information* is countable, whereas in English it is considered singular.

Classroom introduction

The sub-corpus 1 coursebooks cover plural nouns in the first unit in grade seven. However, countable and uncountable nouns are not covered until the last quarter of grade nine. The sub-corpus 2 coursebooks continue with the scaffolding approach by covering plural nouns in the first unit of grade three then irregular plurals are introduced in the following unit and in the beginning of grade five regular and irregular plurals. In grade seven countable and uncountable nouns are covered and repeated at the beginning of grade eight.

Previous research

In her quasi-longitudinal study of the written learner corpus ICLE, Thewissen (2013, p. 90) found that plurals, being one of the earliest taught grammatical rules, were quite accurate at the

CEFR B levels. However, subtle errors in omitted or 'disruptive plurals' continue well into the higher proficiency levels. While examining non-native speech, it has been recorded that many new Englishes deviate from standards in pluralisation (Dröschel, 2011, p. 182; Mollin, 2006, p. 124). Both Dröschel (2011), investigating the Swiss English Database, and Mollin (2006), investigating the Corpus of Euro-English, concluded that although they found evidence of pluralisation deviance, it was not widespread throughout the data and could be traced back to multiple uses by single participants.

Findings

There were 160 non-standard instances of omitted or inserted plural *–s.* Out of the 60 words used, 48 were count nouns which were used 1 to 8 times. In 16 of those instances, a plural *-s* was added and in 32 instances the plural *-s* was omitted. Examples of the three most used words follow:

> (102) and uhm yes I did it for one and a half **year** and after that I break up <1>
>
> (103) and then then he draws another **pictures** and then she was and then she was <29>
>
> (104) the first book I tried to read I started it three or four **time** but always four five Seiten (pages) <2>

Furthermore, non-count nouns accounted for 13 NSU-15 instances and included bloods (n=1), fishes (n=1), perfumes (n=2), informations (n=1) and hairs (n=8). It is interesting that the word *information* occurred only once in the entire corpus and that it was used with a plural -s. Both Rosenberger (2009) and Dröschel (2011) explored the possibility of the pluralisation of

information being a Swiss English characterisation but concluded that despite its high frequency in their corpus, it was used by a limited number of participants.

Subsequently, five irregular plural nouns account for 41.55% of all NSU-15 in the CSC with variations of the 5 irregular nouns *feet* (n=2), *children* (n=3), *man* (n=4), *woman* (n=2) and *person* (n=55). *Person* is an irregular noun which takes the plural form *people*. Since *persons* can correctly be used in official or legal language, accuracy can be especially challenging for L1 German learners because the German *Person* becomes *Personen* in its plural form. Nevertheless, there were only 16 NSU-15 instances with *person/persons* and 39 with *people/peoples* with all but 5 NSU-15 instances with a redundant *-s* inserted. An interesting correlation between the NSU-15 use of *people* and *person* and proficiency levels was observed. Of the 39 uses of NSU-15 with *people*, the majority, 48.71%, were from participants with a B1 level, whereas the majority 50% of the 16 NSU-15 used with the noun *person* were used by participants with a B2 level. This could indicate that accuracy in the use of *people* precedes *person*.

Besides nouns, the determiners *this* and *that* were used a total of 9 times as a NSU-15.

    (105)   Important for them and if they have **this** things then they are happy <11>

    (106)   it's more about architecture and ceiling construction and **that** things and I don't want to lose <116>

The raw numbers are interesting to observe; however, significance in relation to standard use can indicate areas for further investigation. In Table 5.23 below we see that *hair* and *person*, for example, occurred quite frequent in the corpus due to the picture description element which generated a wealth of comparable vocabulary. Here we see that both *people/person* and *hair*

have a considerably higher rate of non-standard usage. This could be an indication that other irregular nouns might also have been used less accurately if they were highlighted. A follow up study with a picture description using other irregular nouns could substantiate this.

Table 5.23 Percentage of NSU-15 compared to CSC

Word	Frequency	No of NSU-15	% NSU-15
This/that			
These/those	1161	9	0.76%
Woman/women	146	2	1.37%
year	263	7	2.66%
Man/men	111	4	3.60%
Child/ children	40	3	7.50%
People/person	518	55	10.61%
hair	44	8	18.18%
information	1	1	100%

Turning to the sub-corpora, the distribution of NSU-15 and proficiency levels is shown below in Table 5.24. In sub-corpus 1, a total of 101 NSU-15 were used by 38 participants with the B2 level participants using the most NSU-15 per person at 3.90. Sub-corpus 2 on the other hand, only used 59 NSU-15 by 22 participants. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-15. There was not a significant correlation between the two variables found, $r(5) = -.78$, $p = .118$. Thus, the most probable reason for the difference is the presumed didactical introduction of countable and uncountable nouns at the end of classroom instruction in sub-corpus 1, as opposed to the scaffolding approach and repetition of plural and countable noun instruction in sub-corpus 2.

Table 5.24 CEFR and NSU-15 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-15	Mean NSU-15 per participant
1	10	66.66	A2	23	2.30
1	14	73.68	B1	31	2.21
1	11	78.57	B2	43	3.90
1	3	100.00	C1	4	1.33
2	6	60.00	A2	18	3.00
2	12	57.14	B1	32	2.66
2	4	40.00	B2	9	2.25
Total	**60**			**160**	**2.66**

Conclusion

This section has highlighted the most frequently used words that were used in a non-standard way. This knowledge can be an indicator that special attention should be given to these words when teaching and deciding when to comment on grammatical correctness when correcting speech. The use of *people*/*person* and *hair* were most prominent and were used by participants from all proficiency levels. The presumed introduction and intensity of classroom instruction was also stated as a possibility for differences between the sub-corpora.

5.1.17 NSU-16 Lexical choice

Description

In this category of non-standard usages, word choice encompasses grammatical and lexical words. The subcategories which are investigated below include in this order: German interference, antonyms, phonetic similarities, base form variations, countries, pronouns, adverbs, numbers and quantities, high semantic verbs, and a miscellaneous category.

Classroom introduction

As stated in 5.1.7, the participants were exposed to at least 2,000-3,500 words in their coursebooks alone. However, 'Comprehending a word is not same thing as producing a word'. (Carter, 1998b, p. 191)

Previous research

Common sense might have one believe that the more we learn the better we become in a linear progression. However, language acquisition is more complicated, in Larsen-Freeman's words, 'language performance and development are complex, nonlinear, dynamic, socially situated processes' (Larsen-Freeman, 2009, p. 10). Contributing to the understanding of these processes, Thewissen (2013) suggests different error categories fall into a range of developmental patterns with strong and weak development as well as non-progressive patterns. Her research results indicate that there could be a grammatical accuracy threshold at B1 level because she found that over 70% of the tagged errors did not markedly change between levels B2-C2 in her quasi-longitude corpus study (Thewissen, 2013, p. 95). Lexis errors seem to change at higher proficiency levels and not necessarily diminish. The findings below show examples of this.

Findings

There were 807 instances of NSU-16 in the whole corpus with 424 used by participants from sub-corpus 1 and 383 used by sub-corpus 2 participants. This accounts for 30.54% of all total NSU in the corpus. These NSU-16 were used by 93 of the 96 participants at least one time. This highlights the importance of lexical choice because even if many ELF features or non-standard usages can be considered non-impeding, it is the opinion of the author that many of the following lexical choices NSU-16 can hinder understanding and thus communication.

Of the 3 participants who did not use any NSU-16, two were from sub-corpus 1 and one from sub-corpus 2. They were investigated for commonalities. Two of the participants were found to be considered native speakers with at least one native English-speaking parent and only one NSU between them. The remaining participant was a quad-lingual Croatian who also used only one NSU in their brief interview. It was evident that the non-native participant had advanced language skills and therefore a good command of lexical agility. It appears that all three were the exception with a higher degree of standard lexical repertoire (dexterity) than the corpus norm.

5.1.17.1 Distinct German interference

The first category of words that were explored were of distinct German interference where 144 instances of 25 different NSU-16 words were found. Eight were found in both sub-corpora and quite common:

Table 5.25 Common German interference words

Words used	Presumed intended words	German equivalent	Frequency
1. become	get, receive	bekommen	5
2. fault	mistake	Fehler (French = faux)	4
3. it gives, it has, you have, it was, they was	there are, there were	es gibt, es hat	12
4. made/make holiday	went/go on holiday	Ferien machen	3
5. meet us	meet each other	uns treffen	4
6. lucky	happy	glücklich	5
7. other	else, different, another	anders	21
8. grammatic(s), grammatical(ly)	grammar, grammatical(ly)	Grammatik	60

The first two examples are clearly due to the idea of *false friends* where words in two languages are very similar but have a different meaning. In German *bekommen* means *to get, receive or obtain* and *Fehler* means *mistake, error or flaw*. *Fehler* is also very close to the French word *faux* which means *wrong* or *false*. The next three examples are direct translations. *There are* in

German is *es gibt* (it gives) or *es hat* (it has). In German holidays are made (*machen*) and we meet us (*uns*). The last two examples represent a word in one language having multiple meanings when in the other language it does not. The German *glücklich* means both *happy* and *lucky* and *anders* translates to English as *other*, but can also mean *else, different* and *another*. Lastly, grammar translates to *Grammatik* and was used by 40.62% of the participants in a non-standard way. CEFR levels were investigated and NSU-16 with *grammar* was used by all CEFR levels with 32% of the A2 level participants, 42.5% B1, 45.6% B2 and 16.66% C1. A probable cause for the lower level of A2 level participants could be avoidance strategies.

The Swiss curriculum encourages functional multilingualism by striving to use all the students' language competence to achieve advanced communication skills and efficiently expand their multilingual repertoire[28]. This multilingual awareness might negatively affect the use of false friends. It is not uncommon in Switzerland for German L1 English learners both young and old to automatically substitute a French word when they attempt to converse in English while not noticing that they were using the wrong foreign language.

Moving on to the NSU-16 words used only by sub-corpus 1 participants, most of them have a phonological based German interference and some of the participants were quite creative. Table 5.26 below lists the words.

---

[28] Deutschschweizer Erziehungsdirektorenkonferenz (D-EDK) (Eds.). Lehrplan 21. Fachbereichslehrplan Sprachen, S. 4.

Table 5.26 Sub-corpus 1 Phonological German interference

Words used	Presumed intended words	German equivalent	Frequency
1. break up	Leave (left)	aufbrechen	1
2. card	map	Karte	1
3. control	check	kontrollieren	1
4. learn	meet	kennenlernen	5
5. man	husband	Mann	1
6. woman	wife	Frau	2
7. wife	woman	Frau	3
8. haven	harbour	Hafen	3
9. quality	jellyfish	Qualle	2
10. scribe	write	schreiben	1

In the first example the verb *aufbrechen* sounds similar to *break up*, but means to leave, depart or forcefully break open. It was used in the example below.

(107)   And after that I **break up** this apprenticeship and then I went (uhm) for a month

working in a factory <1>

An alternative interpretation of this NSU-16 could have been the use of the phrasal verb *broke off*. To validate the decision to classify this as German interference, the corresponding transcript was reviewed, and 20 instances of German code-switching were found including one use of the word *aufbrechen* several sentences later to describe the same occurrence was found.

(108)   I think uhm <G> **s'abbräche** vo de Lehr der=4 </G> uhm was good for me <1>

The second and third examples from Table 5.26 are examples of false friends where the words are pronounced very similarly but have different meanings with *Karte* meaning *map* and *kontrollieren* meaning *to check*. In Example 4, the second part of the German word *kennenlernen* which means *meet* or *become acquainted with* was used 5 times by four different participants. Examples 5, 6 and 7 often cause confusion in beginners because in German *Mann* means both *man* and *husband* and the German *Frau* means both *woman* and *wife*. Example 8

was an interesting word choice because the participant modified the German word *Hafen* which means *harbour* to *haven*. In German, the consonant *v* can produce both an English *v* and an *f* sound, so it appears the participant just switched the sound in an attempt to produce an English word. Example 9 shows the most creative attempt to communicate by using the word quality to mean *jellyfish*. In German, a *jellyfish* is a *Qualle* and is pronounced exactly like the first syllable of quality. Lastly, in Example 10 the word *scribe* was used to mean *write*. Here the similarities to the German word *schreiben* are more orthographically than phonetically based.

The last group of German interference words were from sub-corpus 2. The first four examples have a phonological basis and in the last three examples, the confusion was in meaning. In Table 4.46 below, Examples 1 to 3 are false friends with the spelling and pronunciation almost identical in both English and German. Example 4 sounds similar with different spelling. The German meanings are represented in the presumed intended words column below.

Table 5.27 Sub-corpus 2 German interference

Words used	Presumed intended words	German equivalent	Frequency
1. land	country	Land	1
2. sensible	sensitive	sensibel	1
3. familiar	family	familiär	1
4. yet	now	jetzt	3
5. so	such a	so ein	1
6. deep	thick, heavy	schwer, stark,	1
7. strong	different, many	schwer, stark	2

In Example 5, the more general word *so* was used to mean *such a*, but it was not possible to decipher if *so* was used intentionally as an English discourse marker or as a German word. Examples 6 and 7 meanings stem from the German words *schwer* and *stark* which have multiple meanings which do not correspond exactly with English so learners might have difficulty choosing the best word to express intended meaning.

All of the NSU-16 in the above section except *quality* and *haven* are typical examples of German L1 interference which are frequently encountered in and out of Swiss classrooms. This type of NSU -16 would certainly cause some conversational misunderstanding if one of the parties involved had limited understanding of German.

5.1.17.2   Antonymy

Within the NSU-16, 6 instances of participants using the opposite of the intended word were found. The examples below illustrate this with the intended meaning in brackets.

(109)   and he has **learned** (taught) me to drive a uhm uh uh tractor and mo motorcycle and car <131>

(110)   he **came** (went) down the hill with a Forst vehicle <133>

(111)   they are uhm **close**[29] (open) um yeah **close** (open) in Barcelona they are uhm uhm hey hello how are you they don't know me but they are very friendly to me <55>

(112)   it was very expensive also in Italy we uhm **buy buy** (paid) for a person for a night uhm thirty euro <29>

(113)   yes of course we had we **stay up** (get up) in the morning <58>

The examples are from different participants with the first two from sub-corpus 2 participants with a B1 proficiency level and the following three from sub-corpus 1 participants with an A2 proficiency level. These antonymous examples are reminiscent of word association tests where the 'structural semantics approach' declares that words 'do not exist in isolation' (Carter,

---

[29] The pronunciation of close was /kləuz/.

1998b, p. 19). These beginner and low intermediate learners seem to have pulled on their language resources and produced lexis with a relation to the intended words.

5.1.17.3   Phonetic similarities

This section deals with seventeen NSU-16 where phonetic similarities appear to have determined the NSU-16 choice. The words used have phonetic similarities with words that would make contextual sense.

First language vocabulary is not stored in the brain separately but is linked together 'based on membership to semantic, phonological, graphical, syntactic and other classes' (Aitchison, 1994 in Schmitt and McCarthy, 1997, p. 174). Although Aitchison maintains that 'there is no general agreement as to how the various lexicons are organized in the minds of bilingual and multilingual speakers' (Aitchison, 2012, p. 268), it would appear to be a logical conclusion that there would be many similarities between languages. Factors such as the degree of bi- or multilingualism and languages spoken would also need to be taken into consideration when interpreting data from multilinguals.

It has been confirmed (Browman, 1978; Korait and Lieblich, 1974; Rubin, 1975) that native English speakers recall the beginnings and endings of words better than the middle part. This has been labelled 'the bathtub effect' by Aitchison (2012) and is like a person in a bathtub with their head and feet out of the water. The exposed parts are remembered and the middle part under the water forgotten, with the head protruding higher from the water and thus remembered at a greater frequency (Aitchison, 2012, p. 158). The corpus participants in this study have exercised similar behaviour. In 10 of the 17 examples, the beginnings and 6 examples the endings were phonetically similar.

Table 5.28 below showcases the examples with Examples 114-123 with the beginnings. Here it is interesting to note that number 118 represents 8 instances found of *print* being substituted for *paint* in the corpus from 5 different participants, three from sub-corpus 1 and 2 from sub-corpus 2. The abundance of instances is certainly due to interview part two where all participants were asked to describe the pictures. Example 122 is also interesting because there is the possible interference of L3 French. *Portrait* in French is pronounced with a silent *t* which sounds exactly like *portray*. As mentioned in Section 1.3.3 French is an obligatory subject in all Swiss schools so it is plausible that most participants would have some French vocabulary in their mental lexicon.

Table 5.28 Examples of phonetically based NSU-16

No.	Example	Sub-corpus
114	she sees the picture and and thinks yeah I would it would it look better with uh **kirby** (curly) hair	1
115	**many** (maybe) I would say five thousand or more dollars a month	1
116	we saw the empire state building **state** (statue) of liberty yeah stores uhm then central park	1
117	easier I think because when I speak English I am always a little bit I **things** (think) nervous	1
118	a story okay a woman came to a **printer** (painter) and wanted to and and ask him to make a picture of her	1 and 2
119	the woman could see the picture but she didn't like it then sh he **drove** (drew) another one and then finally she liked it	2
120	it's **healthy** (helpful) when you are in another country you it's yes **healthy** (helpful) when you understood what the other people say or something and or when you can help in Lucerne somebody	2
121	we wanted to shoot some animals there uhm and I also I love uhm this the the the mountains you see and the uhm really big uhm **plate** (place)	2
122	I think this is uh an artist who makes **portrays** (portraits) of people I would say	2
123	it very bad and then the woman is very angry **why** (when) he draws her bad and then he has to paint it again	2
124	and it was the climate it was uhm very cold and also it was irgendwie (somehow) uhm twenty or **flower** (lower) also weniger (fewer) degrees and uhm yes it was	1

125	so this this eh semester **premester** (semester) I **felt** (fell) back in the fifth eh class yes that was eh a very hard experience for me	1
126	you have there uh many **icelands** (islands) and uhm you can spend with a uh uh boat the time	1
127	one from New Orleans one from New York one from flo Florida the other is from Las Vegas the other is from uh yes the other is from England then one is from Irland one is even from **island** (Iceland)	2
128	there it's very nice station uh uh and uh so we did a lot of things uh uh about uh drive with the **squad** (quad) yes quad	2
129	then they say yeah you can make yourself you know **where's** (there's) the things and then	2
130	yes this taught me a lesson that I'm you should uhm love the things you have and not be **kinky** (unkind) or yes do bad to other people	2

Examples 124 to 130 are examples of endings with phonetic similarity. Number 125 has two examples with the participant using the correct word and then changing it to a non-word. In Example 128 we see the opposite; first, the wrong word is spoken and then self-corrected. The interpretation of the last example was chosen due to the context of the conversation prior to the example and could have been just an oddly chosen word or more probably have had an association with being kind. In conclusion, these NSU-16 were used evenly by participants of both sub-corpora with the majority 60% at a low-intermediate proficiency range.

5.1.17.4 Base form variations

A total of 19 words were used by 23 participants 26 times as non-standard variations of base form words with the majority being verb-noun-adjective confusion. In addition, five of the words used were not standard words and are emphasized with italics in Table 5.29 below. The use of the word *conversate* by a Swiss German youth is quite interesting because it is considered to be a non-standard word that has been in use for over 200 years and currently associated with

African American slang and hip hop music[30]. This is the second indication that Swiss youth are gleaning non-standard vocabulary and grammatical phases from popular music. The first indication was with the use of *she don't like* which was made popular by Justin Bieber.

Eighteen of the words were only used once or twice by an individual participant and the remaining words *impressive* and *paint* which were used three and five times, were related to the construct of the corpus interviews. As Table 5.29 below shows, the majority, 17 out of 27, of this type of NSU-16 usage was used by sub-corpus 1 participants.

Table 5.29 Base form NSU-16 words

Word used	Intended word	Frequency	Sub-corpus
1. confidential (adj)	in confidence (n)	1	1
2. *conversate* (v)*	converse (v)	1	1
3. different (adj)	difference (n)	1	1
4. discuss (v)	discussion (n)	1	1
5. fame (n)	famous (adj)	1	1
6. *impresent*, impressed (v)	impressive (adj)	3	1
7. meditation	mediate (v)	1	1
8. record (v)	recording (n)	1	1
9. slippery (adj)	slid (v)	2	1
10. true (adj)	truth (n)	1	1
11. *uncorrect*	incorrect (adj)	1	1
12. understood (adj)	understanding (v)	1	1
13. capital (adj)	capitalize (v)	1	2
14. eastern (adj)	east (n)	1	2
15. interesting (adj)	interested (adj)	1	2
16. *kindfull*	kind (adj)	1	2
17. *understoodness*	understanding (n)	1	2
18. upside (prep.)	up (adj)	1	2
19. paint (n)	painting (n)	5	sub-corpus 1 (2x) sub-corpus 2 (3x)

Taking a closer look at confusion with *impressed* and *impressive*, the participants were given examples of how to use them in written and spoken form during the interview:

---

[30] https://www.merriam-webster.com/words-at-play/is-conversate-a-real-word

Topic 2:   A country you have visited which has impressed you. Describe your visit and say why you found the country particularly **impressive**.

The majority (n=64) of the participants chose topic two and a form or *impress* was used 65 times in the corpus. This equates to a NSU rate of 4.61% and is most likely due to the lower proficiency levels of the three participants involved, see examples below.

> (131) I was in Croatia in the national park uhm the name of this park was or is uhm Kirk and that was very **impresent** for me <44>
>
> (132) rollercoaster yes also the most of them and yeah yes it it was very uhm yeah **impressed** <5>
>
> (133) a cruise yes from Italy to Greece and Croatia and back so that was really interesting for me uh because uh the big ship and so and yeah I've seen four uh five uh villages or uh how do you say Orte uhm in one week so and yeah it was that was really **impressed** <119>

The next noteworthy word *paint* was one that all participants were compelled to use to describe part 2 of the interview. By doing this it is possible to get a better understanding of how these participants who represent the emerging workforce deal with regular verbs which also have a noun form. The following examples show the difficulties they had.

> (134) yes okay it's uh there's a woman and she went to a painter store draw and he make a **paint** from her and paint uhm the reality <42>
>
> (135) the woman uhm wants him to paint her and then she looked at the picture and I think uhm the picture of or the **paint** of her is more beautiful <59>

(136) she fights with him I think then the artist makes one more **paint** about her <136>

(137) yes so I think the woman will have uh uh **paint** of of uh herself <135>

(138) uh also he made her lucky on the **paint** and I think a little bit uh uh hübscher <135>

Turning to overall non-standard verb use, in Section 5.1.1 it was stated that 18.06% of the 72 instances of paint being used as a verb with infinitive or third person singular were non-standard. Now, when all of the 195 verb and noun instances are calculated paint(s)(ed) (n=89) + painting (n=53) + painter (n=53) with all 25 NSU associated with the word *paint*, a NSU ratio of 12.82% is found. The question of why some verbs are learned more accurately than others arises. To investigate, a list of the thirty most frequently words was created. First, a word list search with lemmas was conducted to determine the thirty most frequent verbs. Then, an AntConc search with * was conducted to find any other associated words and pre tagged NSU were manually counted. The results are illustrated in Table 5.30 below. It is surprising that 19 of the 30 most frequent verbs had a non-standard rate of over 5% when all uses of the word were considered and are marked below in bold.

Table 5.30 Thirty most frequent verbs in CSC

Verb	Frequency	Number of NSU associated with	NSU %
Guess	67	0	0.00%
Love	59	0	0.00%
Use	197	1	0.51%
Like	1210	10	0.83%
Think	1046	12	1.15%
Know	538	8	1.49%
Learn	299	7	2.34%
Mean	84	2	2.38%
Work	137	4	2.92%

Speak	467	19	4.07%
Understand	286	14	4.90%
Ask	59	3	**5.08%**
Get	209	11	**5.26%**
Take	99	6	**6.06%**
Do, don't	1121	69	**6.16%**
Come	134	9	**6.72%**
Drive	77	6	**7.79%**
See	252	20	**7.94%**
Try	69	6	**8.70%**
Look	222	21	**9.46%**
Talk	169	16	**9.47%**
Put	22	2	**9.52%**
Want	302	29	**9.60%**
Say	355	36	**10.14%**
Show	115	14	**12.17%**
Paint	195	25	**12.82%**
Tell	70	9	**12.86%**
Go	650	90	**13.85%**
Eat	73	15	**20.55%**
Make	231	60	**25.97%**
Draw	84	23	**27.38%**

Table 5.30 includes all 6 of the highly semantic verbs stated in Section 5.1.7 which are marked in dark blue. They have an NSU rate between 5.26% for *get* and 25.97% for *make*. Contrary to Cogo and Dewey (2012, pp. 70-73) finding innovative uses of highly semantic verbs collocates, the majority of non-standard use of highly semantic verbs in the CSC involved the inversion of another highly semantic verb as the examples below illustrate.

(139) five days with a friend we didn't **do** (go) sightseeing but mostly shopping <55>

(140) still need to work on like my grammar I **do** (make) a lot of little mistakes <63>

(141) rooms and and play with the children **put** (take) put them in the school <38>

The inaccuracy levels of the 30 most frequent verbs raises the question if this could be an indication that Swiss students have difficulty with lexical accuracy in general or only with certain verbs. Further investigation such as the compilation of a larger, more comprehensive corpus is necessary to fully understand the implications.

5.1.17.5 Countries

There was a total of 31 NSU-16 found involving the inversion of countries and nationalities with 21 instances found in sub-corpus 1 and 9 instances found in sub-corpus 2. This phenomenon is not extraordinary. However, it is interesting that this inversion was found evenly distributed in all proficiency levels indicating that accurate use of countries and nationalities by the participants in this study is not proficiency level bound. As can be seen in Table 5.31 below, increased NSU-16 use of Switzerland and its surrounding countries was recorded; however, they also had increased frequency.

Table 5.31 Country and nationality inversion

Word used	Intended word	Frequency	Sub-corpus	CEFR level
British	English	1	1	C1
Croatian	Croatia	1	2	B1
Egypt	Egyptian	1	1	B1
England	English	1	1	A2
French	France	6	3x sub-1, 3x sub-2	2x A2 4x B1
Germany	German	3	1	2x A2 1x B1
Greek	Greece	1	1	B2
Italian	Italy	1	1	B2
Italianer	Italian	1	1	A2
Japan	Japanese	1	1	B1
Korea	Korean	1	1	B1
Swiss	Switzerland	8	6x sub-1, 2x sub-2	2x A2 4x B1

Switzerland	Swiss	4	1x sub-1, 3x sub-2	1x B1 3x B2
Turkeys	Turks	1	2	A2

5.1.17.6 Pronouns

The pronouns, *I*, *he*, *she*, *we*, *they*, *me*, *him*, *her*, *us*, *them*, *you*, *your*, *my*, *myself*, *herself* and *yourself* were used in 40 different combinations a total of 62 times in a non-standard way. NSU-16 instances occurring in part 2 of the interviews where the male artist and female client were described accounted for 22.58% (n=14) of the NSU-16 with pronouns. In total, the direct inversion of gender *he-she* occurred 15 times and the inversion of object *it* and gender occurred 19 times which represents 54.83% of the NSU-16 with pronouns. This is in line with results found by Felix and Hahn (1985, p. 232) in their longitudinal study of German high school students' pronoun errors where gender substitution was most frequent.

5.1.17.7 Adverbs and adjective inversion with *-ly*

The term adverb is used for a wide variety of words with different kinds of use. Their general function is to modify (add meaning to) sentences, clauses or various parts of clauses, but not nouns (Swan, 2016, p. 193). Adjectives on the other hand can express a variety of meaning to nouns and verbs (ibid. p. 183). There can also be some confusion between adverbs and adjectives because some adverbs end in *-ly* and some adverbs and adjectives have the same form (ibid. p. 194)

There was a total of 29 instances of adverb and adjective with *-ly* inversion found which were used by 24 participants and involved 17 words. Two categories of inversion occurred: an adjective or noun instead of an adverb and an adverb instead of an adjective.

Most of the 5 participants who used the 5 instances with an adjective or noun instead of an adverb were B2 proficiency level participants from both sub-corpora such as in the example below.

(142)   I would say she uh he does this **professional** (professionally) he is in the building so he won't be on the street like uh in Barcelona where these artists are on the street <119>

The remaining 19 participants who used an adverb instead of an adjective were also distributed evenly between both sub-corpora; however, the proficiency range was far greater with the following number of NSU-16 for each level A2=10, B1=6, B2=4 and C1=4. Interestingly, one word 'normally' was used in a non-standard way 8 times by 8 different participants from all proficiency levels.

(143)   yes it's a different it's in Sweden is very cold and in Switzerland is **normally** <20>

(144)   he did a really a really uh imaginary picture uh he made her hair look other than **normally** so when she got back to the picture and saw it she was very impressed <33>

(145)   and the people are also when they have uh a good job also a **normally** job they can also uhm enjoy that's mine meine Meinung that the people can better enjoy the life <44>

(146)   and something like that and camel meat also and when you were in Sydney we ate **normally** like yeah meat and something like because there were were normal restaurants like here <56>

(147) oh my god after this one week I just want to go home and eat something

normally <58>

(148) women too for two years guys for three years uhm and it's just like you are

normally like kinda learn a job there which you can use afterwards <63>

(149) because I'm just like more with these people and not with a grammatical uh

speaker they're like just friends speaking **normally** English <132>

(150) not not uh **normally** holidays but it's very interesting you make different things

you go uh walking walking uh up to mountains <137>

Example 147 is an instance where the decision to mark a word as non-standard was not straightforward. Therefore a highly experienced English lecturer was consulted, and the decision was made to mark it as non-standard after reviewing the whole content of the interview section.

In conclusion, a corpus search revealed 40 instances of *normal/normally* which equates to a non-standard rate of 20%. This number appears important and merits further investigation. In earlier research, Rosenberger (2009) dismissed the possibility of adverb/adjective inversion as being a feature of Swiss English due to the relatively low frequency of 1-2% non-standard usages found with the 61 pairs of adjectives and adverbs he analysed (ibid. p. 195). His data revealed the opposite ratio of inversion with 75% adjective instead of adverb compared to 17% in the CSC and 25% adverb instead of adjective compared to 82% in the CSC.

5.1.17.8 Numbers and quantities

There were 10 instances of NSU-16 used with numbers. Half of them were a cardinal number instead of an ordinal number and the other half were multiplicative number and ordinal number inversions.

    (151)   it was a special restaurant they have from the uh **twice** world war uh battle <128>

With 60% of these NSU-16 used by A2 proficiency level participants, it appears to be lower proficiency level bound.

The category of quantities includes all NSU-16 related to *all*, *some* and *any* as well as *less*, *little*, *few*, *a lot of*, *many* and *much* which amounted to 82 instances. The results were divided up into five smaller groups for comparison. Firstly, 18 instances occurred where *much*, *very* and *many* were used instead of *a lot* or *a lot of*. Interestingly, all proficiency levels were involved; however, the majority 70.58% were used by sub-corpus 2 participants.

    (152)   for tourists everything was quite dirty and there was **many** (a lot of) rubbish and uh also there were <118>

    (153)   went to holidays in my childhood I I went there quite **much** (a lot) with my parents and <119>

Next, there were 8 instances of *all* and *every* inversion with 50% of the participants at an A2 proficiency level and 7 out of 8 from sub-corpus 1.

    (154)   yeah divorced and I see my father **all** (every) two weeks and yes he <9>

(155)   the bus was very old and not yes so the **all** (everything) was old it was so yes

        <21>

Any and *some* inversion was the focus of the third group with 15 instances. Here 11 out of 14 instances were from sub-corpus 1 with only 2 at A2 proficiency level.

(156)   because it's really difficult for the Chinese people to come to Switzerland or go

        everywhere (anywhere) it's really hard for them <56>

(157)   you get to know the sister as a friend because you don't know **someone** (anyone)

        else <12>

The fourth and largest group dealt with all variations of *much* and *many* NSU-16 with 34 instances in all proficiency levels and divided equally between the sub-corpora.

(158)   she was too thin then she drank too much had too **much** (many) muscles then

        she had various implants and used Botox <145>

(159)   I don't know I don't know there are so **much** (many) different places so (uh)

        maybe in ten twenty years <33>

The last group concerned *little*, *small* and *less* with 7 instances all from sub-corpus 1. Here all instances were used by participants between B1-C1 proficiency levels.

(160)   probably non-native because uh considering that they're so **little** (few) well I

        don't know how many native people there are here but (uhm) well <35>

(161)   I think so or America but I don't talk so much so cuz yeah I have a **little** (small)

vocabulary <32>

The increased use of NSU-16 with the expression of quantities in sub-corpus 2 could indicate that there is a need to review when and how these items are taught.

5.1.17.9   High semantic word confusion

The overuse of high semantic verbs belongs to the original list of ELF features and were assumed to be present in the CSC. As mentioned in Section 3.3.7, instead of overuse, substitution was found. There were 87 instances that fell into this category with the words *make* (n=41), *do* (n=5), *go* (n=5), *put* and *take* (n=2), *tell* (n=8), *talk* (n=6), *say* (n=9), *look* (n=6), and *watch* (n=6). Starting with *make* which was substituted 24 times for *do/did* as in the following examples with intended words in brackets.

(162)   they correct me and not want to say you **make** (do) it false I think it's <56>

(163)   you have to be you have to **make** (do) one job uh one Ausbildung (internship)

<50>

The remaining 17 uses of *make* were substituted for *go*, *have*, *learn*, *visit*, *put*, *see* and *take*. There was much less confusion with the use of *do*. It was only used twice instead of *make*. This is probably because of the German *machen* which means to make.

The next word to consider is *go* which was used 5 times as a substitute. Twice for *took* and once each for *get*, *do* and *were*. *Put* and *take* were also substituted twice. Although they are considered highly sematic (Clark, 1978), *put* occurs just 22 times and *take* 98 times in the CSC.

(164) Salzburger Nocken is with uhm cran cranberries I think cranberries I think and uh it's with uh eggs and they uh when you **took** (put) that in the uh Backofen (oven) <135>

(165) clean the rooms and and play with the children **put** (take) them in the school <38>

Next, *Say* and *tell* are documented by Swan and Smith (2001, p. 49) as being misleading cognates for German L1 speakers. Three instances were found of their substitution, as well as 5 instances of substitution for *show* which were used during the picture description. However, all instances with *show* were used by a single participant at an A2 CEFR level.

(166) he he **tells** (shows) this girl the photo and she was shocked she was not uhm not begeistert also (thrilled) I don't know <127>

(167) and in English I have to think about the words how I can **tell** (say) them maybe think about the words [L] it's like this so yeah <139>

Reversing the cognates, *say* was used in place of *tell* 7 times and twice with *ask*. However, only 3 out of nine were used during the picture description from 8 participants including all CEFR levels with all but one above an A2 level.

(168) he wanted his medication and we ss **said** (told) him we can't help him <106>

(169) she asked me that oh can you knock on the window and **say** (ask) if it's okay if we take a picture <101>

Continuing with verbalisation, *talk* was substituted for *speak* 5 times and once for *say*. Five of the 6 participants were a CEFR level B1 and the last B2.

(170)   but he can't **talk** (speak) German so he **talked** (spoke) only English <1>

Lastly, we will explore *look* and *watch* which were substituted 6 and 5 times each. Participants ranged fairly evenly between A2-B2 CEFR levels. *Look* was substituted for *watch* (n=4), *show* (n=1) and *view* (n=1), whereas watch was substituted for *see* (n=4) and *visit* (n=1).

(171)   sometimes I also **look** (watch) films or so in English <123>

(172)   sits there then he lets the woman **watch** (see) the picture that he painted <55>

This combination of NSU-16 with *look*, *watch* and *see* is interesting because it appears that the participants relied on both their L1 Swiss German and Standard German vocabulary. For example:

Swiss German   Mir tuent fernseh luege.

Exact English translation   We do television watch.

Standard German   Wir sehen/schauen fern / Fernsehen.

English   We see/watch television.

As can be seen, the Swiss German *luege* is similar to *look* and then standard German *sehen/schauen* is similar to *see/watch*.

5.1.17.10   Miscellaneous

The remainder of the NSU-16 instances fell under the category of miscellaneous and were single instances. There were two instances which are worth mentioning. English loan words have a tradition of finding their way into Swiss German (Bon, 1948, p. 232) and a fairly recent addition is *foodä* which is used as a verb to mean *eat* (Rickenmann, 2005, p. 123). In the example below the participant appears to have used *food* instead of *eat*.

> (173)   yes in the pool [L] and yeah you could sit around and yeah and you sit always in the water uhm yes uh huh you could always **food** everything <17>

The second example is a very good attempt to express complex thoughts. It was assumed that the participant meant people who are spiralling downwards internally as they try to improve their online image.

> (174)   yeah uhm I have lots of friends or not not close friends luckily but I I know lots of people who are uhm getting **down spiral** of wanting to have a better image on the outside and being feeling worse and worse on the inside and uhm they are just struggling with and they're using actually what their own personality because of things like that and if you I don't know anyone famous who has gone through something but I guess it's it's always the same trying to be better more famous for some people they earn their money with that </B145>

To sum up the use of NSU-16 between sub-corpora and CEFR levels, Table 5.32 gives an overview of all 807 instances. As mentioned at the beginning of this section, all but 3 participants used at least one NSU-16. In the last column, we see that the amount of NSU-16 per person hovers between 7 and 9 for most CEFR levels except for higher numbers for the C1

participants in sub-corpus 1 and A2 level participants in sub-corpus 2 and subsequent lower

levels in the highest CEFR levels of each sub-corpora. This fluctuation was confirmed when a

Pearson correlation coefficient was computed to assess the linear relationship between the

participants' CERF proficiency levels and use of NSU-16. There was not a significant

correlation between the two variables found, $r(5) = -.87$, $p = .054$.

Table 5.32 CEFR and NSU-16 use

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-16	Mean NSU-16 per participant
1	15	100.00	A2	109	7.26
1	19	100.00	B1	170	8.94
1	13	92.85	B2	109	8.38
1	2	66.66	C1	34	17.00
1	1	100.00	C2	2	2.00
2	10	100.00	A2	132	13.20
2	21	100.00	B1	175	8.33
2	10	100.00	B2	73	7.30
2	2	66.66	C1	3	1.50
Total	**93**			**807**	

Conclusion

There are a multitude of factors which influence language learners' lexical choices. This section

has attempted to analyse, categorise and illustrate the non-standard usage found in the CSC

corpus. First, distinct German interference was analysed with 25 words supplying 144 NSU-16

instances. Eight words considered common L1 errors were found in both sub-corpora divided

evenly between the CEFR levels with only the A2 level producing fewer. This was suspected

to be due to avoidance strategies. Ten mainly L1 phonologically influenced words were found

in sub-corpus 2 with the participants exhibiting a degree of creativity. The seven words used

solely in sub-corpus 2 on the other hand were slightly more complex and the interference was

not only phonological but meaning based. It is hypothesised that the use of multilingual

awareness techniques in the Swiss classroom might contribute to false friends and L1 interference.

Next, 6 instances of antonymy were illustrated before moving on to the introduction of phonetic similarities. The 'bathtub effect' (Aitchison, 2012) was considered to have contributed to the NSU-16 use of 17 words. The use of phonetic similarities was divided evenly between the sub-corpora. Furthermore, it was found that 60% of these NSU were used by low-intermediate proficiency participants.

After investigating the 26 instances of base word variation, the 30 most frequent words were investigated, and it was found that 19 of the 30 had a non-standard rate of over 5% when all occurrences in the CSC corpus were calculated with all NSU using that verb. This indicates the need for further studies to assess the need for future curriculum focus. The NSU-16 with countries and pronouns were found to be used by all proficiency levels and no significant difference between the sub-corpora. Moving on to adverb and adjective inversion, it was found that primarily B2 level participants replaced an adverb with an adjective or noun, whereas a range of proficiency levels replaced an adjective with an adverb.

During the investigation of numbers and quantity, it was observed that 60% of NSU-16 with numbers were used by the lower proficiency levels and 70% of quantity related NSU-16 were used by sub-corpus 2 participants. Further investigation would be necessary to discover the reasons for the difference. High semantic words were explored next with 87 instances found of high semantic words being confused with other high sematic words. This is contrary to the ELF hypothesis that high semantic words are overused. Lastly, two examples of NSU-16 miscellaneous lexical choices were illustrated to represent lexical dexterity.

5.1.18 NSU-17 Present, past perfect aspect and simple inversion

Description

The present perfect tense is formed with the auxiliary *have* or *has* and the past participle of the verb. The past perfect is formed with the auxiliary *had* and the past participle of the verb. Each represent time in the past. The present perfect has a connection with the present, whereas in the past perfect actions began and ended in the past. There were no instances of perfect continuous NSU in the CSC corpus. Both forms require command of the irregular verbs and 'the acquisition of the present perfect is regarded as the single most challenging task for non-native speakers (Fuchs et al., 2016, p. 297).

Classroom introduction

The sub-corpus 1 coursebooks cover present perfect towards the end of grade nine and past perfect is not covered. In the sub-corpus 2 coursebooks, perfect tenses are introduced in conjunction with a repetition of the past simple in the middle of grade seven and repeated in the beginning of grade eight.

Previous research

In their trilingual Swiss English Database, both Rosenberger (2009, p. 165) and Dröschel (2011, p. 229) recorded non-standard usages of the perfect aspect with 77% of those underuse and 23% overuse. Dröschel concluded that 'Swiss speakers of English generally tend to avoid rather than overuse the perfect aspect' (2011, p. 231). The tendency for German L1 learners to underuse the present perfect is reinforced by Fuchs et al. (2016, p. 323) and Davydova (2011, p. 277).

Findings

A total of 24 instances of NSU-17 were found and divided into 3 categories, present or past perfect insertion, present or past perfect omission and present perfect instead of past perfect. The examples below illustrate this with intended tense in brackets.

(175) it's completely different than here at breakfast we **have eaten** (ate) uh some Bohnen <111>

(176) with my family uhm **I wasn't** (haven't been) alone on holiday till yet so always with my family <114>

(177) I went to Rome like two or three ye years ago and **I've** (had) **never been** in a big city before so it was quite a new experience <61>

Contrary to the Swiss English Database 75% (n=15) of the NSU-17 were insertions and 25% (n=5) omissions. In addition, there were 4 instances of present perfect instead of past perfect. As can be seen in Table 5.33 below, the distribution of perfect aspect NSU-17 between the sub-corpora differs considerably. Only 4 sub-corpora 1 participants used NSU-17 in 5 instances with the C2 example made by a native speaker. The low usage is probably related to presumed late introduction of perfect aspect in the sub-corpus 1 curriculum. The sub-corpus 2 numbers, on the other hand, show attempts to use the perfect aspect in all proficiency levels and can probably be contributed to a scaffolding teaching approach as indicated in the coursebook syllabi. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-17. There was not a significant correlation between the two variables found, $r(5) = -.71$, $p = .175$.

Table 5.33 CEFR and NSU-17 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-17	Mean NSU-17 per participant
1	1	6.66	B1	1	1.00
1	2	10.52	B2	3	1.50
1	1	33.33	C1	1	1.00
2	4	40.00	A2	7	1.75
2	4	19.04	B1	5	1.25
2	3	30.00	B2	7	2.33
Total	**15**			**24**	**1.60**

In their study on intermediate to advanced German L1 learners, Fuchs et al. (2016) found that early learners used present perfect more often, thus more native like. Other variables such as time abroad and length of university English studies were not found to be significant. With only 5 instances of NSU-17 in sub-corpora 1 the frequency of perfect aspect was investigated to determine if this was also evident in a mixed ability corpus. Because the CSC was not tagged for parts of speech, a calculation of the ratio of simple past to perfect aspect is not feasible. However, an investigation into the ratio of perfect aspect to non-standard usages is. AntConc searches of *have*, *haven't*, *has*, *hasn't*, *hadn't*, *had* and *'ve, *'d, *'s were conducted and non-perfect uses eliminated. Thus, 186 instances of perfect aspect were found. This includes the insertion and inverted usages of NSU-17. Table 5.34 shows the distribution. We see that there is a general increased use of perfect aspect with an increase in CEFR levels and lower frequency of perfect aspect use in A2 and B1 levels in sub-corpus 1. However, when a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of perfect aspect. There was not a significant correlation between the two variables found, $r(5) = -.19$, $p = .753$.

Table 5.34 Number of perfect aspect in CSC

Sub-corpus	No of participants	% of participants who used perfect aspect	Proficiency level	No of perfect aspect uses	Mean perfect aspect use per participant
1	1	6.66	A2	1	1.00
1	7	36.84	B1	12	1.70
1	9	64.28	B2	36	4.00
1	2	66.66	C1	15	7.50
1	1	100.00	C2	6	6.00
2	6	60.00	A2	18	3.00
2	15	71.42	B1	31	2.06
2	9	90.00	B2	57	6.33
2	3	100.00	C1	10	3.33
Total	**53**			**186**	**3.50**

The 186 instances equate to a perfect aspect use of 1.88 per thousand words (ptw) which is less than the 4.9 ptw found by Rogatcheva (2016, p. 150). Her specialized corpus of German L1 learners was a written corpus which helps explain the higher number of perfect aspect use. However, if compared to the number of words in the German sub-corpora of the Spoken LINDSEI corpus (Gilquin et al., 2010, p. 27) and the number of perfect aspect found by Fuchs et al. (2016, p. 335) the use of perfect aspect at 1.88 ptw in the CSC is comparable to the 2.20 ptw instances of perfect aspect in the LINDSEI-GE. The increased use of perfect aspect in early English learners by Fuchs et al. (2016, p. 322) is reinforced when the CSC sub-corpora are examined. Sub-corpora 1 which can be considered late starters used the perfect aspect 1.44 ptw compared to the use of 2.32 ptw by the early English starters.

Two factors still need to be considered when interpreting the data. First, 21 of the instances in sub-corpus 1 and 8 instances from sub-corpus 2 were used by quasi native speakers with at least one parent a native speaker. Secondly, 8 of the instances were ambiguous because of the interview part two storytelling picture descriptions it was not possible to determine if *he's*

finished was meant as *he is finished,* or *he has finished*. These examples were included. Therefore, if we were to eliminate those 29 instances, the results would show lower values.

Conclusion

The use of perfect aspect was found to be similar to the LINDSEI German sub-corpora in two respects. Firstly, the amount of perfect aspect used are comparable with 1.88 ptw in the CSC and 2.20 ptw in LINDSEI-GE. Secondly, the early starters in the CSC produced more instances of perfect aspect than the late starters. When NSU-17 were evaluated a 75% rate of insertion was found which contradicts Rosenberger's findings in the Swiss English Database. Differences in the use of perfect aspect between the sub-corpora might be attributed to the late classroom introduction towards the end of their formal classroom instruction period.

5.1.19 NSU-18 Redundant use of continuous forms and *-ing*

Description

Choosing the correct verb form can pose a problem because learners have several options. Continuous verb structures are formed with the verb *to be* and adding *–ing* to the verb. The continuous form signifies the speaker's interpretation of a continuous or temporary action. This form can be used to describe actions in the past, present or future (Swan, 2016, p. 3). Another use of *–ing* is when it is used like an adjective or noun and some verbs are followed by an *–ing* and others are followed by an infinitive. Most of these need to be learned with the individual words. This section will investigate redundant use of *-ing*, whereas omissions will be dealt with in the following section.

In written High German there is no present continuous form. It is possible, though, to add the terms *gerade*, *eben*, *noch*, *nun*, or *jetzt* to indicate something is happening now, at the time of

speaking (Krause, 1997, p. 3). Some German dialects and Swiss German use the verb *to be* with a verb to construct a continuous form; however, the word order differs from English and High German. See example below:

Standard English: I am looking for my glasses.

High German: Ich suche (gerade) meine Brille.

Swiss German: Ich bi mini Brille am sueche.

In the Swiss context, having numerous variations can lead to an redundant use or omission of present continuous depending on which variation of German the speaker calls upon for reference before the English form is internalized.

Classroom introduction

As mentioned in 5.1.11, the sub-corpus 1 coursebooks, cover present simple as the first item and is followed with present continuous towards the end of grade seven. In the sub-corpus 2 coursebooks, on the other hand, tenses are introduced early and then reviewed repeatedly until the various aspects have all been covered. Additionally, the sub-corpus 1 coursebooks cover verbs with prepositions and gerunds at the end of grade eight. However, there is no indication that continuous forms are taught in the sub-corpus 2 coursebooks according to the scope and sequence.

Previous research

Previous research suggests that advanced German L1 learners of English use the continuous in speech more frequently than in writing (Dose-Heidelmayer and Götz, 2016, p. 235). Nevertheless, their use of continuous was found to be significantly less than native speakers when results from LINDSEI-GE and LOCNEC were compared (ibid. p. 236). Results from the

Swiss English Database were minimally explored due to their low frequency across all Swiss languages but give insight into the Swiss German L1 data. Both studies found more non-standard uses of continuous overuse than underuse.

Findings

There were 69 instances of NSU-18 found and they were divided into three categories. First, 25 instances had a full continuous form with verb to be and *-ing*. In 20 instances this was used instead of the present simple and in 5 instances instead of the past. See examples below.

> (178) yes uhm no no uhm when I am drunk I **am speaking (speak)** much better
>
> English because there are the <17>
>
> (179) I just had an English teacher with uh a boyfriend who **was speaking (spoke)**
>
> English as a mother language yes <110>

The second category was the use of *-ing* without the verb to be. There were 40 instances with *-ing* instead of the present simple in 35 cases and in 5 instances instead of the past simple as seen in examples below.

> (180) then the doctor said yeah he has to look at this picture and uhm **speaking**
>
> **(speak)** with the other doctors and blah blah so <125>
>
> (181) and they were pretty nice people and we laughed and **drinking (drank)** a little
>
> bit there yeah they were nice peoples <126>

In the last category, the verb *to be* was substituted for the preposition *for* in 4 instances. The example below illustrates this.

(182)   and then I went uhm for a month working in a factory **for earning (to earn)**

money uhm and uhm then I started <1>

In Table 5.35 below we see that twice as many sub-corpus 1 participants used twice as many NSU-18. One main difference is in the B1 level use, with 47.36% participants from the sub-corpora 1 averaging 2.33 NSU-18 per person compared to only 14.28% participants from sub-corpus 2 averaging 1.33 NSU-18 per person. This fluctuation was reflected in the results when a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-18. There was not a significant correlation between the two variables found, $r(5) = -.65$, $p = .227$.

Table 5.35 CEFR and NSU-18 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-18	Mean NSU-18 per participant
1	6	40.00	A2	8	1.33
1	9	47.36	B1	21	2.33
1	7	50.00	B2	13	1.85
1	1	33.33	C1	4	4.00
1	1	100.00	C2	1	1.00
2	5	50.00	A2	6	1.20
2	3	14.28	B1	4	1.33
2	4	40.00	B2	12	3.00
Total	36			**69**	**1.91**

Conclusion

The CSC NSU-18 of 0.70 ptw is twice the reported number of the Swiss English Database with 0.35 ptw from Swiss German L1 participants. This can be because the database includes written data and assumed higher proficiency levels. A more suitable comparison with LINDSEI-GE can be made where both corpora data included interviews with the same picture description and

thus equal opportunities to produce continuous forms. Here similar findings were made with 0.69 ptw in the LINDSEI-GE. Dose-Heidelmayer and Götz, (2016, p. 242) however, found that relatively few participants used the majority of NSU. The NSU-18 in the CSC were more evenly spread between the participants which might be due to overall lower levels of proficiency.

5.1.20 NSU-19 Omission of continuous forms and *-ing*

Description and Classroom introduction

The same grammar description and time when taught applies as in Section 5.1.19 above.

Findings

There were 21 instances of NSU-19 found. Of those, only 10 were the omission of present continuous, see Example 183 below, the remaining 11 were noun like *-ing* (n=3), see Example 184 below, *-ing* forms after a preposition (n=4), see Example 185 below, *-ing* forms as adjectives (n=2), see Example 186 below and nouns (n=2), see Example 187 below.

(183) then she's **try (trying)** to write back uh Iron Man catched the shield because his reflexes were so fast <138>

(184) that's uhm exactly what you do not learn in school in school we it's all about **express (expressing)** yourself precise uhm precisely and as beautiful as possible <144>

(185) learning it and sometimes I can't or I don't want to just think about the grammatical because then it would stop me **to talk (talking)** and when I'm there like I was in London and I just wanted something to buy <56>

(186) I say Rome uhm the vati Vatican oder yes and uhm it is very **interest (interesting)** <9>

(187)   yes so I think the woman will have uh a **paint (painting)** of of uh herself so he

engaged uh uh a painter <135>

Although continuous aspect is considered difficult for German L1 learners, the NSU-19 are concentrated in both sub-corpora at the lowest proficiency level with little difference between the sub-corpora. Furthermore, a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-19. There was a significant correlation between the two variables found, $r(5) = -.93$, $p = .018$. Thus, omission of continuous forms and *-ing* is marginal and appears to be proficiency bound to lower levels of the CSC.

Table 5.36 CEFR and NSU-19 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-19	Mean NSU-19 per participant
1	5	33.33	A2	5	1.00
1	2	10.53	B1	2	1.00
1	2	14.29	B2	2	1.00
2	3	30.00	A2	6	2.00
2	3	14.29	B1	3	1.00
2	2	20.00	B2	3	1.50
Total	**17**			**21**	**1.23**

Conclusion

The omission of continuous and *-ing* is in line with the research findings mentioned in Section 5.1.19 with 22.47% of the NSU concerning continuous and *-ing* falling in the omission category. In conclusion, although the picture description could have given the participants ample opportunity to use continuous forms, only 12 (13.48%) NSU-18/19 were used relating to the picture description were found.

5.1.21 NSU-20 Word order, redundant or omitted words

Description

This category of non-standard usages became a general category for instances of word order, redundant and omitted words. Both omitted and redundant words are not uncommon in spoken interactions as speech is being constructed as it is spoken, and thoughts can change mid-sentence. Word order, on the other hand, follows a more defined structure. There were discrepancies found in sentence beginning and endings, time, place, adjectives, adverbs, adjectives, subject verb alignment and other miscellaneous instances that were deemed non-standard.

Classroom introduction

Although word order was not mentioned in the scope and sequence of the sub-corpus 1 or 2 coursebooks, it is assumed that it was taught throughout both coursebook series.

Findings

There were 265 instances of NSU-20 found with 98 word omissions, 69 redundant words and 98 instances of non-standard word order. Starting with the 98 omissions, 49 different words were used by 55 participants which amounts to 57.69% of sub-corpus 1 participants and 56.81% sub-corpus 2 participants. Of the 49 words, 22 were used multiple times and the remaining 27 only once. Within the 22 groups of words used multiple times, there were only 4 instances of a participant omitting the same word twice. Hence, we can presume that omitting words can be considered widespread among this group of participants. The three most common word omissions were *it* (n=10), *when* (n=6) and *time(s)* (n=5).

In the concordance lines below, definite patterns can be seen. The pronoun *it* has been omitted before *is* in 60% of the lines and all are mid-sentence omissions except for the first line where *it* would end the sentence. Next, *when* was omitted when referring to past or future time and in half of the lines it was omitted in a comparison. Lastly, *time* was omitted 3 times when referring to the number of occurrences and omitted twice when referring to past periods of time.

I go shopping in the city there	**it** is uhm not more expensive as there
yeah but it's okay when when then	**it** is finished and I make good money then it's okay
yes it's uhm they have no sel and the bread and so	**it** is not sel it is so so kein Geschmack
in eh summer is too cold and sometimes	**it** is raining yes and uhm and in summer is eh
it has something in common but I think there	**it** is really the poor are really poor
four seasons and in Sweden is in eh summer	**it** is too cold and sometimes
because uhm I don't use	**it** that much that I can say uh I'm a user
another one and then finally she liked it and showed	**it** to the other
and we got everything food and in the hotel	**it** wasn't that good uhm you can imagine in Spain [L] it isn't
I couldn't after I take I I I was not used to	**it** I got really really like not crazy but uh yeah
I think my voice is is yeah is not the same as	**when** I speak German
uh I can't uh express me the much as	**when** I talk in German so sometimes I it's quite difficult
it's another language it's not so comfortable as	**when** I talk Swiss German but I really like English its uhm first
there and the friends didn't want to go with us even	**when** we were saying let's go there or just making something
and yeah they in the city	**when** you walk there are many funny peoples
learn or use I don't know but	**when** I married or with my wife in the uh I don't know in
then after the picture uh after he drew it the second	**time** uh the woman gets uh beautifuler and after that the girl
yes it was there yeah some	**time** before the holidays yeah it was really warm
yes we went one	**time** to (uh huh) what's the name outlet
this uh because the people in the uh in the early	**times** before uh I think thou thousand years or more uh they
think English uh helped me also uh just in different	**times** uh when I play games videogames

Next, redundant words were found in 69 instances and used by 41 participants which amounts to 48.07% of sub-corpus 1 participants and 36.36% sub-corpus 2 participants. There were 38 different words used with 15 being used multiple times and 23 only once. Within this category, there are two types of redundancy. First, there are redundant words with a grammatical significance as seen in the examples below:

(188)   I was reading a lot of books in in Ireland and she recommended **me** some some

   of those and she gave me one and uhm <141>

(189)   good when you when you doesn't understand you can explain **you** <B135>

(190)   yes of course that's **more** cheaper that here <B58>

(191)   it's uh there's a woman and she go to a painter store **draw** and he make a paint

   from her and paint (uhm) <42>

In Examples 188 and 189, the L1 German transfer is suspected because *recommend* and *explain* can be followed by a pronoun in German, but not in Standard English. The third example (190) is the redundant use of *more* in a superlative one-syllable word. The fourth sentence (191) is from an A2 level speaker who used phrases and key words to describe the picture situation. This type of NSU-20 could require more effort from the listener to comprehend. The second category is of false starts, misspeaking and mid-sentence directional changes which are characteristic of spoken discourse (Schachter et al., 1991, p. 362). The following examples illustrate this.

(192)   six infinity stones those are stones that created the universe with it and gave

   (it) life to it and you can also use them to destroy half of it <138>

(193)   not from a country who speaks English because uhm **(I)** here in Lucerne when I

   meet some peoples from I don't know India or from China then I speak spoke

   with them in English <24>

(194)   I'll ask you like where are you from what are you doing here because yeah my

   (French) uh my friend with I went with she has coloured blond <63>

(195) yeah I thought I have to die [L] and my little brother cried and **(we)** because it

was uhm in the evening it was ten o'clock and the ranger went away so we were

alone in the desert <107>

Examples 192 and 193 illustrate false starts which are then corrected. In Example 194 the speaker misspoke with *French* instead of *friend* but corrected themselves. The speaker in Example 195 appears to have started with **we** and then realised that more information was necessary to explain '*we were alone in the desert'*. This category of NSU-20 can be considered non-impeding.

Non-standard word order is the last category, and the 98 instances were divided up into 9 sub-categories with the sub-category name expressing the area of word order non-conformity: fronts (n=10), endings (n=12), adverbs (n=5), subject-verb-object (n=19), time (n=14), place (n=8), questions (n=5), adjectives (n=13) and miscellaneous (n=12). An example of each sub-category follows. Many of these examples can hinder comprehension and thus increase the need for repetition or further explanation.

(196) yes because uhm yeah *my dream is it to be free* (**it is my dream to be free**) and

uhm in the garden <125>

(197) speak a little bit more with my hands and something like that because *you can't*

everything express (**you can't express everything**) with your with your uhm

language <11>

(198) when I have *the language English* (**the English language**) </B6>

(199) I don't know I like the spe speciality *ajvar it is ajvar called* (**it is called ajvar**)

<142>

240

(200) *we were three times there* (**we were there three times**) that's a beautiful nation too a lot of nice <126>

(201) or *go home back* (**go back home**) alone <63>

(202) and then the girl asked me *do I have ever build a bomb* (**have I ever built a bomb**) <139>

(203) he is a very good drawer he can *exactly draw her face* (**draw her face exactly**) and her mimic he can draw it and ahh <125>

(204) *need help and then I ask her uhm to help to* (**I asked her if she needed help**) going into the lift and up to her <104>

As can be seen in Table 5.37 below, 81.25% (n=78) of the CSC participants used NSU-20. All proficiency levels were found in the sub-categories above. However, the percentage of participants who used NSU-20 gradually decreased as the proficiency levels increased except for the C1 participants. This is due to the low number of C1 participants and the type of NSU-20. Five out of the six C1 NSU-20 can be categorised as non-impeding. Further, we see that although the percentage of participants decreases, the average NSU-20 per person fluctuates with sub-corpus 2 using more NSU-20 overall. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-20. There was not a significant correlation between the two variables found, $r(5) = -.82$, $p = .085$.

Table 5.37 CEFR and NSU-20 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-20	Mean NSU-20 per participant
1	14	93.33	A2	39	2.78
1	17	89.47	B1	71	4.17
1	9	64.29	B2	45	5.00
1	3	100.00	C1	6	2.00
2	10	100.00	A2	35	3.50

2	19	90.48	B1	49	2.57
2	6	60.00	B2	20	3.33
Total	**78**			**265**	**3.39**

Conclusion

This general category encompasses three sub-categories. Omissions and redundant words can be loosely placed in the field of disfluencies where many can be considered non impeding. The omittance of words was found to be widespread throughout the corpus and redundant words were used slightly less. Non-standard word order, on the other hand, was found to be impeding.

5.1.22 NSU-21 Word order inversion of adverbials of frequency and focus

Description

This category is a sub-section of the previous section dealing with word order. Although Swan (2016, p. 196) claims that it is impossible to state reliable rules for the position of adverbials in sentence structure due to the complexity of the area, the word order inversions in the CSC were found to be straightforward. Firstly, adverbials of frequency and focusing are usually placed mid-position between the verb and its object as in:

He often speaks English on holiday.

My friend also wants to go on holiday.

Mid-position adverbials usually go before one-word verbs, after auxiliary verbs and after *am/are/is/was/were* (Swan, 2016, p. 199). Swiss German speakers often use adverbs of frequency and the focusing adverb *also* in a non-standard manner. This is probably due to L1 interference because in Swiss German as well as High German the adverbial generally follows the verb, although word order with adverbs of frequency differs in Swiss and High German as the following examples show.

High German	Er spricht im Urlaub oft Englisch.
Word for word translation	He speaks in the holiday often English.
Swiss German	Er reded oft[31] Englisch i de Ferie.
Word for word translation	He speaks often English in the holiday.

The focusing adverbial *also*, however, remains the same in Swiss and High German.

High German	Mein Freund möchte auch in den Urlaub fahren.
Word for word translation	My friend wants also in the holiday drive.
Swiss German	Min Fründ wett au i d'Ferie gah
Word for word translation	My friend wants also in the holiday go.

Classroom introduction

The sub-corpus 1 coursebooks cover adverbs of frequency quite early when only the present simple tense has been taught in grade seven. The sub-corpus 2 courses books also cover adverbs of frequency in grade seven, but they had already covered most of the tenses by then.

Previous research

Durham found that the participants in the German part of the Swiss English Database used additive adverbials at a rate of 51% *also*, 40% *as well* and 9% *too* and suggests that this is due to *also* being a lexical item in German (Durham, 2007, p. 223). Her results backed up previous research from Fjelkestam-Nilsson (1983).

---

[31] *Viele* or *meischtens* used more often in Swiss German than *oft*

Findings

There were 45 instances of NSU-21 found. 12 of those involved the inversion of adverbs of frequency and 33 the inversion of focusing adverbials. The CSC contains a total of 380 adverbs of frequency which equates to a non-standard rate of 3.15%. The following CSC excerpts exemplify those NSU-21.

(205)   it was really interesting because (uhm) me and my sister we **have always** arguments and so <12>

(206)   I am interested in texts I **read often** the text from new singles <50>

Of the 33 NSU-21 involving focussing adverbials, 32 used *also* and 1 used *just*. The following words were used recurrently from 2 to 8 times and equate to a 2.9% NSU-21 rate of focus adverbials. The calculation includes all uses of the focus adverbials *also, just, even, only, mainly, mostly either, or, neither* and *nor* as listed in Swan (2016, p. 199).

(207)   yes a lot (uhm) at first I **had also** problems to sit the whole day in school <1>

(208)   but they are very friendly to me and (uhm) I **like also** the the city because it's always <55>

(209)   (eh) maybe you **speak also** in in English <25>

(210)   we had to **go also** (uhm) like I was in the city in the downtown <112>

(211)   and (uhm) I **think also** here is school you have so <1>

Although the classification of NSU-21 was straightforward, the use of *also* was found to be interesting. Firstly because of its prominence over *as well* and *too* and its use as a German code-switching discourse marker. *Also* was found 618 times in the CSC. Of those 146 were used as

a German discourse marker. During transcription, this was quite easy to interpret because there was a change in vocal tonality and obvious contextual meaning.

As mentioned above, Durham observed *as well* being used 40% of the time. The findings in the CSC show a very different outcome. With the German discourse markers are eliminated, *also* is used 78.01%, *as well* only 5.28% and *too* 16.68%. Just as Durham associated the higher use of *as well* with its use being closer to native like, the increased use of *also* and *too* might be an indication of the lower intermediate proficiency level of the CSC.

Table 5.38 below depicts the distribution of the adverbial word order inversion. Here we see that the majority of NSU-21 were used by participants with a B1 proficiency level. This distribution loosely reflects the distribution of all NSU, see Table 5.1 p. 142, which implies that NSU-21 was not correlated to proficiency levels but can be considered a feature of the CSC because 28 or 29.16% of the participants used this structure with 12 participants using it more than once. This was confirmed when a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-21. There was not a significant correlation between the two variables found, $r(5) = -.73$, $p = .155$.

Table 5.38 CEFR and NSU-21 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-21	Mean NSU-21 per participant
1	3	20.00	A2	4	1.33
1	6	31.57	B1	10	1.66
1	2	14.28	B2	3	1.5
1	1	33.33	C1	3	3.00
2	3	30.00	A2	6	2.00
2	7	33.33	B1	11	1.57
2	6	60.00	B2	8	1.33
Total	**28**			**45**	**1.60**

Conclusion

The non-standard use of NSU-21 was found to have been used by 29.16% of the CSC participants. However, when all the uses of adverbs of frequency and focus are calculated the rate of NSU-21 is only 2.83%.

5.1.23 NSU-22 Omission of will future

Description

The future can be communicated by using *present simple, present continuous*, *going to* or *will/shall*. 'When we are simply giving information about the future or talking about possible future events which are not already decided or obviously on the way, we usually use *will*. This is the most common way of talking about the future' (Swan, 2016, p. 35). *Will* is also used to announce decisions and make promises, threats and offers.

Classroom introduction

As mentioned in 5.1.12, *Will* is covered in the sub-corpus 1 coursebooks at the beginning of grade nine and is taught in the middle of grade seven in the sub-corpus 2 coursebooks.

Findings

There were 16 instances of USU-22 found where *will* was omitted. Every instance of will omission was preluded with the pronoun *I*. The speaker was expressing future actions with 11 instances expressing an immediate action such as in Example 1 below and 5 instances expressing an action further in the future. All but two instances were used in the initial interview phase and most of them to begin the monologue about travelling to a country.

1	okay uhm I	**talk**	about uhm Russia I lived there for a half
2	oh okay [L] uhm I	**describe**	the film Aviator you know this it's a very
3	hum … uhm I	**do**	topic two
4	uhm no if I go to another uh country I	**just go**	go travel independently
5	tomorrow I	**go**	
6	yeah economics and then maybe then I	**search**	a job in New York or so yeah
7	mmm em I'll take I	**take**	topic two
8	… I think I	**take**	the topic two
9	uhm I	**take**	China I was there and uhm for two years
10	I'll choose topic two yes I	**take**	I think the US that's w what really impressed me
11	uh huh … hum … uhm okay I think I	**take**	the country actually uhm can I start
12	… okay I think I	**talk**	about topic 2
13	okay uhm I	**talk**	about topic two a country
14	uhm okay I	**talk**	about a country
15	people in this in Switzerland because I	**work**	after this school I work in the same
16	after this school I	**work**	in the same place like jetzt and we

To understand the overall frequency of *will* in the CSC concordance searches for *'ll* and *will* were conducted. The results show that there were 183 instances. This calculates to an NSU-22 rate of 8.74%. If we then evaluate the NSU-22 rate of the most frequently use verb (take) with *will* or *'ll* there is a NSU-22 rate of 18.51%.

There were twice as many instances of *will* future omission in sub-corpus 1 than sub-corpus 2 as can be seen in Table 5.39 below. It is also interesting to see that the omission of *will* was used primarily with intermediate users at a B1 and B2 proficiency level. The late introduction of the *will* future might have contributed to this high rate of non-standard use. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-22. There was not a significant correlation between the two variables found, $r(5) = -.42$, $p = .473$.

Table 5.39 CEFR and NSU-22 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-22	Mean NSU-22 per participant
1	1	6.66	A2	1	1.00
1	3	15.78	B1	4	1.33
1	4	28.57	B2	5	1.25
2	2	20.00	A2	2	1.00
2	4	40.00	B2	4	1.00
Total	**14**			**16**	**1.14**

Conclusion

The NSU-22 rate of will was found to be high at 8.74% and the NSU-22 rate with the most frequent verb *take* even higher at 18.51%. This indicates that possibly more attention should be paid to the *will* form in the classroom with special emphasis on the phrase *I'll take*.

5.1.24 NSU-23 Prepositional possessive phrase

Description

Possessive phrases are often used to talk about possessions, relationships and physical characteristics (Swan, 2016, p. 124). They are formed by adding an *'s* to the noun. In very long expressions an *of*-structure is preferred, but not usual in shorter structures and is not used with people (ibid.).

Classroom introduction

The sub-corpus 1 coursebooks cover possessive *'s* twice. First at the end of grade seven and again at the end of grade eight. Sub-corpus 2 coursebooks do not cover them until the middle of grade nine.

Findings

There were 17 instances of NSU-23 found in the corpus which were used by 12 different participants, which is equivalent to 12.5% of all participants. This phenomenon appears to be unique because there were no similar instances found in either the 1994 BNC spoken corpus or the German LINDSEI corpus. The NSU-23 can be divided into two groups. The first group uses an *of*-structure where an *-s* would usually be used as seen in the examples below. The last two examples might be deemed acceptable because of the length of the phrase.

Friends there and she also uhm knows	**the family of my friends** and the owner of the hotel
uhm not really my also	**the son of my uncle** is here in Switzerland but we
and the	**the kids of the brother** eh speak eh Swiss German but
speak eh Swiss German but	**the kids of their sister** they don't speak Swiss German
and this is Mark uhm	**friend of uhm of my brother** he's
and after that the girl shows it to	**the friends of her** like uh she drew me
doesn't have a really good relationship and with	**the friend of his mother** so with his girlfriend he has
uh I speak sometimes uh with uh uh with uh	**the sister of my girlfriend**
him this evening something and	**the father of a friend** called us we should go to him
yeah without cows uh uh	**the brother of my father's girlfriend** was a trucker
and yes have a good time then the third week we visited	**the family of my father's girlfriend** it was also really

The second group uses the preposition *from* instead of *of* to form a *from*-structure as seen below. Again, none of these structures were found in either the 1994 BNC spoken corpus or German LINDSEI spoken corpus.

Uncles in America two uh th	they're are two **brothers from my mother**
	the **cousins from the uhm from the brother**
and then	**the sister from my father**
yes I think it's	**the cousin from my dad** and he went there like 20 years
yeah I have family in Austria and uh	**the cousin from my grandmother**
or like also in the same movie uhm a girl	his **sister from the superhero** made some shoes that

Two hypotheses were formed to explain why the participants used these forms. Firstly, and most logical is the Swiss German L1 influence and transfer as in the example below of *my father's sister*.

Swiss German   *d'schwöster vo mim Vater*

the sister from my father

High German   *die Schwester meines Vaters*

the sister mine father

The addition of *vo* (from) in Swiss German makes it unique and would most probably not be found in a German learner corpus from Germany or Austria.

The second hypothesis stems from a discussion with students who mentioned the use of the now common sayings to express close friendship *brother from another mother* and *sister from another mister*. As seen below, a Google ngrams search shows that the sayings were non-existent until 1995 when *brother from another mother* first appeared and *sister from another mister* appeared in 2005. It is possible that some CSC participants used the *from*-structure to express possession due to exposure to these sayings or and in connection with Swiss German L1 transfer. This is the third indication that Swiss youth might be gleaning lexical variation from popular vernacular expressions.

Google Books Ngram Viewer

my father's sister,brother from another mother,sister from another mister,sister fr...

1800 - 2019 ▾   English (2019) ▾   Case-Insensitive   Smoothing ▾

ℹ Replaced **my father's sister** with **my father 's sister** to match how we processed the books.
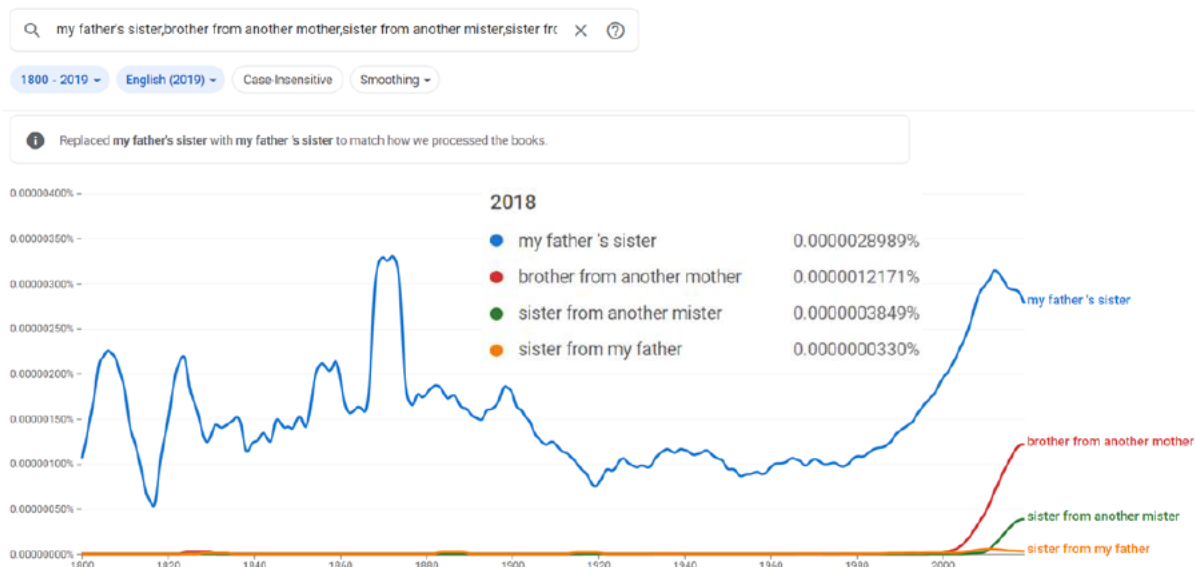
Turning to the distribution in the sub-corpora, in sub-corpus 1 10 instances were found which were made by only 5 participants with 8 out of 11 having an A2 proficiency level. The instances in sub-corpus 2, on the other hand, were more prominent in the B1 and B2 levels and made by 7 different participants. This suggests that there is a difference in how the participants in each sub-corpus use prepositional possessive phrases. A Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-23, there was a significant correlation found between the two variables, $r(5) = -.92$, $p = .023$.

Table 5.40 CEFR and NSU-23 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-23	Mean NSU-23 per participant
1	3	20.00	A2	8	2.66
1	1	5.26	B1	1	1.00
1	1	7.14	B2	1	1.00
2	2	20.00	A2	2	1.00
2	3	14.28	B1	3	1.00
2	2	20.00	B2	2	1.00
Total	**12**			**17**	**1.41**

A concordance search of *'s* revealed that out of the 1,898 instances of *'s* only 16 were used as possessives and were almost equally divided between the sub-corpora. This means that over 50% of the attempts to express possession with an *'s* were non-standard.

Conclusion

Prepositional possessive phrases with *of*-structure were found to be inserted. In addition, the use of *from*-structure was found and it was hypothesised that this could be from Swiss German transfer or influenced by the use of new terminology including *from a brother* or *sister*. This finding could signify the beginning of language change or a current fad. Nevertheless, further investigation would be necessary to make conclusions.

5.1.25 NSU-24 Negation inconsistencies

Description

In Standard English, there are basic rules for forming negative verb forms. The following 5 are stated by Swan (2016 p. 217).

- We make negative verb forms by putting *not* after an auxiliary verb.

- *Do* is normally used if there is no other auxiliary verb.

- *Do* is followed by the infinitive without to.

- *Do* is not used with another auxiliary verb.

- *Do* is not normally used with *be*.

Classroom introduction

Negation is covered throughout both the sub-corpus 1 and 2 coursebook series.

Previous research

It has been documented by Swan and Smith (2001, p. 41) that there is no equivalent in German for the English auxiliary *do* and that negation is done by adding *nicht* (not) after the verb. Therefore, it is most probable that the majority of the NSU-24 found in the CSC corpus were due to L1 interference.

Findings

There were 53 negation inconsistencies found which all but 7 could be divided into 3 major groups. The groups were *do* plus another *auxiliary*, omission of *do/did* where necessary and inconsistencies involving *past/present* and verb *to be*. The remaining 7 instances involved word order (n=2), double negatives (n=2) and use of *no* instead of *not* (n=3). Firstly, there were 17 instances of NSU-24 found where the use of *don't* or *didn't* and the auxiliary *can* or *must* were used. Within this group, there were definite differences found in the language use and proficiency levels. The 11 participants with an A2 level all followed the example below:

(212) **I don't can** say that in French then I say that <135>

The remaining 6 participants with higher proficiency levels 4=B1, 1=B2, 1=C1 also used *do* plus an auxiliary, however, slightly more complex with past tense, use of *nobody*, and *must not can* as the examples below show.

(213) just uh yeah and uh yeah nobody **nobody can't** make something <50>

(214) very strict eh with rules yes so we uhm yes we we **didn't can** drink beer <33>

(215) but she told me yes when I have holidays **don't must** be here and she <38>

Next, the omission of *do* was found 22 times. Here the majority (n=13) had a B1 proficiency level and the remaining an A2 level. In this group, there were no clear differences between NSU-24 found and proficiency level. All 22 NSU-24 from this group can be considered typical German L1 influence with *not* placed after the verb as the examples below demonstrate. This reinforces previous observations by Swan and Smith (2001, p. 41).

(216) I think it's difficult yes **I have not** the biggest vocabulary I think <137>

(217) they speak and speak and you yes **I speak not** so good Spanish the basics <36>

(218) yes yeah more I use it then when I can but yeah I **I learned not** English yeah <136>

The last group of 7 instances can also be considered learner errors with the majority having an A2 proficiency level. Here we see *do* and the verb *to be* as well as present and past tense inversion.

(219) and now uh I'm ausgezogen **I am don't** living at home anymore <126>

(220) was finished she she uh and she see that picture she **she don't was** happy because it's not so beautiful for <129>

(221) I thought my speaking was not perfect but **I don't had** to think I just talked <22>

Although NSU-24 were evenly distributed between the sub-corpora, Table 5.41 below reveals that within the sub-corpora there are differences. In sub-corpus 1, the lower proficiency levels are dominant with slightly more B1, whereas A2 participants dominate NSU-24 in sub-corpus

2. Three participants account for multiple NSU-24 in sub-corpus 1 compared to 5 multiple NSU-24 participants in sub-corpus 2. Furthermore, a Pearson correlation coefficient was computed to assess the linear relationship between the participants' CERF proficiency levels and use of NSU-24. There was a significant correlation between the two variables found, $r(5) = -.93$, $p = .018$. This indicates that the non-standard use of expressing negation is a feature of lower proficiency levels and can be considered a learner error.

Table 5.41 CEFR and NSU-24 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-24	Mean NSU-24 per participant
1	7	46.66	A2	11	1.57
1	7	36.84	B1	12	1.71
1	3	21.42	B2	3	1.00
1	1	33.33	C1	1	1.00
2	7	70.00	A2	18	2.57
2	5	23.80	B1	7	1.40
2	1	10.00	B2	1	1.00
Total	**31**			53	1.70

Conclusion

The non-standard use of negation was found to be primarily due to German L1 influence and linked to lower proficiency levels. It is interesting to observe that in the CSC the use of *I don't can* can be attributed to A2 learners and B1 learners made similar errors, but with slightly more complexity.

5.1.26 NSU-25 *The* or *it* instead of personal pronouns

Description

Personal pronouns refer to the person or people speaking, being spoken to and other people or things. They can function as a subject or object in a sentence, whereas possessive determiners

are used at the beginning of noun phrases and express possession. 'When talking about someone's possessions, or parts of their body, we usually use possessives, not *the*.' (Swan, 2016, p. 142). The exception is when we talk about something that happens to the body like *getting hit in the head* (ibid.).

Classroom introduction

The sub-corpus 1 coursebooks cover personal pronouns in the first chapter of their first year of study in grade seven. While the sub-corpus 2 coursebooks cover personal pronouns in the first unit in grade five in the third year of studying English.

Findings

There were only 5 instances found in the corpus and were highlighted due to their obstructing comprehension. As can be seen in the examples below, in the first two NSU-25 the personal pronoun has been substituted for *the* or *it.* The last three examples all substitute *the* for possessive determiners. It is interesting that this type of NSU is not proficiency level bound. Table 5.42 also shows that it was found in both sub-corpora. It can only be assumed that these NSU were slips of the tongue or German L1 interference. Although *the* has been inserted, it is not typical of German speakers as in Section 5.1.4 where the insertion of *the* is redundant because a substitution has been made here.

> (222)   the artist must draw a new picture of him and **it** (he) made a fake picture and
>   the woman tells him <121>
>
> (223)   the girl wanted to make something that impresses the others or **the** (her)
>   friends or so and she went to a <12>
>
> (234)   oh no it wasn't **the** (their) first language <58>

(225) just German was **the** (their) first language <58>

(226) we put this on **the** (our) body and played in the rain it's also fun you make uh

mud fights <121>

Table 5.42 CEFR and NSU-25 frequency

Sub-corpus	No of participants	% of participants	Proficiency level	No of NSU-25	Mean NSU-25 per participant
1	1	6.66	A2	2	2.00
1	1	33.33	C1	1	1.00
2	1	4.76	B1	2	2.00
Total	**3**			**5**	**1.66**

Conclusion

This type of NSU is not a typical learner error and the low frequency does not currently merit

further study. However, it reinforces the importance of corpus studies to find such instances

that might not have been recognised.

5.2   Summary of NSU findings

Chapters 4 and 5 have investigated the 2,642 non-standard usages found in the CSC corpus.

First, the relationship between demographic, identity-related and educational variables and

NSU per hundred words (phw) were investigated before examining 25 categories of non-

standard usages in greater detail. A summary of variables and NSU findings was given in

Section 4.3, as well as a summary of lexicogrammatical features of ELF findings in Section

5.1.10. This section will summarize the remaining findings. Further conclusions on all aspects

of the thesis will be drawn in Chapter 7.

As one of the objectives of this investigation was to determine the effect of early English on non-standard usages, findings related to differences between the sub-corpora will be addressed first. They will be followed by findings from the 25 NSU which were found to be proficiency bound and suggested areas of German L1 influence will be listed. Next, interesting patterns which were found will be summarised before specific words with high NSU rates are highlighted.

5.2.1 Early versus late English starters

In general, the participants in sub-corpus 1 used a greater number of NSU which signifies less accuracy in their spoken output. Further, it was found that sub-corpus 1 participants used twice and many present simple/simple past inversions (NSU-10) with regular verbs than sub-corpus 2. However, the use of NSU-10 with irregular verbs was equally divided between the sub-corpora. This suggests that the early starters had a greater command of regular verb use. Then, the use of *will* future instead of *want* (NSU-11) was found primarily in sub-corpus 1, used by A2-B2 level participants and only one A2 level sub-corpus 2 participant. This signifies a greater level of German L1 interference in sub-corpus 1. Turning to verb *to be* insertions and omissions (NSU-12), sub-corpus 1 participants with higher CEFR levels continued to insert the verb *to be* while sub-corpus 2 participants' insertions were concentrated in the lower CEFR levels. This was not the case with omission of verb *to be* where sub-corpus 2 participants, had higher NSU-12 use in all CEFR levels. This indicates that insertion and omission of verb *to be* could be affected by multiple variables and should therefore continue to be investigated separately.

The plural -s (NSU-15) was also found to be used by considerably more participants in sub-corpus 1. Non-standard usages in lexical choice (NSU-16), on the other hand, were found to be used by almost all participants in both sub-corpora. A difference was noted in the creativity and

complexity of lexis being higher in sub-corpus 2. An interesting observation was made that 70% of the 82 instances of quantity related NSU-16 (*all*, *some*, *any*, *less*, *little*, *few*, *a lot of*, *many* and *much)* were used by sub-corpus 2 participants. Further investigation would be needed to fully understand the reasons for this.

Sup-corpus 2 participants were also found to use perfect aspect more often which resulted in a higher number of perfect aspect related NSU-17. Lastly, twice as many instances of insertion of the continuous form (NSU-18) were used by sub-corpus 1 participants. However, lower CEFR level sub-corpus 2 participants omitted continuous forms more. It was speculated that the lower CEFR level participants could have used avoidance strategies.

Overall, sub-corpus 1 participants appear to have a marginally higher rate of basic grammatical errors when we reflect on the areas in which they used the highest number of NSU; namely third person singular -s, present/past tense inversion and plural -s. Sub-corpus 2 participants, on the other hand, appear to have exhibited more grammatical and lexical complexity with their increased use of continuous forms and perfect aspect.

In their 5-year longitudinal study of Swiss students, Pfenninger and Singleton (2019) suggest 'An earlier age of learning proved beneficial only for children raised as biliterate simultaneous bilinguals receiving substantial parental support, as opposed to monolinguals and nonbiliterate bilinguals (simultaneous or sequential)' (ibid. 2019). As these findings contradict the findings in the CSC, the studies were compared. Firstly, the CSC is comprised of a sample of the student population in Swiss classrooms, whereas the Pfenninger and Singleton (2019) study appears to include only Swiss students on a Matura path to university. This is suggested because they state that the second round of testing was conducted in secondary school classrooms over a two-

week period during English lessons when the students were between 18-19 years of age (Pfenninger and Singleton, 2019 p. 214) and only Swiss students studying for a Matura remain in secondary school past the age of 16. As mentioned in section 1.3.3, at the commencement of Pfenninger and Singleton's study in 2009, only approximately 20% of the population continued their education at upper-secondary school. In addition, to gain entry to upper-secondary school, either an entry test or the equivalent of proficiency of 84% across the subjects, German, French, English and Math was necessary. Thus, it can be assumed that the participants in the Pfenninger and Singleton (2019) study represent a group with a different scholastic background than the participants in this paper.

Another large difference is the mode of assessing language use. Pfenninger and Singleton (2019) conducted extensive tests on receptive and productive skills to evaluate language skills while the CSC concentrated on the oral production of non-standard language in interviews. Thus, it is not possible to directly compare the results of the two studies. Nonetheless, the results of each study present a different aspect of earlier language learning being better or not. Pfenninger and Singleton's (2019) study has suggested that for the top scholastic 20% of the population, the age at which English is taught has little difference on the results of tests conducted in the short or long term. On the other hand, the CSC suggests that early English has a positive effect on the accuracy of spoken English in the general public.

5.2.2 Proficiency bound features

There were 11 features that were found to be proficiency bound in the CSC. This information can be helpful in the investigation of developmental stages and CEFR levels.

- Omissions of indefinite articles *a/an* were concentrated at level A2, but omissions of definite article *the* were fewer in number and found at all CEFR levels

- All A2 level participants used present/past inversions, whereas use decreased as proficiency levels rose

- In sub-corpus 1 the insertion of the verb *to be* was concentrated in B1-B2 levels, but in sub-corpus 2 it was used primarily by A2 level participants

- Accuracy in plural use of *people* was found to precede accuracy in plural use of *person*

- The substitution of an adjective or noun instead of an adverb was used primarily by B2 level participants, whereas substitution of an adverb instead of an adjective was used by all levels

- 60% of all cardinal and ordinal numbers inversion, as well as multiplicative and ordinal number inversions, were used by A2 level participants

- Standard and non-standard use of perfect aspect was found to correlate with higher proficiency levels

- Omission of continuous forms was found to be used primarily by participants with the lower proficiency levels of A2-B1

- The large NSU-20 category of non-standard word order, redundant or omitted words was found to be used by fewer participants as proficiency levels increased with the exception of 3 C1 level participants.

- Omission of *will* future was found to be used mostly by B1-B2 participants

- Non-standard negation was found to be primarily used by participants with the lower proficiency levels of A2-B1.

5.2.3   German L1 interference

A strong German L1 interference was found to be correlated with the following non-standard uses:

- Confusing the relative pronouns *who* and *which*

- Insertion of redundant prepositions (especially *at*, *for*, *in* and *of*)

- Future tense *will* or *would* instead of *want*

- Omission and substitution of prepositions (especially substitution of *from*, *on* and *at*)

- Lexical choice – 5.45% of all NSU found in the corpus were attributed to German L1 interference in lexical choice

- Redundant words (especially *me*, *you* and *more*)

- Word order inversion of adverbials of frequency and focus

- Prepositional possessive phrases with *from*

- Omission of *do* in negation

This list is not exhaustive. As in all language learner interlanguage, each person's linguistic repertoire is varied and decisions made while speaking can be drawn from their individual experiences.

5.2.4   CSC patterns

While investigating the CSC corpus, several patterns were observed which are worth mentioning. First, there was a correlation found between 76.47% of the verbs used in a non-standard way with third person singular -s and their non-standard use in present simple and past simple inversion. This could be valuable information for language teachers who could ideally monitor learners' inter-development of these two grammatical aspects more closely. Another pattern found related to present simple and past simple inversion was that of the 20% of

participants who did not use an inversion, the majority came from multilingual homes. Thus, past tense accuracy and multilingualism might be correlated.

Although the CSC is void of all but one canonical question tag, the corpus participants were found to have possibly used invariant tags with 12.16% of all instances of *yeah* and *yes*. *Yeah* was used by 70% of the participants and *yes* was used by 45% in this way. It was suggested that these invariant tags were used in two ways. First, to draw the listener back into the conversation and secondly to give final verification of an affirmative answer or statement.

Three indications of possible non-standard vernacular language transfer were found. First, it was speculated that the use of 'she don't like' might be associated with the popular Justin Bieber song with the same title. It was further found that 17 of the 19 similar instances with *he/she/it-don't* were similarly found in popular music lyrics. The second instance was the use of the verb *conversate*. It was only used once and was found to be a non-standard word commonly used in African American slang and hip hop. Lastly, the use of *from* in prepositional possessive phases as in 'my brother from another mother' was found to be a possible non-standard vernacular language transfer as the use of *from* in prepositional possessive phases has increased in the past 20 years.

The last pattern found was the amount of differences in the findings of the CSC when compared to the findings of the Swiss English Database. As mentioned in Section 2.2.5, Dröschel (2011, p. 324) lists eight 'potential characteristics of the English spoken in Switzerland'. Each feature will be compared and briefly commented on below. As the Swiss English Database was trilingual, only the German results will be commented on.

- non-native use of articles, as in *the English is a very important language*.

The non-standard use of articles was also found in the CSC but the findings were opposite of those in The Swiss English Database. Dröschel (2011 p. 170) found that 63.5%[32] of all non-native uses were omissions, whereas the CSC found only 30.2% were omissions and 69.8% were the insertion of primarily the definite article. Furthermore, non-standard article use in the CSC was found to be bound to lower proficiency levels. This difference warrants further investigation to determine if alternative uses of articles will become a Swiss English characteristic.

- Non-native plural marking, as in *we need to contact many different person*.

In the Swiss English Database, there is only a limited amount of data from the German speakers therefore it is difficult to make comparisons with the CSC. Findings in the CSC suggested that some aspects of pluralisation might be proficiency bound such as the accuracy of people/person, so the probability of becoming a Swiss English characteristic is perceived as slight.

- reclassification of non-count nouns, as in, *we need more informations.*

There was only one instance of *information(s)* in the CSC and it was plural. None of the other 5 nouns which were examined in the Swiss English Database were found in the CSC. The CSC did have 8 instances of *hairs*; however, it is generally considered a common German L1 interference. Therefore, the CSC findings do not support reclassification of non-count nouns as a potential Swiss English characteristic.

---

[32] As an exception this percentage includes results from all three languages German, French and Italian because single language data was not available.

- non-native use of third person singular S, as in, *he go to London once a month.*

With 64% of all CSC participants using third person singular -*s* in a non-standard way, it might be in the process of becoming a Swiss English characteristic. However, it seems unlikely because although it was used throughout the CSC, the majority of the NSU were used by participants in the lower proficiency levels.

- Adjuncts of backward span, as in *I live here since 20 years.*

The CSC did not have any instances of *for* being used instead of *since* or the reverse. Furthermore, because of the minimum instances of *since* (n=5) in the whole CSC, the data is not comparable.

- Non-native formulation of conditionals, as in *if I would be rich, I would buy it.*

There were initially no non-standard uses of conditional phrases found in the CSC. All instances of *if* and *would* were double checked and again no NSU were found. Thus, comparison is not possible and at this point the probability of becoming a Swiss English characteristic is questionable.

- Insertion of the to infinitive, as in *I am looking forward to see you*.

Of the ten words used to determine overuse of *to* infinitive in the Swiss English Database, only 3 were used in the CSC with *to* or *ing*. *Start* was used equally with *to* and *ing*, *stop* was used only with an *ing* and *like* was used 77% with *to* and 23% with *ing*. These findings are in line with those of the Swiss English Database but they do not indicate a general overuse of infinitive use *to* because of the sole comparable overuse with *like*.

- non-native placement of adverbials, as in *I like very much chocolate*.

Non-standard adverbial placement was found in the CSC and was by used 29.16% of the participants. However, the NSU rate was found to be only 2.83% which suggests that it is not currently becoming a Swiss English characteristic.

The discrepancies and similarities between the Swiss English Database and the CSC might be interpreted as differences between a learner and user corpus. However, most plausible would be that the corpus compilation was not comparable, thus comparison was impaired. As Sinclair (2005) stated 'Only those components of corpora which have been designed to be independently contrastive should be contrasted.' The Swiss English Database included some written elements and the German part was roughly half the size of the CSC. Nevertheless, the Swiss English Database was the only other Swiss spoken corpus to date. It is expected that the comprehensive documentation of the investigation of the CSC and its results will offer new insights and comparable data for further studies.

5.2.5 Words with high NSU rates

Throughout the analysis, it became apparent that certain words were used in non-standard ways more often than others. Furthermore, NSU rates were calculated for most of them. This information can be very beneficial for curriculum and textbook developers as well as practitioners because when the words which prove difficult to master are known, they can be subsequently highlighted in lesson plans. Furthermore, additional attention can be given in the use of oral correction tactics to provide students with the most useful feedback to increase accuracy. With increased accuracy, better communication can be assured and students can benefit from higher marks on exams which could ultimately result in more successful language learning. The following lists highlight words with high NSU rates.

Third person singular

The following nine words were found to be used with a non-standard rate of over 5%. The rate was calculated by dividing the number of NSU-1 with all instances of that word used in the infinitive and third person singular.

ask	draw	leave	live	paint
present (verb only)	show	try		

Prepositions

The following ten prepositions were used in a non-standard way, with the three with a nonstandard rate over 5 % marked in bold.

about	**at**	by	for	**from**
in	of	**on**	to	with

At was associated with the highest non-standard rate of 12.04% and was found to be most problematic in the following phrases where it was used in place of *on*.

at (on) holiday
at (on) the seventh day
at (on) a normal day
at (on) the woman
at (on) the cruise ship
at (on) the chair

Past simple

The following fifteen words were found to be used with a non-standard rate of over 5%. The rate was calculated by dividing the number of NSU-10 with all instances of that word used in the present and past tense.

ask	build	buy	change	draw
drive	eat	fly	go	make
paint	rent	sit	start	visit

Plural -s

The following four words were found to have an NSU rate of over 5% when calculated with their singular and plural use. It is noteworthy that all are irregular.

hair	foot	child	person

Lexical choice

A list of the 30 most frequent verbs found in the CSC resulted in 20 verbs with a nonstandard rate of over 5%. The calculations were made taking into consideration all forms of the verb and all NSU associations. These words were used 470 times with a NSU which accounts for 17.78% of all NSU found in the corpus.

ask	come	do	draw	drive
eat	get	go	look	make
paint	put	say	see	show
take	talk	tell	try	want

The above-mentioned words are basic words that are taught at elementary levels, yet they were the most problematic in terms of accuracy for the participants in the CSC corpus.

The previous summary has highlighted the results of the analysis of the CSC. The following chapter will address research question three which examines the acceptability of the language found within the corpus, after which conclusions will be drawn on the work presented in this thesis in Chapter 7.

Chapter 6 PERCEPTIONS OF ACCEPTANCE OF NON-STANDARD ENGLISH

6.1   Introduction

To determine the extent to which the Swiss accept non-standard features and answer the third research question, a survey was conducted to evaluate the acceptance of the non-standard language produced by the participants in this study.

RQ 3: How is the acceptance of commonly used non-standard features perceived by the emerging Central Swiss workforce?

The Swiss National Science Foundation project 'Language Contact and Focussing: The Linguistics of English in Switzerland' from 1999-2005 concluded that although potential characteristics of Swiss English were found, a Swiss English variety was not evolving. Instead, Swiss English was described as being 'a conglomerate of learner varieties of English which are heavily conditioned by processes of L1 transfer and simplification' (Dröschel, 2011, p. 331). The present study has investigated Swiss learner language with the aim of documenting its current state. If the non-standard features produced are accepted, this could be an indication that innovations are occurring. Bamgbose describes an innovation as being 'seen as an acceptable variant, while an error is simply a mistake or uneducated usage' (Bamgbose, 1998, p. 2) and further, the acceptability factor is 'the ultimate test of admission of an innovation' (ibid. p. 4). Thus, data which reveals current opinions about acceptability of non-standard usage is valuable.

Two previous surveys of acceptance are pertinent to this study. First Murray (2003) conducted a survey of Swiss English teachers' attitudes and acceptability judgements of Euro-English. She

asked native and non-native teachers of English to decide if eleven sentences containing typical Euro-English features were either acceptable or unacceptable. Acceptability was generally divided between sentences that broke 'explicit grammatical rules taught in Standard English teaching materials' and those that did not (Murray, 2003, p. 157), with those that broke the rules being rated with a higher unacceptability. Furthermore, discrepancies between native and non-native teachers were found with native teachers being more accepting. Regional differences were likewise noted concerning French or German specific loan words with the acceptability of regionally used loan words being higher.

The second study was conducted by Mollin (2006). Her study expanded on Murray's by using five of Murray's sentences and eight other fictitious sentences to represent Euro-English. The inclusion of seven correct sentences and asking the respondents to correct any sentences they deemed unacceptable was used as a refined means to predict acceptance of Euro-English more precisely when combined with the respondents' metadata and information on attitudes towards English (Mollin, 2006, p. 165). Although Mollin (2006) found some Euro-English characteristics with high acceptability rates, she concluded that her survey results were consistent with Murray's and 'one cannot speak of an institutionalization of Euro-English' (Ibid., p. 190).

Both Murray (2003) and Mollin (2006) investigated the acceptance of Euro-English by conducting surveys with elite English speakers, Murray with Swiss English teachers and Mollin with university faculty across Europe. The rationale for this approach being that acceptance of Euro-English would begin at the top with those that use English with other Europeans most often. In Central Switzerland, a different viewpoint can be surmised. With the influx of international companies with English as the company language to the daily interactions with

tourists in all public areas, English is spoken by all classes of people with varying degrees of competence. I hypothesis that it is the masses who will create change with 46% of all residents in the Swiss German speaking part of Switzerland above the age of 15 using English regularly[33]. Similar to language changes being documented from the bottom up, for example, teenagers (Tagliamonte, 2006) or Black English Vernacular (Labov, 1972b) in North America, this group's interactions are being heard and perhaps imitated much more broadly than the elite who might prefer conformity. Thus, the decision to survey the emerging workforce was made to discover their acceptance of non-standard English used by their equals.

6.2   Methods

After preliminary analysis of the CSC corpus was conducted, ten sentences were chosen which depicted a variety of common non-standard usages found in the corpus. The decision to use sentences from the corpus was based on the belief that the results would verify the acceptance of those non-standard usages within the context of real spoken language more accurately. These ten sentences, see Table 6.1 below, became the basis for the online survey which was conducted from 12.11.2019-19.06.2020. The survey link was sent to selected teachers of students studying for a Matura and those in an apprenticeship. Unfortunately, parallel to the collection of data, the COVID-19 virus restrictions in Switzerland affected the access to some educational institutions and thus access to Matura students was limited. Therefore, the survey was opened up to a small group of students at a higher vocational institute and university.

---

[33]https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/sprachenreligionen/sprachen.assetdetail.15384
164.html accessed 04.06.2021

Table 6.1 Survey sentences with non-standard usages

1. I think she **want** (wants) to say it in another way.

2. In Switzerland, **the** (~~the~~) most people can speak English.

3. Five minutes later the wind came and blew faster and then we **go** (went) into a museum.

4. I need English when I'm **in** (on) holiday.

5. In the underground trains in London, there are too many **peoples** (people) for me.

6. It is important to use the right **grammatic** (grammar).

7. My friend is in **gymnasium** (high school) and I can **speak with her English** (speak English with her).

8. He painted another one and then finally showed (**it**) to the others.

9. It was the first time I went to another country alone and it was (**an**) interesting time.

10. I don't know how many people **speaking** (speak) the language.

The survey introduction informed respondents that the sentences were taken from interviews and had a non-standard aspect with my interest being how acceptable they thought the sentences were when they are spoken in everyday situations, see Appendix 5 for full survey. The respondents were informed about the non-standard aspects to emphasise the element of choice and control that they have over their opinions of acceptability.

A 4-point Likert scale was used to force the respondents to decide how acceptable each sentence was with three options of acceptability and one not acceptable: very acceptable, acceptable, somewhat acceptable, not acceptable. The reason for using gradient levels of acceptability and one level of unacceptability was the belief that language use can be acceptable up to a certain point and afterwards no longer.

The acceptability rate was calculated by using the Smart Survey[34] approach of assessing a satisfaction rate. This approach was used because of the similarities to satisfaction and acceptability. First, values were assigned to the four points of the Likert scale: very acceptable 3 points, acceptable 2 points, somewhat acceptable 1 point, not acceptable 0 points. Next, the *maximum possible value* was calculated by multiplying the number of respondents and the highest possible value, 205 x 3 = 615. Then, the *actual total value* was calculated by multiplying the number of respondents for each selected answer choice by the assigned value from 0-3 and adding them together. Lastly, the *actual total value* was divided by the *maximum possible value* and multiplied by 100 to get the acceptance rate. By using these calculations, we have a better understanding of degree of acceptability than if only the percentages of very acceptable and acceptable were tallied.

A total of 205 respondents took part in the survey. The two variables age and current studies were asked. It was hypothesized that there would be differences in acceptance depending on age group and educational level with the hypothesis that the older and higher educated respondents would find the sentences less acceptable. Starting with education; Table 6.2 below illustrates the stated current studies of the respondents.

Table 6.2 Current studies of survey respondents

Type of studies	Number of respondents	Percentage of respondents
Group 1- Apprentices	133	64.87%
Group 2- Matura students	4	1.95%
Group 3- Higher vocational education	47	22.92%
Group 4- University	19	9.26%
Teachers	2	0.97%

---

[34] https://help.smartsurvey.co.uk/article/satisfaction-rate-calculations?s=Charts-and-Data-Tables

The respondents are somewhat representative of the educational levels of the Central Swiss population as mentioned in Table 3.13. The exception being the underrepresentation of Matura respondents. The survey construct allows for the comparison of respondents undergoing their first educational level, apprentices and Matura students (n=137), and those studying further education, higher vocational education and university (n=66). Thus, variation of acceptance can be observed. In addition, all respondents over the age of 25 could not have had early English in public schools if they attended primary school in Central Switzerland. Additionally, two teachers answered the survey and their answers were calculated in the acceptance rate, but not in the educational variable.

Moving on to the age variable, the ages of the respondents ranged from 15 to 50 years of age. These were divided into three groups; 15-20 year olds to represent the age group of the participants who comprised the CSC, 21-25 year olds to represent respondents who are studying in further education and possibly had early English instruction and lastly 26-33 year olds who are studying further education and most probably did not have early English. The two teachers aged 40 and 50, who completed the survey were not included here and will be mentioned separately. Table 6.3 below gives an overview of the age groups.

Table 6.3 Survey respondent age groups

Age group	Age	Number of respondents
	15	7
	16	6
Group 1	17	34
15-20 year olds	18	39
	19	14
	20	12
Total	**112**	
	21	6
	22	13
Group 2	23	19
21-25 year olds	24	7
	25	17
Total	**62**	
	26	6
	27	6
Group 3	28	4
26-33 year olds	29	2
	33	1
Total	**19**	

6.3   Survey results

Next, the survey question results will be presented, and relevant findings highlighted. The tables corresponding to each survey question were created to give a concise view of the data. Each table begins with the overall results before highlighting the variables education and age. The four groups within the education section of each table stand for:

Group 1 - Apprentices

Group 2 - Matura students

Group 3 - Higher vocational studies students

Group 4 - University students

The three groups within the age section of each table represent the following age ranges:

Group 1- 15 to 20-year olds

Group 2- 21 to 25-year olds

Group 3- 26 to 33-year olds

6.3.1　Question 1 - I think she **want** to say it in another way

The first question is an example of NSU-1 which is the non-standard use of the third person singular -*s*. There were 170 instances of NSU-1 found in the corpus and in Section 5.1.1 it was concluded that the corpus participants used some verbs quite accurately with a NSU-1 frequency of under 1 percent when compared with all uses of that verb in the infinitive and third person singular. These verbs included high frequency verbs such as *speak*, *think* and *like* as well as the verb *to be* and *do*. However, verbs needed to describe specific actions were less accurate. The verb *want* in the survey question below appeared in the corpus as infinitive or 3rd person singular 258 times with 12 NSU-1 which is a non-standard rate of 4.65%.

As can be seen in Table 6.4 below, the sentence was perceived as being very acceptable or acceptable to slightly more than half of the survey respondents at 50.7% and only 11.7% viewed it as unacceptable. This is substantial and quite surprising because it could indicate that zero third person singular has the possibility of becoming a salient feature of Swiss English in the future. At present, it occurs to be unlikely because of its minimal use in the CSC corpus of 0.17 times per 1000 words which is in line with the corpus results Dröschel (2011 p. 213) reported from her data. Mollin (2006, p 187) on the contrary, reported an acceptability rate of only 5%, with lower proficiency rates accounting for acceptance of the sentence ' Do you know where she live' (ibid.). It is possible that acceptance has changed in the past fifteen years since Mollin's survey or there is a vast difference in the perception of acceptability by the Swiss who have been taught with a productive focused curriculum since 2005/8.

When the results are further investigated there is a pattern of less acceptance from the groups with higher education and age, with a higher level of education having the greatest impact, although 26.3% of university students still find the sentence acceptable in an everyday spoken context. Presently 43% of the Swiss population between 25-64 years of age have a tertiary education and it is estimated that this will increase to 50% by 2037 (Demografische Entwicklung und Auswirkungen auf den gesamten Bildungsbereich 2019 p. 32). Therefore, if the Swiss population continues to seek higher education there is a chance that acceptance of zero third person singular could be slightly curtailed. Much will depend on the development of English at the global level and who the Swiss will be communicating with in English because in the highest age group only 15.78% find it not acceptable.

Table 6.4 Survey Question 1, I think she **want** (wants) to say it in another way.

Question 1				
			Response %	Response no.
1. Very acceptable			11.2%	23
2. Acceptable			39.5%	81
3. Somewhat acceptable			37.6%	77
4. Not acceptable			11.7%	24
Acceptance rate	**50.08%**			
Mean	2.5			
Standard deviation	0.84			
Standard error	0.06			
Question 1 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	12.03%	50%	6.38%	10.52%
2. Acceptable	43.60%	50%	38.29%	15.78%
3. Somewhat acceptable	31.57%	-	46.80%	57.89%
4. Not acceptable	12.78%	-	8.51%	15.78%
Question 1 Age variable				
Acceptability	Group 1	Group 2	Group 3	
1. Very acceptable	16.07%	6.45%	5.26%	
2. Acceptable	45.53%	37.09%	36.84%	
3. Somewhat acceptable	35.71%	43.54%	52.63%	
4. Not acceptable	11.60%	12.90%	15.78%	

Both teachers found Question 1 somewhat acceptable. This is perhaps due to the comprehensibility of the spoken message. To sum up, zero third person singular was largely accepted by the respondents particularly by the 15 to 20 year olds with non-acceptability generally increasing with education and age.

6.3.2 Question 2 - In Switzerland, **the** (the) most people can speak English

The second question deals with the insertion of the definite article *the* and is an example of NSU-4. Of the 97 instances of redundant definite article use in the corpus, 10 of them were with *the most*. Besides this non-standard use of *most* like a superlative, definite articles were inserted with places, institutions, collective nouns, times and numbers, the word *life*, languages and miscellaneous items. It was concluded that many NSU-4 were probably due to L1 interference. The word *most* was used in the corpus 56 times with 10 NSU-4 which is a non-standard rate of 17.85%. This high non-standard rate could be correlated with the particular sentence because it has two hurdles for learners. First, the complicated article rules and second German interference. In German, the definite article is used with *most*, whereas in English it is not.

German   In der Schweiz können *die meisten* Menschen Englisch sprechen.

English   In Switzerland *most* people can speak English

Table 6.5 Survey Question 2, In Switzerland, the (the) most people can speak English.

Question 2				
			Response %	Response no.
1. Very acceptable			22.0%	45
2. Acceptable			37.1%	76
3. Somewhat acceptable			26.3%	54
4. Not acceptable			14.6%	30
Acceptance	55.45%			
Mean	2.34			
Standard deviation	0.98			
Standard error	0.07			

Question 2 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	26.31%	25%	12.76%	15.78%
2. Acceptable	42.85%	25%	31.91%	15.78%
3. Somewhat acceptable	21.05%	50%	36.17%	36.84%
4. Not acceptable	9.77%		19.14%	31.57%
Question 2 Age variable				
Acceptability	Group 1		Group 2	Group 3
1. Very acceptable	28.57%		16.12%	15.78%
2. Acceptable	49.10%		27.41%	21.05%
3. Somewhat acceptable	24.10%		33.87%	31.57%
4. Not acceptable	7.14%		22.58%	42.10%

In Table 6.5 above, we see that the acceptability rate is 55.45% with only 30% of the respondents judging the sentence not acceptable. Here it is interesting to observe the steep increase of non-acceptability as education and age increases, although the higher vocational students still find the sentence twice as acceptable as the university students. In addition, both teachers found the sentence not acceptable. This could be because it breaks the grammar rules taught at pre-intermediate level as Murray (2006) suggested in her study.

6.3.3   Question 3 - Five minutes later the wind came and blew faster and then we **go** (went) into a museum

Question 3 is an example of a NSU-10 simple present and past inversion. There are a total of 454 NSU-10 in the corpus with over 80% of the participants using at least one. There were 68 instances of *go/goes* used instead of went. This equals a NSU-10 rate of 12.48% for *go/went*. Although there was a high rate of NSU-10 *go/went*, it was distinctly recognised as not or somewhat acceptable by the vast majority of the respondents and has the lowest acceptability rate of all 10 sentences. Only a small number of mainly young apprentices and a few 26 to 33-year-olds found it acceptable. Most striking is the reduced deviation of differences within the age groups for somewhat and not acceptable. Table 6.6 below illustrates these.

Table 6.6 Survey Question 3, Five minutes later the wind came and blew faster and then we go (went) into a museum

Question 3				
			Response %	Response no.
1. Very acceptable	■		4.4%	9
2. Acceptable	▬		14.6%	30
3. Somewhat acceptable	▬		41.0%	84
4. Not acceptable	▬		40.0%	82
Acceptance rate	**27.8%**			
Mean	3.17			
Standard deviation	0.83			
Standard error	0.06			
Question 3 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	6.76%	-	-	-
2. Acceptable	17.29%	-	10.63%	10.52%
3. Somewhat acceptable	39.09%	25%	44.68%	52.63%
4. Not acceptable	36.84%	75%	44.68%	36.84%
Question 3 Age variable				
Acceptability	Group 1		Group 2	Group 3
1. Very acceptable	7.14%		-	5.26%
2. Acceptable	15.17%		12.90%	26.31%
3. Somewhat acceptable	47.32%		40.32%	31.57%
4. Not acceptable	39.28%		46.77%	47.36%

Both teachers found the sentence not acceptable, again it is presumed because of the grammatical aspect.

6.3.4 Question 4 - I need English when I'm **in** (on) holiday

Question 4 contains a typical Swiss example of NSU-13 substitution of prepositions. Out of a total of 224 NSU-13, 14 were found with holiday and a non-standard preposition. Eight of those are similar to Question 4 with the substitution of *in* instead of *on* holiday. There are only 22 instances of holiday with a preposition in the entire CSC which makes an NSU rate of 63.63%, one of the highest NSU rates in the study. When we look at the corresponding Table 6.7 below, we see that there is an acceptability rate of 42.6% which is the fourth highest among the 10 sentences. However, there appears to be uncertainty with the majority rating the sentence somewhat acceptable in all education and age groups. The uncertainty was also projected by

the teachers with one proclaiming the sentence acceptable and the other somewhat acceptable. The apprentices and youngest age group showed the highest level of acceptability.

Table 6.7 Survey Question 4, I need English when I'm in (on) holiday

Question 4				
			Response %	Response no.
1. Very acceptable			13.7%	28
2. Acceptable			25.4%	52
3. Somewhat acceptable			36.1%	74
4. Not acceptable			24.9%	51
Acceptance rate	**42.6%**			
Mean	2.72			
Standard deviation	0.99			
Standard error	0.07			
Question 4 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	15.70%	-	12.76%	5.26%
2. Acceptable	24.81%	50%	25.53%	21.05%
3. Somewhat acceptable	33.83%	50%	36.17%	47.36%
4. Not acceptable	25.56%	-	25.53%	26.31%
Question 4 Age variable				
Acceptability	Group 1	Group 2	Group 3	
1. Very acceptable	16.07%	11.29%	15.78%	
2. Acceptable	49.10%	29.03%	15.78%	
3. Somewhat acceptable	40.17%	30.64%	52.63%	
4. Not acceptable	25%	29.03%	26.31%	

Prepositions remain difficult for many learners to decipher especially as in Question 4 because of L1 interference. In order to be able to make comparisons, questions in acceptability surveys must be carefully chosen. Murray (2006) also used a sentence with a prepositional substitution. However, it was deemed the second least acceptable sentence with a non-acceptable rate of 77%. Her use of 'I am going **by** (to) the dentist tomorrow' (ibid. , p. 158) does not have an obvious German L1 influence factor nor is it known if it is typically used in Switzerland or fabricated. Thus, it is challenging to come to conclusions on the differences between the English teachers' in her study and learners' acceptability of non-standard prepositional use in this study.

6.3.5   Question 5 - In the underground trains in London, there are too many **peoples**

(people) for me

Question 5 is an example of NSU-15 inserted plural *s*. Of the 160 NSU-15, 55 were used with

person which has the irregular plural *people*. The non-standard use of *peoples* was found in the

corpus 39 times with an NSU rate of 10.61%.

The acceptability rate for Question 5 of 41.3% is fairly high with apprentices and the youngest

age group having the highest rate of very acceptable. However, the higher vocational education

group and the 21-25 year olds were highest when the rating of acceptable is observed in Table

6.8 below.

Table 6.8 Survey Question 5

Question 5			Response %	Response no.
1. Very acceptable			11.2%	23
2. Acceptable			25.9%	53
3. Somewhat acceptable			38.5%	79
4. Not acceptable			24.4%	50
Acceptance rate	**41.3%**			
Mean	2.76			
Standard deviation	0.95			
Standard error	0.07			
Question 5 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	14.28%	-	6.30%	5.26%
2. Acceptable	23.30%	25%	34.04%	26.31%
3. Somewhat acceptable	41.35%	51%	27.65%	42.10%
4. Not acceptable	21.05%	25%	31.91%	26.31%
Question 5 Age variable				
Acceptability	Group 1	Group 2	Group 3	
1. Very acceptable	15.17%	6.45%	10.52%	
2. Acceptable	26.78%	30.64%	21.05%	
3. Somewhat acceptable	44.64%	33.87%	42.10%	
4. Not acceptable	22.32%	29.03%	36.84%	

As in Question 4, it appears that there is uncertainty in the non-acceptability because of the high percentage of somewhat acceptable. One teacher found it somewhat acceptable and the other not acceptable. Further, the corpus analysis revealed that the majority of participants who used *peoples* had a B1 proficiency level as do 41.66% of the corpus. This could be a factor in the level of uncertainty.

In related research, Mollin (2006 p. 181) reported that the acceptability of the plural uncountable nouns *informations* and *equipments* was correlated to the proficiency level of the respondents with the near-native English respondents with the highest rate of unacceptability. Interestingly, she found *informations* was accepted by 23.50% and *equipments* over twice as high at 53.19%. She concluded that difference was probably proficiency bound and did not reflect the acceptance of Euro-English norms (ibid. p. 182).

6.3.6 Question 6 - It is important to use the right **grammatic** (grammar)

Question 6 is an example of an NSU-16 of lexical choice. Of the 807 NSU-16, 144 were directly associated with L1 German interference and 60 of those were used with the word *grammar* with an NSU rate of 50.42%. In German, the word *grammar* translates to *Grammatik*. As we see in Table 6.9, only 9.3% felt that using the German word/pronunciation was considered not acceptable and the vast majority found it either very acceptable or acceptable. This number is also reflected in the acceptance rate of 62.6%, the highest in the survey. Within the educational groups, university students showed the highest rate of not acceptable, as well as the 26-33 year olds. One of the teachers found the sentence acceptable and the other somewhat acceptable. This is interesting because grammatic is not an English word. There is the possibility that the respondents read the sentences quickly and recognised the word which belongs to their repertoire without noticing it was the wrong language.

Table 6.9 Survey Question 6

Question 6			Response %	Response no.
1. Very acceptable	▬ (green)		28.3%	58
2. Acceptable	▬ (blue)		40.5%	83
3. Somewhat acceptable	▬ (yellow)		22.0%	45
4. Not acceptable	▬ (red)		9.3%	19
Acceptance rate	**62.6%**			
Mean	2.12			
Standard deviation	0.93			
Standard error	0.06			
Question 6 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	35.58%	-	23.40%	5.26%
2. Acceptable	38.34%	75%	44.68%	36.84%
3. Somewhat acceptable	21.05%	25%	8.51%	36.84%
4. Not acceptable	6.01%	-	14.89%	21.05%
Question 6 Age variable				
Acceptability	Group 1	Group 2	Group 3	
1. Very acceptable	15.17%	20.96%	21.05%	
2. Acceptable	41.96%	45.16%	42.10%	
3. Somewhat acceptable	25%	20.96%	21.05%	
4. Not acceptable	5.35%	12.90%	26.31%	

Murray (2003) found similar results with the loan words *handy* and *fitness*. A sentence with *handy*, which is commonly used in the German speaking part of Switzerland to mean cell phone, was accepted by 57.1% of the German speaking respondents. Further, the word *fitness* is commonly used to represent *fitness centre* in the French speaking part of Switzerland, was accepted by 66.7% of the French-speaking respondents (ibid., p. 159). The acceptability of these words was only between 33-42% in the opposing language regions. Thus, demonstrating how elite and non-elite users of English accept non-standard language that is used and heard in their daily lives.

6.3.7 Question 7 - My friend is in **gymnasium** (high school) and I can **speak with her English** (speak English with her)

Question 7 is an example of a sentence with two non-standard usages. First *gymnasium* is the German equivalent to *high school* and is classified as a NSU-16, lexical choice, and *speak with her English* is an example of a NSU-20, word order. NSU-16 and 20 account for 40% of all NSU in the corpus and it is not uncommon for more than one NSU to occur in a sentence, especially in the lower proficiency levels. *Gymnasium* was used only four times in the corpus with a 100% NSU rate and word order NSU-20 with subject-verb-object occurred 19 times.

The low frequency of these NSU in the corpus could be an indication of their unacceptability. Question 7 has the second lowest overall acceptability rate at 34.96% and the second lowest rate of very acceptable at 6.3%. Acceptability was highest with the apprentices and lowest age group with none of the university students choosing very acceptable or acceptable. We also see in Table 6.10 below non-acceptability increasing with higher educational levels and age. Interestingly, both teachers rated the sentence somewhat acceptable, which is quite surprising. We know that both teachers were accepting of common L1 interference in Question 6, but the acceptance of non-standard word order was not expected. One possible reason could be the speed in which the survey was completed. On average respondents spend five minutes to read the simple instructions, rate the ten sentences and state their studies and age.

Table 6.10 Survey Question 7

Question 7			
		Response %	Response no.
1. Very acceptable	▬	6.3%	13
2. Acceptable	▬▬	24.9%	51
3. Somewhat acceptable	▬▬▬	36.1%	74
4. Not acceptable	▬▬▬	32.7%	67
Acceptance rate	**34.96%**		
Mean	2.95		
Standard deviation	0.91		
Standard error	0.06		
Question 7 Education variable			
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	9.02%	-	2.12%	-
2. Acceptable	32.33%	25%	14.89%	-
3. Somewhat acceptable	33.08%	50%	36.17%	47.36%
4. Not acceptable	25.56%	25%	46.80%	52.63%
Question 7 Age variable			
Acceptability	Group 1	Group 2	Group 3
1. Very acceptable	9.82%	1.61%	5.26%
2. Acceptable	34.82%	14.51%	15.78%
3. Somewhat acceptable	39.28%	40.32%	23.31%
4. Not acceptable	25%	43.54%	63.15%

6.3.8   Question 8 - He painted another one and then finally showed (**it**) to the others

Question 8 is another example of NSU-20, this time with an omitted object. There were 12 instances of an omitted object and 4 instances with **it** being omitted in the corpus. One can estimate the NSU rate of missing objects by first calculating the number of complete sentences. Number of sentences were estimated by dividing the total words in the corpus by an estimated sentence length of 15 words. This resulted in an estimated missing object NSU rate of 0.18%.

This question had low acceptability with 37.89%. Interestingly, the apprentices who until now had for the most part the highest acceptability rates of all questions had the highest not acceptable rate. The university students, on the other hand, had the highest acceptability rate and the lowest not acceptable rate. Furthermore, the oldest group of respondents still recorded the highest rate of not acceptable as is seen in Table 6.11. Lastly, somewhat acceptable between

36.84 -50% indicates a high level of uncertainty about the acceptability. Both teachers found the omission of the object in Question 8 not acceptable.

Table 6.11 Survey Question 8

Question 8			Response %	Response no.
1. Very acceptable	▬		6.8%	14
2. Acceptable	▬▬		26.3%	54
3. Somewhat acceptable	▬▬▬		40.5%	83
4. Not acceptable	▬▬		26.3%	54
Acceptance rate	**37.89%**			
Mean	2.86			
Standard deviation	0.88			
Standard error	0.06			
Question 8 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	7.51%	-	4.25%	10.52%
2. Acceptable	28.57%	25%	17.02%	36.84%
3. Somewhat acceptable	36.84%	50%	53.19%	36.84%
4. Not acceptable	27.06%	25%	25.53%	15.78%
Question 8 Age variable				
Acceptability	Group 1	Group 2		Group 3
1. Very acceptable	9.82%	4.83%		-
2. Acceptable	31.25%	22.58%		26.31%
3. Somewhat acceptable	39.28%	50%		42.10%
4. Not acceptable	28.57%	22.5%		42.10%

6.3.9 Question 9 - It was the first time I went to another country alone and it was (**an**) interesting time

Question 9 is an example of NSU-3, omitting an indefinite article. In total 34 indefinite articles were found to be omitted in the corpus which is an NSU rate of 1.53%. As we see in Table 6.12 below, Question 9 has a low rate of very acceptable, but an overall rate of 40% acceptability. Within the educational groups, the university group's rating of acceptable was higher than the apprentices which only occurred in two other sentences, Question 5 with *peoples* and Question 8 with the omission of *it*. The majority of the respondents rated the sentence somewhat acceptable including both teachers. Hence, the omission of words seems to have been observed,

but with the essence of the sentence comprehended, the respondents appear to have been somewhat accepting. Murray (2003, p. 158) reported a similar acceptability rate of 38% with the sentence 'You should see doctor'.

Table 6.12 Survey Question 9

Question 9				
			Response %	Response no.
1. Very acceptable			6.3%	13
2. Acceptable			31.7%	65
3. Somewhat acceptable			37.6%	77
4. Not acceptable			24.4%	50
Acceptance rate	**40%**			
Mean	2.8			
Standard deviation	0.88			
Standard error	0.06			
Question 9 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	9.02%	-	-	5.26%
2. Acceptable	29.32%	25%	38.29%	36.84%
3. Somewhat acceptable	42.10%	50%	25.53%	26.31%
4. Not acceptable	19.54%	25%	36.17%	31.57%
Question 9 Age variable				
Acceptability	Group 1	Group 2	Group 3	
1. Very acceptable	9.82%	1.61%	10.52%	
2. Acceptable	31.25%	45.16%	10.52%	
3. Somewhat acceptable	47.32%	24.19%	47.36%	
4. Not acceptable	21.42%	29.03%	42.10%	

6.3.10 Question 10 - I don't know how many people **speaking** (speak) the language

Question 10 is an example of an NSU -18, redundant use of continuous forms. The question contains the verb *speak* in the continuous form without the verb *to be*. Of the 69 NSU-18, 40 examples of the redundant use of continuous without the verb *to be* were found in the corpus. As mentioned in Section 5.1.19, High German does not have a continuous form with the verb to be, however, Swiss German does. This can be confusing for lower proficiency levels but is perhaps more recognisable as non-standard because of the attention usually given to the continuous forms in English classrooms with German L1 students.

At 31.1% it had the third highest rating of overall not acceptable of all the sentences. Further, Table 6.13 shows a general acceptability rate at 36.26%. As in nine of the ten questions, the very acceptable rate is highest in the apprentice group. In the age variable, there is also a gradient increase in the not acceptable rate with the increase in age. This is found in nine out of the ten questions. The teachers agreed with the majority of the respondents that the sentence was not acceptable.

Table 6.13 Survey Question 10

Question 10				
			Response %	Response no.
1. Very acceptable			8.8%	18
2. Acceptable			25.4%	52
3. Somewhat acceptable			31.7%	65
4. Not acceptable			31.1%	70
Acceptance rate	**36.26%**			
Mean	2.91			
Standard deviation	0.97			
Standard error	0.07			
Question 10 Education variable				
Acceptability	Group 1	Group 2	Group 3	Group 4
1. Very acceptable	10.52%	-	4.25%	5.26%
2. Acceptable	26.31%	25%	27.65%	15.78%
3. Somewhat acceptable	31.57%	75%	27.65%	36.84%
4. Not acceptable	31.57%	-	38.29%	42.10%
Question 10 Age variable				
Acceptability	Group 1	Group 2	Group 3	
1. Very acceptable	11.60%	4.83%	10.52%	
2. Acceptable	29.46%	30.64%	-	
3. Somewhat acceptable	33.92%	29.03%	47.36%	
4. Not acceptable	33.92%	35.48%	52.63%	

6.4   Summary of acceptance and conclusion

This survey has demonstrated how the spoken output of Swiss English speakers is accepted by their peers and students of further education. In Table 6.14 below, the sentences are listed in order of acceptability. We see that three of the most acceptable sentences, 6, 2 and 4 have aspects of German L1 interference. At the same time, the NSU in these sentences have a high

NSU rate which means that they are probably being used regularly. As in Murray's study, with the exception of sentence one, sentences which break grammatical rules were less accepted.

Table 6.14 Questions in order of acceptability rate

Questions in order of acceptability rate	Acceptability rate	NSU rate
6. It is important to use the right **grammatic** (grammar).	62.60%	50.42%
2. In Switzerland, **the** (~~the~~) most people can speak English.	55.45%	17.85%
1. I think she **want** (wants) to say it in another way.	50.08%	4.65%
4. I need English when I'm **in** (on) holiday.	42.60%	63.63%
5. In the underground trains in London, there are too many **peoples** (people) for me.	41.30%	10.61%
9. It was the first time I went to another country alone and it was (**an**) interesting time.	40%	1.53%
8. He painted another one and then finally showed (**it**) to the others.	37.89%	0.18%
10. I don't know how many people **speaking** (speak) the language.	36.26%	*
7. My friend is in **gymnasium** (high school) and I can **speak with her English** (speak English with her).	34.96%	0.28%
3. Five minutes later the wind came and blew faster and then we **go** (went) into a museum.	27.80%	12.48%

* NSU rate not calculated

The 50% acceptability rate of sentence one is quite surprising because the sentence clearly breaks a grammatical rule. The use of zero third person has been recognised as being 'particularly widespread' in ELF data (Cogo and Dewey, 2012, p. 49). Furthermore, third person *-s* irregularities have been documented in native English varieties (Britain, 2010, p. 39) as well as non-native varieties (Kortmann, 2010, p. 409) and could indicate upcoming change.

Another factor to be considered is that Swiss German speakers are known to be very accommodating and accepting when speaking their dialect to others outside of their immediate language region. Sometimes, a Swiss German dialect just 10 kilometres away will have noticeable differences due to a multitude of rationales such as municipality borders to geographic borders of mountains and valleys. Thus, language variation acceptance is common not only in spoken, but also in written because Swiss German has no standard for dialectal orthography. This means that each speaker writes how they believe the words should sound (Siebenhaar, 2006, p. 483). In the age of text messaging, accommodation is standard and

tolerance high. This, along with the widespread acceptance of loan words from other languages (Bon, 1948, p. 232), implies a general high tolerance for language variation.

In conclusion, this online survey of Swiss English acceptance has provided insight into the language which is deemed acceptable. First, accepting language peculiarities within a Swiss context is ongoing and their acceptance is not uncommon (Rickenmann, 2005, p. 123). However, the acceptability of non-standard language can be an indication of innovation and language change. Therefore, the acceptability of zero third person should be monitored and investigated further to determine if Swiss English speakers will join fellow Europeans in accepting emerging features of English as Cogo and Dewey (2012) suggest or are in fact still learners who will change their attitude once a higher level of proficiency is achieved.

Chapter 7 **CONCLUSION**

7.1   Contributions made by this thesis

This thesis has contributed to the sociolinguistic field of study by investigating the spoken English of the emerging Swiss workforce by means of a learner corpus and subsequent online survey of non-standard English acceptance. The gathering of extensive metadata and proficiency rating of each participant facilitated an in-depth analysis of the effect of demographic, identity-related and educational variables on the use of non-standard spoken output. The detailed analysis of the frequency of non-standard usages resulted in greater understanding of the effects of Swiss educational policy changes of early English instruction, evidence of proficiency bound learner accuracy and wordlists related to certain types of non-standard usages made by the participants in the study.

The knowledge gained can be used threefold. First, as an alternative example of learner corpus compilation for plurilingual and mixed proficiency level corpora. Secondly, the results confirm the achievements made in the Swiss school system thus far and highlights suggested areas of improvement or development. Finally, the gained understanding of the type and number of non-standard English used by the participants of this study can guide Swiss practitioners when choosing the most effective teaching material and methods to use when encouraging accurate spoken English output.

7.2   Research questions reviewed

RQ 1: What is the relationship between the use of non-standard English features and speakers' demographic, identity-related and educational variables?

Research Question 1 was answered in Chapter 4, Section 4.2. Beginning with the demographic variables explored; gender, age, nationality, native language, parents' mother tongue and languages spoken at home, both gender and age were not found to have a noteworthy effect on non-standard English usage. The relationship between nationality and languages, however, was found to be of importance with the more exposure to languages resulting in higher spoken accuracy of English. This coincides with past research indicating the advantage that children with multilingual backgrounds have when learning a third language in Swiss classrooms (Brohy, 2001 and Haenni-Hoti et al., 2011).

The participants in this study were mildly diverse with 17.7% non-Swiss and thirteen different native languages. It would appear to be very diverse when compared with monolingual European countries. However, depending on the municipality, Swiss classrooms can range from 0 to 79% foreign students (Haag, 2016). Therefore, the participants in this study could represent other municipalities with a similar range of multicultural backgrounds.

Identity-related variables included familiarity with the interviewer as well as participants' views on the importance of grammatical correctness versus understanding, who they will speak English with in the future, feelings when speaking English and self-proclamation of being a learner or user. Familiarity with the interviewer did not appear to affect the accuracy of spoken output. This can be an indication that the interviews were conducted in an impartial manner. The general outcome of the analysis of identity-related variables is that participants who believed in the importance of grammatical correctness and exhibited expectations of speaking with native English speakers produced more non-standard usages than their counterparts who were less mindful of upkeeping native standards. Negative feelings towards speaking English

were also reflected in the increased production of non-standard English, whereas neutrality or positiveness resulted in increased accuracy.

Educational variables encompassed current studies, years of English instruction, English teachers and language of instruction, English and French proficiency levels and stays in English speaking countries. With higher CEFR level scholastic objectives, it was not surprising that the Matura participants had higher CEFR levels of English and French and in general, used less NSU. It was interesting though that the number of years of study as well as English teachers and language of instruction had minimal influence on the use of non-standard English. Nor did the CEFR level of French have a positive effect on English accuracy until the participants reached a B2 level of French. Proficiency levels of English on the other hand, had a direct correlation with the use of NSU. Lastly, stays in an English speaking country had an overall positive effect on the spoken accuracy of the participants with those who had stayed abroad producing one NSU phw less than their counterparts who hadn't travelled abroad. Another factor to consider here is that young adults with extensive travel experience are more apt to come from affluent households which have been reported to stress the importance of education, which could affect educational success (Brenner et al., 2016). This might explain why there was no conclusive evidence found that the length of a stay in an English speaking country had an effect on the number of NSU, it could have simply been the exposure to another culture.

RQ 2: What effect does early English, taught with a productive focused curriculum, have on the number and type of non-standard features used?

In Chapter 5, research question two was explicitly explored. This thesis has brought forth evidence that the participants in this study who started English in the third grade used fewer

NSU and therefore are assumed to use English more accurately than their counterparts who started formal English instruction four years later in secondary school. Hence, it can be confirmed that the decision to start Early English was warranted and has been successful in educating the participants in this study to a level of proficiency where they can be interviewed and communicate their thoughts and opinions about everyday subjects and their use of English.

The success of Early English was not universally beneficial for it was found that the participants who learned French before English achieved higher self-proclaimed levels of French proficiency than those who were taught Early English. This can perhaps be compensated with the increased awareness or ease of speaking foreign languages which is reflected in the increased opinion of feeling or acting neutral or positively when speaking English from 48% for late English starters to 82% for early starters. Positive association with speaking English is also reflected in the increased percentage of early starters who felt they were not only learners but also users of English. In turn, this positivity was directly correlated with the reduced use of non-standard English or in other words increased accuracy.

RQ 3: How is the acceptance of commonly used non-standard features perceived by the emerging Central Swiss workforce?

The online survey, discussed in Chapter 6, was used to gauge the acceptability of ten sentences produced by participants of this study revealed acceptance rates between 27.80%. and 62.80%. The three least acceptable sentences all had a grammatical error which is assumed to have been recognised as such with tendencies for the older the respondent, the less acceptable they were rated. This tendency for the older the respondents, the higher the percentage of a 'not acceptable' rating continued throughout the survey except for the sentence 'I need English when

I'm **in** (on) holiday'. 'In holiday' is an example of a common Swiss German interference of prepositional use which is considered non-standard English.

Acceptance of further German interferences were also high which suggests that the tolerance for these common NSU related to German interference is much higher than those with a more obvious grammatical nonconformity. The one exception is the 50% acceptance of 'I think she **want** to say it in another way'. This sentence also had a high rate of 'somewhat acceptable' (37.6%) and only 11.7% generally thought it was not acceptable. With a mere 15.78% of the most educated and oldest respondents rating it 'not acceptable' this could indicate widespread acceptance of zero third person. Although the results provide evidence of acceptance, it must be mentioned that the placement of this sentence as the first of ten might have had an influence on its acceptability rate. Further studies with perhaps random sentence placement would provide more precise information. The question of Swiss German speakers' tendencies towards linguistic tolerance is also a factor that would have to be investigated before conclusive statements can be made about the acceptance and the possible innovation effect of zero third person singular in the Swiss context.

7.3   Theoretical implications

In general, this study has shown that it is possible to acquire English as a second or third language in the classrooms of Central Switzerland. Furthermore, it supports the recent foreign language policy changes by demonstrating that the current system works well to produce citizens able to function in a plurilingual society who are very accepting of language variation. The survey revealed a high tolerance for non-standard language. This can be interpreted as agility and acceptance of others with understanding at the forefront as desired by the Swiss education authorities. On the other hand, it could be interpreted as a lack of the general

knowledge, or the interest needed to determine if language use is grammatically correct. This disparity between language exposure in a multilingual context and language accuracy requires further research to better understand the theoretical implications and if the driving force lies in the plurilingual society or the language policies which influence the society.

As mentioned above, this study has given evidence that the decision to teach early English in the third grade can be concluded as being successful in terms of accuracy and generally more positive attitudes towards the use of English. On the other hand, some discrepancies and areas of concern were also discovered. Namely, the primary use of inductive teaching methods during the first four years of instruction in primary school followed by the main use of deductive methods in secondary school. Although Pfenninger and Lendl (2017) have addressed the need for more communication between primary and secondary school, it is argued that it is necessary to reconsider the prescribed use of primarily inductive teaching methods in primary school. Cognitive restraints have been argued to be a reason for delaying deductive methods until children are capable of profiting most from deductive teaching methods (Jaekel et al., 2017). This is obviously in the interest of the well-being of the children. However, Pfenninger and Lendl (2017, p. 446) suggest 'the main goal [of English in primary school] is not necessarily an increase in FL (foreign language) proficiency', but to develop attitudes towards languages and culture (ibid.). I argue that if the recommendations for primarily using inductive teaching methods in primary school are expanded to allow deductive teaching of the areas which can later inhibit accuracy, students will ultimately benefit.

This study has identified areas which could be potentially problematic for learners of English in the German speaking part of Switzerland and thus it is recommended that they are addressed fully from the onset of English instruction in order to avoid fossilization and later uncertainties.

- **Third person singular-s** is one of the most important areas because it has the potential to become an innovation and thus be accepted as a Swiss English feature. If this were to happen, the reputation of the Swiss as proficient plurilinguals would be tarnished. It is recommended that great care be taken to provide impeccable examples of this form and provide guidance on forming it correctly as well as modelling techniques as corrective measures from the onset of English instruction in the third grade.

- **Instruction sequence of relative pronouns *who* and *which*** The corpus analysis revealed that participants that were assumed to have been taught *who/that* at least six months before *which* had a better command when using them in spoken discourse than participants who were assumed to have been taught them concurrently. With a non-standard usage rate of 12%, it is an area which merits attention. Therefore, it is recommended that the relative pronouns *who* and *which* are taught separately.

- **Prepositions** were proven to be used by the majority of participants in a non-standard way and remain an area of concern. *At*, *on* and *from* caused the most confusion and it is recommended that the current teaching methodology be reviewed and revised.

- **Word order inversion of adverbials of frequency and focus** Throughout the corpus instances of non-standard adverbial use were recorded and were used by 27% of the corpus participants especially those at a B1-B2 level. This could indicate some fossilization which could be remedied by early recognition and correction instead of relying on students to eventually understand its correct use once they are explicitly taught.

- **Perfect aspect** was found to be used primarily by the early starters. This suggests that the teaching methods have resulted in a measurable improvement over late starters and it is recommended that the methods being used are continued and monitored for possible future need for development.

- **Adverbs and adjective inversion with -ly** The standard use of *normal/normally* proved to be difficult as the NSU rate recorded 20%. Consequently, it is recommended that its use be closely monitored and early corrective measures taken.

- **Past tense** inconsistencies accounted for 17% of all non-standard uses and is an area that needs to be addressed. Survey results suggest that non-standard uses are recognised as such and were found to be not accepted by the majority of respondents. Thus, there is a gap between knowledge and standard use. As an example, the past tense of *rent* was only used in a standard way one out of six times. It is recommended that both receptive and productive skills are used to strengthen the standard use of the past tense. As Davies (1976) reminds us of the need to build sufficient vocabulary through receptive skills before expecting proficient productive skills.

- **Lexis and NSU correlations** Words lists in Section 5.2.5 were created to highlight the words that were associated with third person singular, prepositions, past simple, plurals and lexical choice with an NSU rate of over 5 %. It is highly recommended that these specific words are given adequate focus in the classroom because they have been proven to be problematic for the participants in this study.

The findings of this study have several pedagogical implications. The preceding examples of potentially problematic language areas are a valuable resource, and it is recommended that these observations are used in teacher training, coursebook design, and classroom exercises. The benefits of knowing which grammatical points and lexis are used less accurately could give precise guidance on which language to focus on to improve accuracy. If primary school teachers were trained to promote accuracy early on, it is conceivable that a smoother transition from primary to secondary school could be achieved.

As mentioned on page 145, coursebooks are produced exclusively for the Swiss market. Therefore, it would be an advantage if the publishers were made aware of the CSC and granted access to better understand the spoken output that has been achieved after the use of their textbooks and advised on how to implement the findings of this study to promote accuracy.

The findings from this study would also be valuable for experienced teachers when they are creating their detailed lesson plans. Although there may be restrictions on which coursebooks must be used, the allocation of time and emphasis on certain language features lies within the discretion of each teacher. With a little creativity, practitioners could create simple exercises focusing on the potentially problematic language areas revealed in this study and share them on existing Swiss teacher platforms[35]. The following two suggestions exemplify how the research findings could be used beginning in primary school. First, by integrating the simple rhyme 'He, she, it, S muss mit' into the classroom from day one and displaying it on a classroom wall for constant reference can reinforce the importance of third-person singular correctness. Another suggestion which could easily be implemented from primary school would be to integrate the words with high NSU rates from section 5.2.5 into stories and or writing assignments. For

---

[35] https://www.zebis.ch/

example, the irregular plurals hair and foot could be emphasised with a short story and picture with the keywords in a primary school classroom.

There once was a very hairy bear. He had hair everywhere except for his feet. On one foot he had one black hair and on the other foot, he has two brown hairs. His feet were cold because he didn't have any hair on his feet, so his grandma knitted him some socks.

The bridge between learner corpus research findings and implementation in the classroom has been an ongoing challenge (Chambers, 2019). In the Swiss context, a plausible approach would be to deliver talks or workshop sessions at the Swiss English Teachers Association (ETAS)[36] conference or Professional Development Days to inform the members of the findings and deliver concrete examples that can be immediately implemented in the classroom. As a second step, a corpus linguistics workshop to introduce corpora and the tools to create and understand it could broaden practitioners' understanding of corpora. This study owes its creation to this type of exposure to research over a decade ago.

7.4   Limitations

Although this study achieved the goal of mapping out the linguistic uniqueness of a group of Swiss participants, the group is small and due to the complexity and variance within the linguistic landscape of Switzerland, the conclusions can give guidance for further areas of research but are difficult to apply to all of Switzerland or even the estimated Swiss German population of 5.3 million.

---

[36] https://www.e-tas.ch/

Further, as mentioned in Section 3.7, the corpus was compiled using somewhat structured interviews and not natural occurring spontaneous speech as recommended by Sinclair (1996). Although Gilquin et al. (2010) state that the warming-up set topic and discussion part of the LINDSEI corpus closely resemble natural speech, the picture description is semi constrained even though participants were free to choose the language used to describe the pictures. Nonetheless, the construct of the CSC was designed to be generally comparable with other learner corpora in the LINDSEI suite, although it does not conform completely with the LINDSEI guidelines.

In hindsight, the survey would have supported the study better if it were designed after the analysis was complete. For example, a sentence with the expression 'she don't like' or 'the brother from my mother' could have been included to determine the acceptability of colloquial expressions used in the English speaking entertainment world. Furthermore, gaining further information on the use of the use of perfect tense and expressions with *normal/normally* would have been beneficial in understanding their current use. Unfortunately, due to time contractions, this was not possible.

7.5   Further research

As the first Swiss English learner corpus, it can be used to suggest directions of interest for further study and they are numerous. The Swiss education system has proven to be able to produce English speakers at different levels of capability but nonetheless able to communicate effectively. This study has questioned the didactical approach throughout and suggests that it continue, however, with adjustments in the recommended areas to maximise accuracy which will benefit students and by increasing their measurable ability which in turn reflects the system positively. From the knowledge gained, four suggestions of further areas of study follow.

The first suggested area of further study would be to compile a comprehensive Swiss Learner corpus of written and spoken English with German L1, French L1 and Italian L1 sub-corpora. This could be used to better understand how English is being used in the entire country as well as identify L1 specific features and compare with additional leaner corpora from other language communities. Ideally, such a Swiss English learner corpus would comply completely with the LINDSEI corpus to facilitate comparison. As such an undertaking would require governmental financial support, the realization is not yet conceivable.

A smaller, but not less important area of research which could be addressed is the relation between multilingual exposure in the home and foreign language accuracy. This study suggests there is a correlation. However, a broader study would be able to determine to which degree and if certain L1s are more advantageous than others and if other factors play a role such as social status, education level of parents, number of siblings or extended family in the household. With knowledge gained in such a study, families with a migration background (currently 37.7%[37] of Swiss residents) could be informed or more skilfully consulted on how to encourage foreign language learning for their children.

This study suggested that neutral or positive feelings towards English had a positive effect on the accuracy of spoken English and negative opinions of English resulted in less accuracy. This was combined with the expectation of speaking with native speakers and importance of being grammatically correct. It is hypothesised that this phenomenon could be related to anxiety and further study into what causes foreign language anxiety and how it can be alleviated would be beneficial.

---

[37] https://www.bfs.admin.ch/bfs/en/home/statistics/population/migration-integration/by-migration-status.html

Lastly, a further interesting research area is suggested due to three indications of possible non-standard vernacular language transfer found in the corpus, with 'she don't like', the use of 'conversate' and 'the sister from my father'. Since the corpus was compiled primarily in a school setting, it is assumed the participants attempted to use 'school English', thus refraining from the use of vernacular language or slang. This is suggested as to why the corpus only contains three swear words as the use of all imaginable English swear words are extremely popular in social media, Swiss television and radio where the use of F*** is common and acceptable. It is these semi-forbidden foreign words that have a magnetized effect on youth and it is suggested that investigating the 'Street English' of Swiss youth would give a further understanding of how English is being used to communicate with other Swiss with the same or different L1.

7. References

ACE (2020) *The Asian Corpus of English*. Director: Kirkpatrick, A.; Researchers: Lixun, W., Patkin, J. and Subhan, S. Available at: https://corpus.eduhk.hk/ace/index.html (Accessed: 15 May 2021).

Adolphs, S. and Knight, D. (2010) 'Building a spoken corpus What are the basics?', in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Abingdon: Routledge, pp. 38-52.

Aijmer, K. (2011) 'Well I'm not sure I think... the use of well by non-native speakers', *International Journal of Corpus Linguistics*, 16 (2), pp. 231-254.

Aijmer, K. (2020) 'Spoken Corpora', in S. Adolphs and Knight, D. (eds.) *The Routledge Handbook of English and Digital Humanities*. Abingdon: Routledge, pp. 5-25.

Aitchison, J. (2012) *Words in the Mind: An Introduction to the Mental Lexicon*. 4th ed. John Wily and Sons Incorporated.

Anderson, R. W. and Shirai, Y. (1994) 'Discourse motivations for some cognitive acquisition principles', *Studies in Second Language Acquisition*, 16 (2), pp. 133–156.

Andreu-Anrés, M. A., Guardiola, A. A., Matararredona, M. B., MacDonald, P., Fleta, B. M. and Sabater, C. P. (2010) 'Analysing EFL Learner Output in the MiLC Project, *It's, but Which Tag?', in Campoy-Cubillo, M. C, Bellés-Fortuno, B. and Gea-Valor, M. L. (eds.) *Corpus Based Approaches to English Language Teaching*. New York: Continuum International Publishing Group, pp. 167-179.

Anthony, L. (2019) *AntConc* (Version 3.5.8) (Computer Software). Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software.

Arnet-Clark, I., Bell, N., Bleuler-Lanz, C., Ritter, G., Rüdiger-Harper, J., Stampfli-Vienny, C. and Wirrer, M. (2005) *Young World 1-4*. Baar: Klett und Balmer Verlag.

Baker, D. (2015) 'Tag questions', *ELT Journal*, (July), 69 (3) pp. 314–318.

Bamgbose, A. (1998) 'Torn between the norms: innovations in world Englishes', *World Englishes*, 17 (1), pp. 1-14.

Bardovi-Harlig, K. (2000) *Tense and aspect in second language Acquisition: Form, meaning, and use*. Oxford: Blackwell.

Baumgarten, N. and House, J. (2010a) 'I think and I don't know in English as lingua franca and native English discourse', *Journal of Pragmatics*, 42, pp. 1184–1200.

Baumgarten, N. and House, J. (2010b) 'Discourse markers in high-stakes ELF academic interaction: Oral exams', *3rd International ELF Conference*. University of Vienna. 22-25 May. Vienna.

Baumgartner, C. (2012) 'Wie früh sollen Kinder Fremdsprachen lernen?', *Bildung Schweiz*, 12, pp. 9-10.

Benner, A. D., Boyle, A. E., Sadler, S. (2016) 'Parental Involvement and Adolescents' Educational Success: The Roles of Prior Achievement and Socioeconomic Status', *J Youth Adolescence*. (5 February), New York: Springer Science+Business Media.

Biber, D. (1988) *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. (2000) *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Ltd.

Bieber, J. (2012) *She don't like the lights*. (Sound Recording) (Rodney Jerkins and Andre Lindal).

Bildungs- und Kulturdepartement des Kantons Luzern (2016) *Sprachen Lehrplan 21*. Lucerne: Bildungs- und Kulturdepartement des Kantons Luzern.

Bildungs- und Kulturdepartement des Kantons Luzern (2020) *Lehrplan für die Volksschule des Kantons Luzern*. Bern: Bildungs- und Kulturdepartement des Kantons Luzern.

Bildungsdirection Nidwalden (2008) *Wohin schulwege führen*. Stans: Bildungsdirection Nidwalden.

Bon, P. B. (1948) 'English Words in Swiss German Usage', *American Speech*, (Oct.-Dec.), 23 (3/4), pp. 232-235.

Bordet, L. (2017) 'From vogue words to lexicalized intensifying words: the renewal and recycling of intensifiers in English. A case-study of very, really, so and totally', *Journal in English Lexicology*, 10, doi.org/10.4000/lexis.1125.

Brady, J. (2015) 'Dialect, power and politics: standard English and adolescent identities', *Literacy*, (September), 49 (3), pp. 149-157.

Brand, C. and Kämmerer, S (2006) 'The Louvain International Database of Spoken English Interlanguage (LINDSEI): Compiling the German component', in Braun, S., Kohn, K. and Mukherjee, J. (eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods*. 127–40. Frankfurt am Main: Peter Lang.

Britain, D. (2010) 'Grammatical variation in the contemporary spoken English of England', in A. Kirkpatrick, (ed.) *The Routledge Handbook of World Englishes*. Abingdon: Routledge, pp. 37-58.

Brohy, C. (2001) 'Generic and/or Specific Advantages of Bilingualism in a Dynamic Plurilingual Situation: The Case of French as Official L3 in the School of Samedan (Switzerland)', *International Journal of Bilingual Education and Bilingualism*, pp. 38-49.
.
Brohy, C. (2005) 'Trilingual Education in Switzerland', *International Journal of the Sociology of Language*, (1 January), pp. 133-148.

Browman, C. P. (1978) 'Tip of the Tongue and Slip of the Ear: Implications for Language Processing', *UCLA Working Papers in Phonetics* 42.

Brown, R. (1973) *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Bula, A. and Segura, J. (2019) *People in education 2019.* (ed), Neuchâtel: Federal Statistical Office (FSO).

Bundesamt für Statistik (2010) *Bildungsstatistik 2009 Bildung und Wissen 15*. Neuchâtel.

Bundesamt für Statistik (2008) *Bildungsstatistik*. Available at: http://www.bfs.admin.ch (Accessed 12 June 2018).

Bundesamt, vom 18. April 1999 (Stand am 7. März 2010). *Bundesverfassung der Schweizerischen Eidgenossenschaft Art.18 Sprachenfreiheit*. Available at: http://www. admin.ch/ch/d/sr/101/index.html (Accessed 20 June 2018).

Burnard, L. (2007) *British National Corpus*. Reference Guide for the British National Corpus (XML Edition): http://www.natcorp.ox.ac.uk/XMLedition/URG/ (accessed 1 April 2021).

Buysse, L. (2015) ''Well it's not very ideal …' The pragmatic marker well in learner English'. *Intercultural Pragmatics*, 12 (1), pp. 59-89.

Callies, M. (2015) 'Learner corpus methodology', in Granger, S., Gilquin, G. and Meunier, F. (eds.) *The Cambridge Handbook of Learner Corpus Research*. Cambridge: Cambridge University Press, pp. 35-56.

Callies, M. and Götz, S. (2015) 'Learner corpora in language testing and assessment Prospects and challenges', in Callies, M. and Götz, S. (eds.) *Learner corpora in language testing and assessment*. Amsterdam: John Benjamins Publishing Company, pp. 1-9.

Cambridge Advanced Learner's Dictionary (2005) 2nd ed. Cambridge: Cambridge University Press.

Cambridge English Language Teaching (2010) Available at: http://www.cambridge.org /ch/elt/catalogue/subject/custom/item3646603/Cambridge-International-Corpus-Cambridge-Learner-Corpus/?site_locale=de_CH (Accessed 3 July 2020).

Carlsen, C. (2012) 'Proficiency Level—a Fuzzy Variable in Computer Learner Corpora', *Applied Linguistics*, 33 (2), pp. 161-183.

Carter, R. (1998a) 'Orders of reality: CANCODE communication, and culture', *ELT Journal*, (January) 52 (1), pp. 43-56.

Carter, R. (1998b) *Vocabulary Applied Linguistic perspectives*. 2nd ed. London: Routledge.

Carter, R. and McCarthy, M. (1995) 'Grammar and the Spoken Language', *Applied Linguistics*, 16 (2), pp. 141-158.

Carter, R. and McCarthy, M. (2006) *Cambridge Grammar of English: a Comprehensive Guide*. Cambridge: Cambridge University Press.

Carter, R. and McCarthy, M. (2017) 'Spoken Grammar: Where Are We and Where Are We Going?', *Applied Linguistics*, 38 (1), pp. 1–20.

Centre for English Corpus Linguistics (2006) Available at: http://www.fltr.ucl.ac.be /FLTR/GERM/ETAN/CECL/research%20learner%20corpora.html (Accessed 23 July 2010).

Chambers, A. (2019) 'Towards the corpus revolution? Bridging the research–practice gap.' *Language Teaching*, 52, pp. 460–475.

Clancy, B. (2010) 'Building a corpus to represent a variety of a language', in A. O'Keeffe and M. McCarthy, (eds.) *The Routledge Handbook of Corpus Linguistics*. Abingdon: Routledge, pp. 80-92.

Clark, E. V. (1978) 'Discovering what words can do', in (ed.) *Papers from the Paracession on the Lexicon*. Chicago: Chicago Linguistic Society, University of Chicago, pp. 34-57.

Cogo, A. (2012) 'English as a Lingua Franca: concepts, use and implications', *ELT Journal*, (January), 66 (1), pp. 97-105.

Cogo, A. and Dewey, M. (2006) 'Efficiency in ELF Communication: From Pragmatic Motives to Lexico-grammatical Innovation', *Nordic Journal of English Studies* 5 (2) pp. 59–93.

Cogo, A. and Dewey, M. (2012) *Analysing English as a Lingua Franca A Corpus Driven Investigation*. London: Continuum International Publishing Group.

Coleman, J. (1997) 'Residence abroad within language study', *Language Teaching*, 30, pp. 1-20.

Cook, V. (2002) *Portraits of the L2 User*. Clevedon: Multilingual Matters.

Cook, V. (2008) *Second Language Learning and Language Teaching* 4th edition. London: Routledge.

Cook, V. (2009) 'The goals of ELT: Reproducing native-speakers or promoting multicompetence among second language users?', Lisbon, *ILTEC- Instituto de Linguística Teórica e Computacional*, pp. 9-22.

Corder, S. P. (1967) 'The significance of learner's errors', *International Review of Applied Linguistics in Language Teaching*, 5, pp. 161-170.

Corder, S. P. (1971) 'Idiosyncratic dialects and error analysis', *International Review of Applied Linguistics in Language Teaching*, 9 (2), pp. 147-160.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Council of Europe (2020) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume*, Strasbourg: Council of Europe Publishing.

Crystal, D. (1995) *The Cambridge Encyclopedia of the English Language* 2nd ed. Cambridge: Cambridge University Press.

Crystal, D. (1997) *English as a global language*. Cambridge: Cambridge University Press.

Crystal, D. (2008) 'Two thousand million?', *English Today*, (March) 24 (1), pp. 3-6.

Davies, N. F. (1976) 'Receptive Skills in Foreign Language Learning', *The Modern Language Journal*, (December), 60 (8), pp. 440-443.

Davydova, J. (2011) *The Present Perfect in Non-Native Englishes: A Corpus-Based Study of Variation*. Berlin: De Gruyter Mouton.

de Villiers, J. and de Villiers, P. (1973) 'A Cross-sectional study of the acquisition of grammatical morphemes', *Journal of Psycholinguistic Research*, pp. 267-78.

Demont-Heinrich, C. (2005) 'Language and National Identity in the Era of Globalization: The Case of English in Switzerland', *Journal of Communication Inquiry*, (January), 29 (1), pp. 66-84.

Dewey, M. (2009) 'English as a lingua franca: heightened variability and theoretical implications', in Mauranen, A. and Ranta, E. (eds.) *English as a lingua franca*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 60-83.

Diaz-Negrillo, A. and Fernandez-Dominguez, J. (2006) 'Error tagging systems for learner corpora', *RESLA*, 19, pp. 83-102.

Diaz-Negrillo, A. and Thompson, P. (2013) 'Learner corpora: looking towards the future', in Diaz-Negrillo, A. Ballier, N. and Thompson, P. (eds.) *Automatic Treatment and Analysis of Learner Corpus Data. Studies in Corpus Linguistics*. Amsterdam: John Benjamins, pp. 9-30.

Dörnyei, Z. (2009) 'The L2 Motivational Self System', in Dörnyei, Z. and Ushiodo, E. (eds.) *Motivation, Language Identity and the L2 Self*. Bristol: Multilingual Matters, pp. 9-42.

Dose-Heidelmayer, S. and Götz, S. (2016) 'The progressive in spoken learner language: A corpus-based analysis of use and misuse', *IRAL*, 54 (3), pp. 229-256.

Driscoll, L. (2005) *Common mistakes at PET and how to avoid them*. Cambridge: Cambridge University Press.

Driscoll, L. (2007) *Common mistakes at KET and how to avoid them*. Cambridge: Cambridge University Press.

Dröschel, Y. (2011) *Lingua Franca English*. Bern: Peter Lang AG.

Durham, M. (2007) *English in Switzerland: Inherent variation in a non-native speech community*. Fribourg: Unpublished Doctoral Thesis.

Durham, M. (2016) 'English as a Lingua Franca: Forms and Features in the Swiss Context', *Cahiers de l'Institut de Linguistique et des Sciences du Langage,* 48, pp. 170-118.

Dürmüller, U. (1994) 'Multilingual Talk or English Only? The Swiss Experience', *Sociolinguistica,*8 pp. 44-64.

Dürmüller, U. (2002) 'English in Switzerland: from foreign language to lingua franca', in Allerton, D., Skandera, P. and Tschichold, C. (eds.) *Perspectives on English as a world language.* Basel: Schwabe, pp. 114-123.

EDK (2019) *Schweizerische Konferenz der Kantonalen Erziehungsdirection*. Available at: http://www.edk.ch/dyn/14901.php (Accessed 15 January 2019).

Edwards, J. (2009) *Language and Identity*. Cambridge: Cambridge University Press.

ELFA (2008) The Corpus of English as a Lingua Franca in Academic Settings Director: Mauranen, A. Available at: http://www.helsinki.fi/elfa/ (Accessed 7 May 2021).

Federal Statistical Office (FSO) (2018) *Switzerland's population 2017*. Neuchâtel: Federal Statistical Office FSO.

Federal Statistical Office STAT-TAB - interactive tables (FSO) (2020) Available at: https://www.pxweb.bfs.admin.ch/pxweb/en/px-x-1502040100_131/px-x1502040100_131/px-x- 1502040100_131.px/?rxid=9a55bede-17b3-4f77-87db-52237fc92f74 (Accessed 22 March 2020).

Federal Statistical Office (2019) *Reisen der Schweizer Wohnbevölkerung 2018*. Neuchâtel: Bundesamt für Statistik (BFS).

Felix, S. W. and Hahn, A. (1985) 'Natural process in classroom second-language learning'. *Applied Linguistics*, 6 (3), pp. 223-238.

Ferguson, C. A. (1972) *Language Structures and Language Use*. Stanford: Stanford University Press.

Figueras, N., North, B., Sauli, T., Verhelst, N. and Van Avermaet, P. (2005) 'Relating examinations to the Common European Framework: a manual'. *Language Testing*, 22, pp. 261-279.

Firth, A. (1996) 'The discursive accomplishment of normality. On 'lingua franca' English and conversation analysis'. *Journal of Pragmatics*, 26, pp. 237-259.

Firth, A. (2009) 'The lingua franca factor'. *Intercultural Pragmatics*, 6 (2), pp. 147-170.

Firth, A. and Wagner, J. (1989) 'On Discourse, Communication, and (Some) Fundamental Concepts in SLA Research'. *Modern Language Journal*, 81, pp. 285-300.

Fischer, M., Giancola-Bürer, C., Gutmann, B. and Hahin, T. (2011) *Open World 1-3*. Baar: Klett und Balmer Verlag.

Fjelkestam-Nilsson, B. (1983) 'Also and Too: A Corpus-Based Study of Their Frequency and Use in Modern English'. *Stockholm Studies in English*. Stockholm: Almqvist and Wiksell International, pp. viii + 170.

Flaugergues, A. D. (2016) *Sprachliche Praktiken in der Schweiz: Erste Ergebnisse der Erhebung zur Sprache, Religion und Kultur 2014*. Neuchâtel: Bundesamt für Statistik (BFS).

Flowerdew, L. (2000) 'Investigating referential and pragmatic errors in a learner corpus'. in Burnard, L and T. McEnery, T. (eds.) *Rethinking language Pedagogy from a Corpus Prospective*. Frankfurt am Main: Peter Lang, pp. 145-154.

Freed, B. (1998) 'An overview of issues and research in language learning in a study-abroad setting'. *Frontiers: The Interdisciplinary Journal of Study abroad*, 4, pp. 21-60.

Fuchs, R., Götz, S. and Werner, V. (2016) 'The present perfect in learner Englishes: A corpus-based case study on L1 German intermediate and advanced speech and writing'. in Werner, V., Seoane, E. and Suàrez-Gomez, C. (eds.) *Re-assessing the present perfect*. Berlin: Walter de Gruyter, pp. 297-338.

Gablasova, D., Brezina, V. and McEnery, T. (2017) 'Exploring Learner Language Through Corpora: Comparing and Interpreting Corpus Frequency Information'. *Language Learning*, (June), 67, pp. 130-154.

Garside, R. and Smith, N. (1997) 'A hybrid grammatical tagger: CLAWS4'. in Garside, R., Leech, G. and McEnery, T. (eds.) *Corpus annotation: Linguistic information from computer text corpora*. London: Longman, pp. 102–121.

Garton-Sprenger, J. and Prowse, P. (2011) *New Inspiration 1-3*. Oxford: Macmillan.

Gilquin, G. (2008) 'Hesitation markers among EFL learners: Pragmatic deficiency or difference'. in Romero-Trillo, J. (ed.) *Pragmatics and Corpus Linguistics A mutualistic entente*. Berlin: Mouton de Gruyter, pp. 119-150.

Gilquin, G. (2015) 'From design to collection of learner corpora'. in Granger, S., Gilquin G. and Meunier, F. (eds.) *The Cambridge Handbook of Learner Corpus Research*. Cambridge: Cambridge University Press, pp. 9-34.

Gilquin, G., De Cock, S. and Granger, S. (2010) *Louvain International Database of Spoken English Interlanguage (LINDSEI)*. Louvain-la-Neuve: Presses universitaires de Louvain.

Gnos, C. (2012) *Englischkompetenzen am Ende der 6. Klasse: Überprüfung der Lernziele Sprachstandserhebung im Kanton Luzern*. Luzern: Dienststelle Volksschulbildung des Kantons Luzern (DVS).

Götz, S. (2013) *Fluency in Native and Nonnative English Speech*. Amsterdam: John Benjamins Publishing Co.

Götz, S. (2015) 'Tense and aspect errors in spoken learner English: Implications for Language Testing and Assessment'. in Callies, M. and Götz, S. (eds.) *Learner Corpora in Language Testing and Assessment*. Amsterdam: John Benjamins, pp. 191-215.

Götz, S. and Mukherjee, J. (2018) 'Investigating the effect of the study abroad variable on learner output: A pseudo-longitudinal study on spoken German learner English'. in Brezina, V. and Flowerdew, L. (eds.) *Learner Corpus Research: New Perspectives and Applications*. London: Bloomsbury, pp. 47-68.

Gráf, T. (2017) 'The Story of the Learner Corpus LINDSEI_CZ'. *Číslo.* 2, pp. 22-35.

Granger, S., 1996. 'From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora'. in Aijmer, K., Altenberg, B. and Johansson, M. (eds.) *Languages in Contrast*. Lund: Lund University Press, pp. 37-51.

Granger, S. (2008) 'Learner corpora in foreign language education'. in Deusen-Scholl, N. V. and Hornberger, N. H. (eds.) *Encyclopedia of Language and Education*. New York: Springer, pp. 337-51.

Granger, S. (2013) 'A Bird's-eye view of learner corpus research'. in Granger, S., Hung, J. and Pech-Tyson, S. (eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: Benjamins, pp. 3-33.

Granger, S. (2015) 'Contrastive interlanguage analysis: A reappraisal'. *International Journal of Learner Corpus Research*, (April), pp. 7-24.

Granger, S., Gilquin, G. and Meunier, F. (2005) *The Cambridge Handbook of Learner Corpus Research*. Cambridge: Cambridge University Press.

Gregersen, T. and Horwitz, E. K. (2002) 'Language Learning and Perfectionism: Anxious and Non-Anxious Language Learners' Reactions to Their Own Oral Performance'. *The Modern Language Journal*, (Winter), 86 (4), pp. 562-570.

Grin, F. (1998) 'Language Policy in Multilingual Switzerland: Overview and Recent Developments'. *Cicle de confèrencies sobre política lingüística Direcció general de política lingüística*, Barcellona, (4 December), pp. 1-9.

Grin, F. (2001) 'English as economic value: facts and fallacies'. *World Englishes*. 20 (1), pp. 65-78.

Grin, F. and Korth, B. (2005) 'On the Reciprocal Influence of Language Politics and language Education: The Case of English in Switzerland'. *Language Policy,* 4, pp. 67–85.

Grossenbacher, S. and Vögeli-Mantovani, U. (2010) *Sprachenpolitik und Bildungsstrategien in der Schweiz*, Aarau: Schweizerische Koordinationsstelle für Bildungsforschung.

Gross, M. (2017) *ROMANSH The Romansh language in education in Switzerland*, The Netherlands: Mercator European Research Centre on Multilingualism and Language Learning.

Haag, P. (2016) 'WIL: Schweizer Kinder in der Minderheit', *Tagblatt*. 23.08.2016

https://www.tagblatt.ch/ostschweiz/wil/wil-schweizer-kinder-in-der-minderheit-ld.584606.

Hadikin, G. (2014) 'A, an and the environments in Spoken Korean English'. *Corpora*, 9 (1), pp. 1-28.

Haenni-Hoti, A. and Heinzmann, S. (2009) 'Zwei Fremdsprachen auf der Primarstufe Wie wirksam ist das neue Modell 3/5 mit der Sprachenfolge Englisch-Französisch?'. *Babylonia*, 2 (9), pp. 16-24.

Haenni-Hoti, A., Heinzmann, S., Müller, M., Oliveira, M., Wicki, W. and Werlen, E. (2011) 'Introducing a second foreign language in Swiss primary schools: the effect of L2 listening and reading skills on L3 acquisition'. *International Journal of Multilingualism*, pp. 98-116.

Hägi, S. and Scharloth, J. (2005) 'Ist Standarddeutsch für Deutschschweizer eine Fremdsprache? Untersuchungen zu einem Topos des sprachreflexen Diskurses'. *Linguistik Online*, 24 (3/05) pp. 19-47.

Hardmeier, S. (2017) *EMPFEHLUNGEN zum Fremdsprachenunterricht (Landessprachen und Englisch) in der obligatorischen Schule*, Zürich: Schweizerischen Konferenz der kantonalen Erziehungsdirektoren.

Hega, G. M. (2001) 'Regional Identity, Language and Education Policy in Switzerland'. *Compare*, 31 (2), pp. 205-227.

Heinzmann, S., Müller, M., Oliveira, M., Haenni Hoti, A. and Wicki, W. (2010) *Kurzzusammenfassung des Schlussberichts zur Studie Englisch und Französisch auf der Primarstufe-Verlängerung des NFP-56-Projekts*, Luzern: Pädagogische Hochschule Zentralschweiz Luzern.

Herder, J. G. (1794) *Briefe zu Beförderung der Humanität*. Riga: Johann Friedrich Hartknoch.

Hinrichs, L. and Bohmann, A. (2020) 'Sociolinguistics'. in Adolphs, S. and Knight, D. (eds.) *The Routledge Handbook of English Language and Digital Humanities*. Abingdon: Routledge, pp. 283-305.

Hogg, M. A., Joyce, N. and Dominic, A. (1984) 'Diglossia in Switzerland? A Social Identity Analysis of Speaker Evaluations'. *Journal of Language and Social Psychology*. 3, pp. 185-196.

House, J. (1999) 'Misunderstanding in intercultural communication: Interactions in English as a lingua franca and the myth of mutual intelligibility'. in Gnutzmann, C. (ed.) *Teaching and learning English as a global language.* Tübingen: Stauffenberg, pp. 73-89.

House, J. (2001) 'A stateless language that Europe must embrace'. *The Guardian Weekly*, 19 April 2001.

House, J. (2013) 'Developing pragmatic competence in English as a lingua franca: Using discourse markers to express (inter)subjectivity and connectivity'. *Journal of Pragmatics*, 59, pp. 57-67.

House, J. and Lévey-Tödter, M. (2010) 'Linguistic competence and professional identity in English medium instructional discourse'. in Apfelbaum, B. M. (ed.) *Multilingualism at work*. Amsterdam, Netherlands: John Benjamins, pp. 13-46.

Housen, A., Kuiken, F. and Vedder, I. (2012) 'Complexity, Accuracy and Fluency Definitions, measurement and research'. in Housen, A., Kuiken, F. and Vedder, I. (eds.) *Dimensions of L2 Performance and Proficiency, Complexity, Accuracy and Fluency in SLA*. Amsterdam: John Benjamins Publishing Company, pp. 1-20.

Hudson, R. (1996) *Sociolinguistics*. 2nd ed. Cambridge: Cambridge University Press.

Hülmbauer, C. (2007) 'You moved, aren't?': the relationship between lexicogrammatical correctness and communicative effectiveness in English as a lingua franca'. *Vienna English Working Papers* 16, pp. 3-35.

Hulstijn, J. H. (2010). Linking L2 proficiency to L2 acquisition: opportunities and challenges of profiling research. In *Communicative proficiency and linguistic development: intersections between SLA and language testing research* (S. 233-238). Amsterdam: European Second Language Association.

Hunnius, Y. (2010) 'Bewerbung Englischkenntnisse entscheiden'. *Neue Luzerner Zeitung*, (13 February). p. 15.

Hunston, S. (2002) *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Hyltenstam, K. and Abrahamsson, N. (2012) 'High-Level Acquisition, Learning and Use'. *Studies in Second language Acquisition*, 34, pp. 177-186.

Information services of the Federal Chancellery, (2014). *The Swiss confederation a brief guide*. Available at: www.bundespublikationen.admin.ch (Accessed 21 February 2014).

Jaekel, N., Schurig, M., Florian, M., & Ritter, M. (2017). 'From early starters to late finishers? A longitudinal study of early foreign language learning in school'. *Language Learning*, 67 (3), 631-664.

Jenkins, J. (2000) *The Phonology of English as an International Language*. Oxford: Oxford University Press.

Jenkins, J. (2007) *English as a Lingua France: Attitude and identity*. Oxford: Oxford University Press.

Jenkins, J., Modiano, M. and Seidlhofer, B. (2001) 'Euro-English'. *English Today*, (October), pp. 13-19.

Kachru, B. B. (1985) 'Standards, codification, and sociolinguistic realism: The English language in the outer circle'. in Quirk, R. and Widdowson, H. (eds.) *English in the World: Teaching and Learning the Language and Literature*. Cambridge: Cambridge University Press, pp. 11-30.

Kaneko, T. (2009) 'Use of mother tongue in English-as-a-foreign-language speech by Japanese university students'. *Gakuen*, 822, pp. 25-41.

Kanton Luzern Bildungs- und Kulturdepartement (2020) Kanton Luzern Bildungs- und Kulturdepartement. Available at: https://lu.lehrplan.ch/index.php?code=e|1|3 (Accessed 07.11.2020 November 2020).

Kennedy, G. (1998) *An Introduction to Corpus Linguistics*. Harlow: Pearson Education Limited.

Kerswill, P. (2013) 'Identity, ethnicity and place: the construction of youth language in London'. in Auer, P., Hilpert, M., Stukenbrock, A. and Szmrecsanyi, B. (eds.) *Space in language and linguistics*. Berlin: Walter de Gruyter, pp. 128-164.

King, N., Horrocks, C. and Brooks, J. (2019) *Interviews in Quantative Research*. 2nd ed. London: Sage Publishers.

Kirkpatrick, A. (2007) *World Englishes Implications for International Communication and English Language Teaching*. Cambridge: Cambridge University Press.

Kirkpatrick, A. (2016) 'The Asian Corpus of English – introduction to the special issue'. *JELF*, 5 (2), pp. 225-228.

Kizil, A. S. and Kilimci, A. (2014) 'Recurrent phrases in Turkish ELF learners' spoken interlanguage: A corpus driven structural and functional analysis'. *Journal of Language and Linguistic Studies*, 10 (1), pp. 195-210.

Kollegium St. Fidelis (2020) *Lehr- und Stoffplan*. Stans: Amt für Berufsbildung und Mittelschule.

Korait, A. and Lieblich, I. (1974) 'What does a person in a 'TOT' state know that a person in a 'don't know' state doesn't know'. *Memory and Cognition*, 2, pp. 647–55.

Kortmannn B. (2010) 'Variation across Englishes Syntax', in A. Kirkpatrick, (ed.) *The Routledge Handbook of World Englishes*. Abingdon: Routledge, pp. 400-424.

Krashen, S. (1977) 'Some issues relating to the Monitor Model'. in *On TESOL 77*. Washington DC: TESOL, pp. 144-158.

Krause, O. (1997) 'Progressiv-Konstruktionen im Deutschen im Vergleich mit dem Niederländischen, Englischen und Italienischen', *Sprachtypologie und Universalienforschung* 50 pp. 1-31.

Labov, W. (1972a) *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1972b) *Language in the Inner City Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.

Laitinen, M. (2020) 'Empirical perspectives on English as a lingua franca'. *World Englishes*, 39, pp. 427–442.

Larsen-Freeman, D. (2009) 'Adjusting Expectations: The Study of Complexity, Accuracy, and Fluency in Second Language Acquisition'. *Applied Linguistics*, 1 (11), pp. 1-11.

Leech, G. (2000) 'Grammars of Spoken English: New Outcomes of Corpus-Oriented Research'. *Language Learning*, (December), 50 (4), pp. 675-724.

Leech, G. (2005) 'Adding Linguistic Annotation'. in Wynne, M. (ed.) *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow Books, pp. 17-29.

Leech, G. (2011) 'Frequency, corpora and language learning'. in Meunier, F. DeCock S. and Gilquin, G. (eds.) *A taste for corpora. In honour of Sylviane Granger*. Amsterdam: John Benjamins, pp. 7-31.

Leemann, A. and Siebenhaar, B. (2006) 'Prosodic Features of Spontaneous Utterance-Initial Phrases in Bernese and Valais Swiss German'. *Taipei: Insitute of Linguistics, Academia Sinica*, pp. 127-142.

Lehrplan 21 (2021) Available at: https://www.lehrplan21.ch/ (Accessed 18 April 2021).

Lehrplangruppe Englisch der Bildungsregion Zentralschweiz (2004) *Der Lehrplan Englisch für das 3.-9. Schuljahr*. Luzern: Bildungsplanung Zentralschweiz.

Lehrplangruppe Französisch der Innerschweizer Erziehungsdirektorenkonferenz (2000) *Lehrplan Französisch 5. bis 9. Schuljahr*. Ebikon: Bildungsplanung Zentralschweiz.

LePage, R. and Tabouret-Keller, A. (1985) *Acts of Identity*. Cambridge: Cambridge University Press.

Lightbown, P. M. and Spada, N. (2006) *How languages are learned*. 3rd ed. Oxford: Oxford University Press.

Lindemann, S. (2005) 'Who speaks "broken English"? US undergraduates' perceptions of non-native English'. *International Journal of Applied Linguistics*, 15 (2), pp. 187-212.

LINDSEI (2006) The Louvain International Database of Spoken English Interlanguage. Available at: http://www.fltr.ucl.ac.be/fltr/germ/etan/cecl/Cecl-Projects/Lindsei/lindsei.htm (Accessed 15 January 2015).

Liu, D. and Gleason, J. L. (2002) 'Acquisition of the article the by nonnative speakers of English: Analysis of Four Nongeneric Uses'. *SSLA*, 24, pp. 1-26.

Lorenz, G. (1998) 'Overstatement in advanced learners' writing: Stylistic aspects of adjective intensification'. in Granger, S. (ed.) *Learner English on Computer*. London: Longman, pp. 53-66.

Love, R., Dembry, C., Hardie, A., Brezina, V. and McEnery, T. (2017) 'The Spoken BNC2014 Designing and building a spoken corpus of everyday conversations'. *International Journal of Corpus Linguistics*, 22 (3), pp. 319-344.

Lüdi, G. (2018) 'Das Gesamtsprachenkonzept von 1998-20 Jahre Danach'. *Babylonia*, 3, pp. 14-20.

Lüdi, G. and Werlen, I. (2005) *Eidgenössische Volkszählung 2000 Sprachenlandschaft in der Schweiz*, Neuchâtel: Bundesamt für Statistik (BFS).

Mauranen, A. (2011) 'Learners and Users - Who do we want corpus data from?'. in Meunier, F., DeCock, S., Gilquin, G. and Paquot, M. (eds.) *A taste for Corpora*. Amsterdam: John Benjamins B.V., pp. 155-172.

Mauranen, A., and Ranta, E. (eds.) (2009) *English as a Lingua Franca Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars Publishing.

McArthur, T. (1987) 'The English Languages?'. *English Today*, (11 July), pp. 9-13.

McCarthy, M. (1998) *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press.

McCarthy, M. and Carter, R. (2001) 'Size Isn't Everything: Spoken English, Corpus, and the Classroom. Teachers of English to Speakers of Other Languages', *TESOL*. 35 (2), pp. 337-340.

McEnery, T. and Hardie, A. (2012) *Corpus Linguistics: Method. theory and practice*. Cambridge: Cambridge University Press.

Meierkord, C. (2013) 'English as a Lingua Franca'. in Chapelle, C. A. (ed.) *The Encyclopedia of Applied Linguistics*. Oxford: Blackwell Publishing Ltd. pp. 1-7 DOI: 10.1002/9781405198431.wbeal0375.

Mendikoetxea, A., Murcia Bielsa, S. and Rollinson, P. (2010) 'Focus on Errors: Learner Corpora as Pedagogical Tools'. in *Corpus-based Approaches to English language Teaching*. London: Continuum, pp. 180-194.

Meyer, M. (2018) '"Je Früher Desto Besser?" - zur Neuroplastizität des Spracherwerbs'. *Babylonia*, 3, pp. 74-78.

Modiano, M. (1999) 'International English in the Global Village'. *English Today* 58, (April) 15 (2), pp. 22-28.

Modiano, M. (2009) 'Inclusive/exclusive? English as a lingua franca in the European Union'. *World Englishes*, 28 (2) pp. 208-223.

Mohammadkhani, A., Eslamdoost, S. and Gholamreza'i, S. (2011) 'An investigation of the role of instruction in second language: A case of third person singular -s'. *Procedia - Social and Behavioral Sciences*. (29), pp. 910-916.

Mollin, S. (2006) *Euro-English*. Tübingen: Narr Francke Attempto Verlag.

Moore, J. (2005) *Common Mistakes at Proficiency and how to avoid them*. Cambridge: Cambridge University Press.

Mortensen, J. (2013) 'Notes on English used as a lingua franca as an object of study'. *JELF*, 2 (1), pp. 25 – 46.

Mumford, S. (2009) 'An analysis of spoken grammar: the case for production'. *ELT Journal*, (April), 63 (2), pp. 137-144.

Muñoz, C. (2014) 'Starting age and other influential factors: Insights from learner interviews'. *Studies in Second language Learning and Teaching*, 4, pp. 465-484.

Murray, H. (2003) 'Swiss English teachers and Euro-English: Attitudes to a non-native variety'. *Vereinigung für angewandte Linguistik in der Schweiz*, 77, pp. 147-165.

Neuliep, J. W. and Speten-Hansen, K. M. (2013) 'The influence of ethnocentrism on social perceptions of nonnative accents'. *Language and Communication*, (5 June), 33, pp. 167-176.

Nguyen, D. Q. (2017) SWI Swissinfo. Available at: https://www.swissinfo.ch/eng/migration-series-part-1-_who-are-the-25-foreign-population-in-switzerland/42412156 (Accessed 08 November 2020).

North, B. and Hughes, G. (2003) *CEF Illustrative Performance Samples: For Relating Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF)*, Zurich: Council of Europe.

Norton, B. (2000) *Identity and language learning: Gender, ethnicity and educational change*. Harlow, England: Longman/Pearson Education.

OED (2021) *Oxford English Dictionary online*, Oxford: Oxford University Press.

Oxford English Dictionary (2021). *redundant*, adj. and n. Oxford: Oxford University Press. Available at: www.oed.com/view/Entry/160538 (Accessed: 25 March 2021).

Ogden, C. K. (1930) *Basic English: A general introduction with rules and grammar*. London: Paul, Trench, Trubner.

O'Keeffe, A. and McCarthy, M. (2010) 'Historical perspective what are corpora and how have they evolved?'. in O'Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. Abingdon: Routledge, pp. 3-13.

O'Keeffe, A., McCarthy, M. and Carter, R. (2007) *From Corpus to Classroom*. Cambridge: Cambridge University Press.

Önen, S. (2014) *Lexico-grammatical features of English as a lingua franca: A corpus-based study on spoken interactions*. Istanbul: T.C. İstanbul Üniversitesi Egitim Bilimleri Enstitusu, Unpublished doctoral thesis.

Oswald, S. (2010) *Swiss English, English as a Lingua Franca or Learner English: A Study of the Spoken English of 18-20 Year Old Students from Central Switzerland and Their Attitudes Towards English*. Birmingham: Unpublished Master Dissertation.

Oswald, S. (2014) 'Patterns of English usage in Central Switzerland'. *ETAS Journal*, 31 (3), pp. 16-18.

Pädagogische Hochschule Luzern (2020) PH Luzern Pädagogische Hochschule. Available at: https://www.phlu.ch/studium/studiengaenge.html (Accessed 11 November 2020).

Petkova, M. (2012) 'Die Deutschschweizer Diglossie: eine Kategorie mit fuzzy boundaries'. *Zeitschrift für Literaturwissenschaft und Linguistik*: Lili, 42 (166), pp. 61-87.

Pfenninger, S. E. and Lendl, J. (2017) 'Transitional woes: On the impact of L2 input continuity'. *Studies in Second Language Learning and Teaching*, 3, pp. 443-469.

Pfenninger, S. E. and Singleton, D. (2019) 'Starting Age Overshadowed: The Primacy of Differential Environmental and Family Support Effects on Second Language Attainment in an Instructional Context'. *Language Learning*, (March), 69 (1), pp. 207-234.

Pfenninger, S. and Watts, R. (2019) 'English in Switzerland'. in Hicky, R. (ed.) *English in the German-Speaking World*. Cambridge: Cambridge University Press, pp. 315-333.

Pica, T. (1986) 'Adult acquisition of English as a second language under different conditions of exposure'. *Language Learning*, (December), 33 (4), pp. 465-497.

Pichard, A. (2018) 'Wie Man eine Fremdsprache an die Wand fährt. Eine Replik auf Georges Lüdis Bilanz zum Gesamtsprachenkonzept'. *Babylonia*, 3, pp. 21-25.

Pilipp, F. (1997) 'Scatological Reality and Humor in Martin Walser's "The Inner Man"'. *South Atlantic Review*,(Winter), 62 (1), pp. 65-78.

Pölzl, U. and Seidlhoffer, B. (2006) 'In and on their own terms: the ''habitat factor'' in English as a lingua franca interactions'. *International Journal of Sociology of Language* 177, pp. 151–176.

Powell, D. (2005) *Common mistakes at CAE and how to avoid them*. Cambridge: Cambridge University Press.

Quan, L. and Weisser, M. (2015) 'A study of 'self-repair' operations in conversation by Chinese English learners'. *System*, 49, pp. 39-49.

Quirk, R. (1985) 'International communication and the concept of Nuclear English'. in Smith, L. E. (ed.) *English for cross-cultural communication*. London: Macmillan, pp. 151-165.

Quirk, R. (1990) 'Language varieties and standard language'. *English Today* 21, (January), pp. 3-10.

Rajadurai, J. (2005) 'Revisiting the Concentric Circles: Conceptual and Sociolinguistic Considerations'. *The Asian EFL Journal Quarterly*. (December) 7 (4), pp. 111-130.

Ranta, E. (2006) 'The attractive Progressive - Why Use the -ing Form in English as a Lingua Franca'. *Nordic Journal of English Studies*. 5 (2), pp. 95-116.

Ranta, E. (2009) 'Syntactic Features in Spoken ELF - Learner Language or Spoken Grammar'. in *English as a Lingua Franca Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 84-106.

Reppen, R. and Simpson-Vlach, R. (2020) 'Corpus Linguistics'. in N. Schmitt, N. and Rodgers, M. P. (eds.) *An Introduction to Applied Linguistics* 3rd edition. Abingdon: Routledge, pp. 91-108.

Ribeaud, J. (2013) *Vier Sprachen ein Zerfall*. München: Nagel and Kimche.

Rickenmann, A. (2005) 'Mundart und Kirche'. in F. Helveticum, (ed.) *Dialekt in der (Deutsch)Schweiz – Zwischen lokaler Identität und nationaler Kohäsion*. Menziken: Forum Helveticum, pp. 118-124.

Rogatcheva, S. (2016) *Aspect in learner writing: A corpus-based comparison of advanced Bulgarian and German learners' written English*. Giessen: University of Giessen Dissertation.

Rosenberger, L. (2009) *The Swiss English Hypothesis*. Tübingen: Narr Francke Attempto.

Rubin, D. C. (1975) 'Within word structure in the tip-of-the-tongue phenomenon'. *Journal of Verbal Learning and Verbal Behavior*, 14, pp. 392–397.

Samarin, W. (1987) 'Lingua franca'. in Ammon, U., Dittmar, N. and Mattheier, K. (eds.) *Sociolinguistics: An international handbook of the science of language and society*. Berlin: Walter de Gruyter, pp. 371-374.

SBBK (2021) Schweizerische Berufsbildungsämter-Konferenz (SBBK). Available at: https://www.sbbk.ch/dokumentation/empfehlungen (Accessed 15 May 2021).

Schachter, S., Christenfeld, N., Ravina, B. and Bilous, F. (1991) 'Speech Disfluency and the Structure of Knowledge'. *Journal of Personality and Social Psychology*, 60 (3), pp. 362-367.

Schneider, E. W. (2003) 'The Dynamics of New Englishes: from Identity Construction to Dialect Birth', *Language*, 79 (2) pp. 233-281.

Schoch, B. (2000) *Switzerland - A Modal for Solving Nationality Conflicts?*. PRIF-Report No. 54/2000: Peace Research Institute Frankfurt.

Seidlhofer, B. (2004) 'Research Perspectives on Teaching English as a Lingua Franca'. *Annual Review of Applied Linguistics*, 24, pp. 209-239.

Seidlhofer, B. (2009) 'Orientations in ELF Research: Form and Function'. in *English as a Lingua Franca studies and findings*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 37-59.

Seidlhofer, B. (2010) 'Lingua franca English The European context'. in Kirkpatrick, A. (ed.) *The Routledge Handbook of World Englishes*. Abingdon,: Routledge, pp. 355-371.

Seidlhofer, B. (2011) *Understanding English as a lingua franca*. Oxford: Oxford University Press.

Sewell, A. (2013) 'English as a lingua franca: ontology and ideology'. *ELT Journal*, (January), 67 (1), pp. 3-10.

Shirato, J. and Stapleton, P. (2007) 'Comparing English vocabulary in a spoken learner corpus with a native speaker corpus: Pedagogical implications arising from an empirical study in Japan'. *Language Teaching Research*, 11 (4), pp. 393-412.

Siebenhaar, B. (2006) 'Code choice and code-switching in Swiss-German Internet Relay Chat rooms'. *Journal of Sociolinguistics*, 10 (4), pp. 481–506.

Simpson, R. C., Briggs, S. L., Ovens, J. and Swales, J. M. (2002) *The Michigan Corpus of Academic Spoken English*. Available at: http://micase.elicorpora.info/ (Accessed 19 July 2020).

Sinclair, J. M. (1987) *Looking Up*. London: Collins ELT.

Sinclair, J. (1996) 'Preliminary recommendations on corpus typology', *Technical report, EAGLES (Expert Advisory Group on Language Engineering Standards)*. Available at: www.ilc.cnr.it/EAGLES96/corpustyp/corpustyp.html (Accessed on 27 June 2021).

Sinclair, J. (1991) *Corpus Concordance Collocation*. Oxford: Oxford University Press.

Sinclair, J. (2005) 'Corpus and Text-basic principles'. in M. Wynne, M. (ed.) *Developing Linguistic Corpora: A Guide to Good Practice*. Oxford: Oxbow Books, pp. 1-16.

Sinclair, J. (1997) 'Corpus Evidence in language Description'. in Wichmann, A., Fligelstone, S., McEnery, T. and Knowles, G. (eds.) *Teaching and Language Corpora*. Harlow: Addison Wesley Longman, pp. 27-39.

Singleton, D. (2018) 'What is the best age to start learning an additional language?'. *Babylonia*, 3, pp. 53-56.

Skehan, P. (1996) 'Second language acquisition research and task-based instruction'. in J. Willis, J and Willis, D. (eds.) *Challenge and Change in Language Teaching*. Oxford: Heineman, pp. 17-30.

sprachenunterricht.ch (2010) Available at: http://www.sprachenunterricht.ch/ (Accessed 20 June 2010).

Stadelmann, V. (2019) *Mädchen sind der Schule gegenüber meist positiver eingestellt.* https://www.srf.ch/news/schweiz/getrennter-unterricht-maedchen-sind-der-schule-gegenueber-meist-positiver-eingestellt : SRF Schweizer Radio und Fernsehen.

Stenström, A. B., Andersen, G. and Hasund, I. K. (2002) *Trends in Teenage Talk: Corpus compilation, analysis and findings*. Amsterdam: John Benjamins Publishing Company.

Swan, M. (1997) 'The Influence of the Mother Tongue on Second language Vocabulary Acquisition and Use'. in Schmitt, N. and McCarthy, M. (eds) Vocabulary: description, acquisition and pedagogy. Cambridge: Cambridge University Press, pp. 156-180.

Swan, M. (2016) *Practical English Usage*. 4th ed. Oxford: Oxford University Press.

Swan, M. and Smith, B. (2001) *Learner English: A Teacher's Guide to Interference and Other problems*. 2nd ed. Cambridge: Cambridge University Press.

Swiss Confederation Federal Statistics Office (2011) *The New Census*. Neuchâtel: Federal Department of Home Affairs FDHA.

Swiss Conference of Cantonal Ministers of Education (2007) *Koordination des Fremdsprachenunterrichts auf der Sekundarstufe II Bericht zuhanden der EDK* (22.8.2007), Bern: Swiss Conference of Cantonal Ministers of Education.

Szmrecsanyi, B. (2003) 'Be going to versus will/shall Does Syntax Matter?', *Journal of English Linguistics*, (December), 31 (4), pp. 295-323.

Tagg, C. (2009) *A Corpus Linguistic Study of SMS Text Messaging.* University of Birmingham, UK: Unpublished PhD thesis at the Department of English.

Tagliamonte, S. A. (2006) 'So cool, right? Canadian English entering the 21st century'. *Canadian English in a global context: Theme issue of Canadian Journal of Linguistics*, 51, pp. 309-331.

Tagliamonte, S. A. (2016) *Teen Talk The language of Adolescents*. Cambridge: Cambridge University Press.

Tanaka, K. and Ellis, R. (2003) 'Study-abroad, Language Proficiency, and Learner Beliefs about Language Learning'. *JALT Journal*, (May),1, pp. 63-85.

Tan, P. K. W. and Tan, D. K. H. (2008) 'Attitudes towards non-standard English in Singapore'. *World Englishes*,. 27 (3/4), pp. 465–479.

Tayfoor, S. (2004) *Common mistakes at First Certificate and how to avoid them*. Cambridge: Cambridge University Press.

The Swiss Education Server (2015) Available at: http://www.educa.ch/dyn/67389.asp (Accessed 12 June 2015).

Thewissen, J. (2013) 'Capturing L2 Accuracy Developmental Patterns: Insights From an Error-Tagged EFL Learner Corpus'. *The Modern Language Journal*, 97(S1), pp. 77–101.

Thomas, M. (2006) 'Research synthesis and historiography: The case of assessment of second language proficiency'. in Norris, J. M. and Ortega, L. (eds.) *In Synthesizing Research on Language Learning and Teaching*. Amsterdam: John Benjamins, pp. 279-298.

Thompson, P. (2005) 'Spoken Language Corpora'. in Waynne, M. (ed.) *Developing Linguistic Corpora: A Guide to Good Practice*. Oxford: Oxbow Books, pp. 59-70.

Toolan, M. (1997) 'Recentering English: New English and global'. *English Today* 52, (October), pp. 3-10.

Tottie, G. and Hoffman, S. (2006) 'Tag questions in British and American English'. *Journal of English Linguistics*, 34 (4), pp. 283-311.

Trudgill, P. (1974) *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, P. (2004) *New Dialect Formation: The Inevitability of Colonial Englishes*. Edinburgh: Edinburgh University Press.

Turton, N. D. and Heaton, N.D. (1996) *Longman Dictionary of Common Errors*. London: Longman.

UNESCO, (1953) *The use of vernacular languages in education*. Paris: United Nations Educational, Scientific and Cultural Organization.

Van der Walt, J. L. and Van Rooy, B. (2002) 'Towards a norm in South African Englishes'. *World Englishes*, 21 (1), pp. 113-128.

VOICE (2013) The Vienna-Oxford International Corpus of English (version 2.0 XML) , Director: Seidlhofer, B. ; Researchers: Breiteneder, A., Klimpfinger, T., Majewski, S., Osimk-Teasdale, R., Pitzl, M. L., Radeka, M. Available at: http://www.univie.ac.at/voice (Accessed 15 April 2021).

Wardhaugh, R. (2006) *An Introduction to Sociolinguistics*. 5th ed. Oxford: Blackwell publishing.

Watts, R. (1999) 'The ideology of dialect in Switzerland'. in Bloommaert, J. (ed.) *Language ideological debates*. Berlin: Mouton de Gruyter, pp. 67-103.

Watts, R. J. (2018) 'Early L2 Teaching: Discursive and Factual Arguments'. *Babylonia*, 3, pp. 57-62.

Weinreich, U. (2011) *Languages in Contact*. Amsterdam: John Benjamins Publishing Company.

Werlen, E. (2006) 'Neuland Grundschul-Fremdsprachenunterricht Empirische Ergebnisse der wissenschaftlichen Begleitung Fremdsprache in der Grundschule in Baden- Württemberg'. *Babylonia*, 6, pp. 10-16.

Whinnom, K. (1971) 'Linguistic hybridisation and the 'special case' of pidgins and creoles'. in *Pidginisation and Creolisation of Languages*. Cambridge: Cambridge University Press, pp. 91-115.

Widdowson, H. (1994) 'The Ownership of English'. *TESOL Quarterly*. (Summer) 28 (2), pp. 377-389.

Wierszycka, J. (2013) 'Phrasal Verbs in learner English: a semantic approach. A study based on a POS tagged spoken corpus of learner English'. *Research in Corpus Linguistics*, 1, pp. 81-93.

Woodrow, L. (2006) 'Anxiety and Speaking English as a Second Language'. *RELC Journal*, 37, pp. 308-328.

WrELFA (2015) *The Corpus of Written English as a Lingua Franca in Academic Settings*. Director: Mauranen, A. Compilation manager: Carey, R. http://www.helsinki.fi/elfa (Accessed 23 May 2021).

Zentralschweizer Bildungsserver (2020) Available at: http://rigi.zebis.ch/zebis5.htm (Accessed 19 June 2020).

8. Appendices

Appendix 1 European Common Reference Framework Levels

Global scales[38]

PROFICIENT USER	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
INDEPENDENT USER	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.
BASIC USER	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

---

[38] https://www.coe.int/web/common-european-framework-reference-languages/Table-1-cefr-3.3-common-reference-levels-global-scale (accessed 02.05.2021)

Qualitative aspects of spoken language use[39]

	RANGE	ACCURACY	FLUENCY	INTERACTION	COHERENCE
C2	Shows great flexibility reformulating ideas in differing linguistic forms to convey finer shades of meaning precisely, to give emphasis, to differentiate and to eliminate ambiguity. Also has a good command of idiomatic expressions and colloquialisms	Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).	Can express him/herself spontaneously at length with a natural colloquial flow, avoiding or backtracking around any difficulty so smoothly that the interlocutor is hardly aware of it.	Can interact with ease and skill, picking up and using non-verbal and intonational cues apparently effortlessly. Can interweave his/her contribution into the joint discourse with fully natural turn taking, referencing, allusion making etc.	Can create coherent and cohesive discourse making full and appropriate use of a variety of organisational patterns and a wide range of connectors and other cohesive devices.
C1	Has a good command of a broad range of language allowing him/her to select a formulation to express him/ herself clearly in an appropriate style on a wide range of general, academic, professional or leisure topics without having to restrict what he/she wants to say.	Consistently maintains a high degree of grammatical accuracy; errors are rare, difficult to spot and generally corrected when they do occur.	Can express him/herself fluently and spontaneously, almost effortlessly. Only a conceptually difficult subject can hinder a natural, smooth flow of language.	Can select a suitable phrase from a readily available range of discourse functions to preface his remarks in order to get or to keep the floor and to relate his/her own contributions skilfully to those of other speakers.	Can produce clear, smoothly-flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices.
B2	Has a sufficient	Shows a relatively high degree of	Can produce stretches of	Can initiate discourse, take	Can use a limited number

	RANGE	ACCURACY	FLUENCY	INTERACTION	COHERENCE
	range of language to be able to give clear descriptions, express viewpoints on most general topics, without much con-spicuous searching for words, using some complex sentence forms to do so.	grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes.	language with a fairly even tempo; although he/she can be hesitant as he or she searches for patterns and expressions, there are few noticeably long pauses.	his/her turn when appropriate and end conversation when he / she needs to, though he /she may not always do this elegantly. Can help the discussion along on familiar ground confirming comprehension, inviting others in, etc.	of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some "jumpiness" in a long con-tribution.
B1	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circum-locutions on topics such as family, hobbies and interests, work, travel, and current events.	Uses reasonably accurately a repertoire of frequently used "routines" and patterns asso-ciated with more predictable situations.	Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.	Can initiate, maintain and close simple face-to-face conversa-tion on topics that are familiar or of personal interest. Can repeat back part of what someone has said to confirm mutual understanding.	Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
A2	Uses basic sentence patterns with memorised phrases, groups of a few words and formulae in order to communicate limited information in simple everyday situations.	Uses some simple structures correctly, but still systematically makes basic mistakes.	Can make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident.	Can answer questions and respond to simple statements. Can indicate when he/she is following but is rarely able to understand enough to keep conversation going of his/her own accord.	Can link groups of words with simple connectors like "and", "but" and "because".

	RANGE	**ACCURACY**	**FLUENCY**	**INTERACTION**	**COHERENCE**
A1	Has a very basic repertoire of words and simple phrases related to personal details and particular concrete situations.	Shows only limited control of a few simple grammatical structures and sentence patterns in a memorised repertoire.	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.	Can ask and answer questions about personal details. Can interact in a simple way but communication is totally dependent on repetition, rephrasing and repair.	Can link words or groups of words with very basic linear connectors like "and" or "then".

Appendix 2 Participant Profile

Surname: First names:
Age: Gender: male □   female □

Nationality: Swiss □   other _____
Country: Switzerland
Native language: Swiss German □   other _____
Father's mother tongue: Swiss German □   other _____
Mother's mother tongue: Swiss German □   other _____
Language(s) spoken at home: (if more than one, please give the average % use of each)

Education:
Primary school - medium of instruction: German □ other _____
Secondary school - medium of instruction: German □ other _____

Current studies:
Institution:
Medium of instruction: German □ other _____

Years of English at school:

Stay in an English-speaking country:
Where?
When?   How long?

Other foreign languages in decreasing order of proficiency:

Interviewer:
Relation with subject: Familiar □   Vaguely familiar □   Unfamiliar □

Appendix 3 Participant information sheet

Perceptions of English as a Lingua Franca in the Swiss Context

This study is part of a research project by the Department of English Language and Applied Linguistics in the University of Birmingham.

Description of the study:

I am investigating the spoken English of 18-20 year old Swiss students. I would like to understand more about the English you speak. I will attempt to answer the following questions:

1. How do Swiss students see themselves when they speak English and what unique features do they use?

2. How does learning English in primary school influence Spoken English?

The knowledge gained from this study could be helpful in making decisions about foreign language policy in Central Switzerland.

Invitation:

I would like to invite you to take part in my study. I have chosen you because you are between 18 and 20 years old and either started learning English in the 3rd or 7th grade. Participation is voluntary and participants will not be paid.

I would like to interview you for about 15 minutes. I will be asking you to speak about general subjects that might interest you and a few questions about your views on English. There are no right or wrong answers. I am interested in how you use English to communicate. The interview will be recorded and later transcribed into words which I will investigate with a computer programme.

The data will be treated confidentially and kept securely for a period of 10 years before being destroyed. Your identity will be known only to the researcher. If you have any questions please feel free to ask. You are free to withdraw at any time with no consequences.

Results:
The results of this study will be the basis for my PhD. If you are interested in the outcome of the study or have any further queries please do not hesitate to contact me in the first place or my PhD supervisor Paul Thompson.

Susanne Oswald
PhD Candidate University of Birmingham

Dr Paul Thompson

PhD supervisor
University of Birmingham

Consent form: Perceptions of English as a Lingua Franca in the Swiss Context

<u>Statements of understanding/consent</u>

- I confirm that I have read and understand the participant information leaflet for this study.

- I have had the opportunity to ask questions if necessary and have had these answered satisfactorily.

- I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.

- If I withdraw my data will be removed from the study and will be destroyed.

- I understand that my personal data will be processed for the purposes detailed above, in accordance with the Data Protection Act 1998.

- Based upon the above, I agree to take part in this study.

<u>Name, signature and date</u>

Name of participant………………………   Date……………   Signature………………..

Name of researcher
Susanne Oswald   Date……………   Signature………………..

Appendix 4 Example of transcribed and tagged interview

To demonstrate the transcription and tagging process, an interview was chosen that represents the average participant as close as possible. The following interview is from a Swiss German native speaker who had 10 years of English instruction, spent 3 months in the USA and 3 months in the UK. They consider themselves both a learner and user of English and feel that they will speak with both native and non-native speakers of English in the future. Furthermore, they feel that both grammar and being understood are important and was evaluated as having a B1 proficiency level. The participant used 26 non-standard usages with the corpus average at 27.5 and 2.7 NSU per hundred words with the corpus average at 2.96.

<file sc="2">

<S2>

<A107> first of all I'd like you just to talk for a few minutes about three to five minutes and I have three topics to choose from (uh) one two or three </A107>

<B107> … I choose topic two </B107>

<A107> okay good </A107>

<B107> so I <MW> in in </MW> the USA with my family for four weeks we travelled around with the motorhome it was impressive so and it was really cool and yeah we travelled (uhm) from LA around to San Francisco and down the east coast yeah it was really cool and I <NSU-21> saw also </NSU-21> up there with his little child that was really amazing [L] </B107>

<A107> where did you see that </A107>

<B107> (uhm) in the forest so <NSU-17> we've </NSU-17> walked around and we sought and then we went away because it's really dangerous so </B107>

<A107> (uh huh) </A107>

<B107> yeah and another really impressive moment was when we were in the desert and we saw I don't know how <NSU-16> it </NSU-16> say it we went in the underground and then we saw through different holes the sun came down and </B107>

<A107> like a cave </A107>

<B107> yeah I don't know how I don't <trun>kn</trun> I can't remember the name because it's four years it was four years ago so yeah </B107>

<A107> which state was it in </A107>

<B107> I can't remember [L] yeah </B107>

<A107> it was beautiful though [L] </A107>

<B107> yeah it was really amazing yeah </B107>

<A107> (uh huh) </A107>

<B107> yeah and <cleared throat> another moment was when I got <NSU-16> stang </NSU-16> by a scorpion yeah it wasn't really nice but in the end we found out it's just like a <NSU-16> stang </NSU-16> of <NSU-16> the </NSU-16> bee </B107>

<A107> (uh huh) </A107>

<B107> so it's not poison or poisoning yeah </B107>

<A107> (uh huh) </A107>

<B107> yeah I thought I have to die [L] and my little brother cried and <NSU-20> we </NSU-20> because it was (uhm) in the evening it was ten o'clock and the ranger went away so we were alone in the desert and it was just a station with (uh) with <NSU-16> the </NSU-16> woman and she we <NSU-10> have </NSU-10> to go to the hospital it's really dangerous and so </B107>

<A107> (uh huh) </A107>

<B107> but then she called the ranger and he told her it's just like a <NSU-16> stang </NSU-16> of <NSU-16> the </NSU-16> bee it's not dangerous if she's not allergic </B107>

<A107> (uh huh) </A107>

<B107> to <NSU-15> bee </NSU-15> and yeah so <MW> it was it was </MW> okay [L] </B107>

<A107> so you survived [L] </A107>

<B107> yeah [L] </B107>

<A107> [L] oh so would you like to go back to the United States again </A107>

<B107> yeah I want to go back and <trun>tra</trun> travel around in different states yeah </B107>

<A107> (uh huh) </A107>

<B107> USA is really cool maybe I want to live there a while and come back to Switzerland </B107>

<A107> (uh huh) </A107>

<B107> I don't know [L] </B107>

<A107> would you like to work there </A107>

<B107> yeah maybe </B107>

<A107> (uh huh) </A107>

<B107> yeah but I don't know [L] </B107>

<A107> what did you think of the food there </A107>

<B107> (uhm) we cooked by <NSU-16> ourself </NSU-16> and when I <NSU-10> were </NSU-10> there in the family in Miami I gained like ten kilograms I was like oh my gosh [L] but when I came back I lost it in two weeks so </B107>

<A107> really </A107>

<B107> yeah it was I was just like oh my god I weigh a lot [L] so I ate there just really we ate also good things but it was like yeah the family was really busy so they gave their children just like hamburger and <F> pomme frites </F> and some <trun>da</trun> sometimes also fruits

and salads stuff like that but it was yeah I don't know (uh) it was not that but it was better than in England so [L] </B107>

<A107> okay why what happened in England </A107>

<B107> they can't cook </B107>

<A107> [L] </A107>

<B107> no really it was disgusting oh yeah it was not that good </B107>

<A107> well I think some people can cook in England </A107>

<B107> I don't know it was just I ate <NSU-13> in </NSU-13> the school so they cook for many people and it was like yeah the food didn't look really good <SC> I sometimes we </SC> didn't know what it was so </B107>

<A107> [L] </A107>

<B107> we just ate it (uh) and guessed what it could be so yeah </B107>

<A107> that sounds </A107>

<B107> yeah [L] </B107>

<A107> sounds interesting [L] okay well thank you </A107>

</S2>

<P>

<A107> I'd like to move the second part now and that's where I have pictures here four pictures and I'd like you to look at the pictures and then make up a story about them </A107>

<B107> … (uhm) this girl went to an artist maybe it's I don't know (uhm) yeah this girl went to an artist and wanted a portrait of herself and he painted her but when she saw the painting she was it yeah she thought it's not <SC> that really </SC> nice and then he painted her again and much prettier not as she really looks because yeah maybe <NSU-20> 0 </NSU-20> of her friends and then she showed her friends the picture and they <NSU-10> look </NSU-10> at the

picture and I think yeah I don't know they see that it isn't like <MW> her her </MW> real look I think yeah I don't know [L] </B107>

<A107> thank you that's fine </A107>

</P>

<E>

<A107> and then last part just some questions about English </A107>

<B107> yeah </B107>

<A107> so when you speak English what is more important grammatical correctness or being understood </A107>

<B107> I don't know when I speak English I just speak and don't I think it's <MW> more more </MW> the <NSU-16> understoodness </NSU-16> than the <NSU-16> grammatic </NSU-16> but maybe it's also important to use <MW> the the </MW> right <NSU-16> grammatic </NSU-16> but yeah I don't know </B107>

<A107> and why is it important </A107>

<B107> well because when we use the wrong <NSU-16> grammatic </NSU-16> maybe the other when you speak with someone he thinks (uh) I think she <NSU-1>want </NSU-1>to say in another way so </B107>

<A107> (uh huh) </A107>

<B107> yeah maybe it's yeah maybe he understands what I wanted to say but I had to say it in another way </B107>

<A107> okay </A107>

<B107> yeah </B107>

<A107> and who do you think you'll speak English with after you've completed your studies or in the future with native or non native speakers </A107>

<B107> (uhm) you know I have to use <NSU-13> of </NSU-13> English in the surgery but there we have just one native English speaker so he can understand me I think pretty well </B107>

<A107> (uh huh) </A107>

<B107> and but <MW> I I </MW> don't use English really often (uhm) when I'm home or with my friends I just use it when I travel around the world so there are also native speaker or <NSU-16> no </NSU-16> native speaker if they speak another yeah language </B107>

<A107> (uh huh) </A107>

<B107> and just <NSU-20> 0 </NSU-20> English us to understand and yeah </B107>

<A107> okay so both (uh) </A107>

<B107> yeah I say both </B107>

<A107> and have your English teachers been native or non native English speakers </A107>

<B107> (uhm) <NSU-16> no </NSU-16> native yeah </B107>

<A107> so all Swiss teachers </A107>

<B107> yeah I think so yeah </B107>

<A107> (uh huh) and do you feel or act differently when you're speaking English </A107>

<B107> no I don't think so yeah </B107>

<A107> so </A107>

<B107> it's like it's just like (uh) it's just like German for me because I spoke it really often I was half a year in countries where you speak just English and <MW> it's it's </MW> normal for me <cleared throat> when I came home I spoke some words in English I was just like oh no I want to say in German [L] and yeah </B107>

<A107> [L] okay and last question do you consider yourself a user or a learner of English or both </A107>

<B107> what's a learner of English </B107>

338

<A107> do you consider yourself a learner like you're learning it or that you use it so how do you what do you think </A107>

<B107> I think both yeah I don't know we have the bilingual English here in school so </B107>

<A107> (uh huh) </A107>

<B107> I learn English <MW> every every </MW> lesson new words and I use it also I don't know [L] yeah </B107>

<A107> both okay thank you very much </A107>

</E>

</file>

Appendix 5 Online survey Acceptance of spoken language

Acceptance of spoken language

1. Introduction

This survey is part of a research project by the Department of English Language and Applied Linguistics in the University of Birmingham. I am investigating the spoken English of Swiss students. I have conducted almost 100 interviews and chose 10 sentences from those interviews with a non-standard aspect. I am interested in knowing how acceptable you think the sentences are.
Thank you in advance!
Susanne Oswald

How acceptable do you think the following 10 sentences are when they are spoken in everyday situations?

Please mark them accordingly.

1. I think she want to say it in another way. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

2. In Switzerland the most people can speak English. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

3. Five minutes later the wind came and blew faster and then we go into a museum. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

4. I need English when I'm in holiday. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

5. In the underground trains in London, there are too many peoples for me. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

6. It is important to use the right grammatic. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

7. My friend is in gymnasium and I can speak with her English. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

8. He painted another one and then finally she liked it and showed to the others. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

9. It was the first time I went to another country alone and it was interesting time. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

10. I don't know how many people speaking the language. *

Very acceptable	Acceptable	Somewhat acceptable	Not acceptable
○	○	○	○

11. How old are you? *

- ○ 16
- ○ 17
- ○ 18
- ○ 19
- ○ 20
- ○ Other (please specify):

12. What are you studying? *

- ○ Apprecticeship
- ○ Matura
- ○ Other (please specify):

Finish Survey