Wesleyan University The Honors College

Language Deterioration or Language Revitalization?

Exploring the Effects of ICTs and English Proficiency on Linguistic Diversity

by

Eleonore Jane Finkelstein
Class of 2014

A thesis submitted to the
faculty of Wesleyan University
in partial fulfillment of the requirements for the
Degree of Bachelor of Arts
with Departmental Honors in Sociology

Middletown, Connecticut April, 2014

TABLE OF CONTENT

ACKNOWLEGEMENTS

First of all, I would like to thank my advisor, Daniel Long, who has been my academic advisor, my major advisor and my thesis advisor during my time at Wesleyan. He has always been a great mentor. His knowledge of quantitative research has been extremely helpful throughout this process.

I would also like to thank my wonderful cousin and professor, Estefania Santacreu-Vasut, for whom I worked for as a research assistant last summer at U.C. Berkeley. She initiated me to the field of sociolinguistics, which she applies to her own personal research on economics and languages.

Thank you to my roommates and friends, who have been an incredible source of support during this process.

Finally, thank you to my parents for their constant words of encouragement and for teaching me the importance of my native language and culture.

INTRODUCTION

The study of languages has always been an object of fascination for natural scientists and social scientists alike. Many linguists have exclusively studied the structures and patterns of languages to better understand them purely as communication tools. However, the field of linguistics is broadening and social sciences are being applied to the study of linguistics. Indeed, 'language as a carrier of culture' is the underlining concept of sociolinguistics, which focuses on the specifics of a language within a societal context. Language is a powerful method for diffusing perspectives and values, especially because many of us are largely unaware of its influence on our thoughts (Buttjes 1991). Indeed, in addition to being a means for communication, language acts as a marker of identity. The cultural and social properties of a language contribute to an individual's sense of self. In addition, languages can become a source of ethnic pride. Indeed, a community's cultural cohesion is cemented by a shared language.

However, with globalization and the standardization of the world, cultures are slowly unifying into one global culture, predominantly one that is mirrored on western ideals. In the past decades, one of the principle mechanisms driving globalization has been the revolutionary innovation of Information and Communication Technologies (ICTs). They have increased the frequency of communication and the quantity of recipients of information, transforming the dynamics between people from different cultures. In addition, English is today's Lingua Franca; in other words, it is the main language of global communication. Indeed, English is not only the language of business and trade, but also the language

of the Internet. From the QWERTY keyboard to the US's authority on domain name, English websites are greatly favored, allowing information in English to circulate faster throughout the world. Both ICTs and English as the Lingua France are attributed to the global decline of languages. However, in recent years there has also been a surge in linguistic revitalization movements. Indeed, if language is tied to identity formation, could language communities be holding on to their native-tongues, as a way to resist globalization, ICTs, and English supremacy?

Overview of the thesis

Through my literature review, I will explore the different top-down and bottom-up factors contributing to the changes in linguistic diversity. In particular, I will focus on the effects of ICTs and English proficiency, which are often accredited to the intensification of globalization.

Furthermore, through a cross-sectional quantitative analysis, I plan on testing the effects of ICTs and English proficiency on changes in languages, while controlling for country-level variables. My analyses will use a large sample of countries to identify the general trends in the data. The first chapter will look at the association between Information and Communication Technologies and Linguistic Diversity. The second chapter will compare the differences between the relationship of ICTs with indigenous languages and its relationship with the percentage of immigrant languages. Finally, the third chapter will focus on English proficiency and its association with ICTs, with linguistic diversity, and with indigenous languages. Finally, I will provide concluding remarks on the results of the analyses and broadening the study with a discussion on issues of language rights.

LITERATURE REVIEW

Language decline is a global and pressing issue, resulting in the extinction of approximately twenty-five languages each year (Hagène 2009). Several scholars have argued that languages act as evolving organisms that change as they come into contact with other languages and cultures (Behme 2008; Canagarajah 2007). The dominant language in a country, due to its political, economic, and social status, will usually replace the less dominant languages (Batibo 2009; Brenziger 2009; Hagège 2009). This can happen as a result of societal pressures, experienced by speakers of a lower status language, to switch to a more advantageous one. It can also come from governmental policies, frequently imposing strict language rules to favor the more economically viable language.

Language preservation is the result of a communal effort to maintain a minority language. Marginalized languages are usually discriminated against. However, the response to these prejudices has sometimes been the cause of revitalization movements, elicited by the speakers of these minority languages, as a way to empower their community. Indeed, Wardhaugh (1987) claims that, "in the modern state, language can therefore easily become a symbol of either unity or resistance" (1987:3).

I. Factors that lead to changes in linguistic diversity

Languages are not stable systems and are susceptible to change, decline, and revitalization throughout time. There are a little less than 7,000 languages in the

world today. That's less than half the number languages in the 1500[1]. Some linguists believe that by the end of the 21-century, 50 to 90 percent of the existing languages will have disappeared (Steger, 2009). In this section, I will explore the intricacies behind language decline the world through two different mechanisms— 1) top-down, 2) bottom-up. Before I list these numerous factors, it is important to note that a dominant language is not always a majority language. A dominant language is one that has economic, political, and social supremacy over other languages. It is the language that is used in state bureaucracy and is associated with wealth and upper class. This creates severe inequalities between those who speak the dominant language and those who are at an economic, politic, and social disadvantage by not speaking it. Although in most cases it is also the most spoken language (the majority language), it some cases it is only spoken by a small advantageous elite. Stanley G.M. Ridge's (1996) case study of South Africa clearly illustrates the difference between a dominant and a majority language. Even though the official language of South Africa is English, making it dominant, a large portion of the population does not speak it, which means it is not the majority language.

Top-down factors

When there is so much cultural and linguistic diversity within a country, it becomes difficult to reduce conflict and maintain social cohesion, which generally becomes an issue in a heterogeneous society. Heugh (2013) explains, "diversity offers opportunity, equality and a democratic utopia, yet, it is susceptible to conflict and

[1] There were 14,500 languages in the year 1500 (Steger 2009).

violence" (2013: 9). Furthermore, nations or regions with a high GNI per capita are negatively correlated with language diversity. These countries generally have one or two indigenous languages as opposed to hundreds of different ones. On the other hand, countries with a low economic growth usually have the highest rates of language diversity (Romaine 2009, Batibo 2009). Thus, language diversity is seen as economically restricting.

In order to resolve the issue of language diversity, governments tend to implement language policies that either encourage or discourage the use of a language. This top-down constraint to linguistic choice has led to the disappearance of many languages. Language policies that impose a language serve the purpose of unifying the population, modernizing the country, and having more efficient communication within the country and with the outside world. One way of doing so is by passing an official language act, and, thus, encouraging the use a certain language. This explicitly designates the language that should be spoken by the population of the nation. Some states use implicit language policies, which Herriman and Burnaby (1996) describe as using a certain language in state documents and bureaucracy without officially declaring it as the state language. This latter form of language policy functions similarly to the explicit official language policies because it excludes non-speakers from all governmental participation, pressuring them to switch to the unofficial state language. In some cases, language policies explicitly discourage the use of a language. Vulgarakis and Dawei (2011) have studied the phenomenon of "linguistic genocide", which they describe as the deliberate destruction of a language, as opposed to a natural dying out of a language. This does not involve the actual

killing of language speakers but consists of a strict reinforcement of language policies that prohibits the use of a language. They studied the example of Turkey's language policy, established in the beginning of the 20th century, which banned anyone from speaking Kurdish in public places, educational institutions, and governmental buildings.

Furthermore, Thomas Ricento (2009) and Tollefson (1991) both argue that one of the biggest issues with language planning is that it treats languages as if they were devoid of a sociocultural and historical framework and, as a result, regards them simply as tools for communication. By doing this, the government is not concerned with the ethical or moral implications of preventing a community from speaking their native tongue.

However, according to Herriman and Burnaby (1996),

> problems of miscommunication, as trivial or serious as they may be, are not the focus of a need for language policy except in the broadest sense. It is, instead, where rights, freedoms and power are associated with language that policies become important. (1996: 8)

In other words, it is *because* languages have a certain sociocultural and political meaning that language policies exist. They argue that they are used not to solve inadequate communication between groups, but to designate power and status to a certain language and subculture, and consequently, to repress minority rights. They explain that language "represents the individual's most powerful lien on the group" (Herriman and Burnaby, 1996:10). This means that without the ability to speak their language, the particular social group loses its identity and eventually disappears. For

example, the US has only recently made English its official language, through a language policy act, even though it has always implicitly been the case. This happened not long after issues of immigration rights were brought up. According to Herriman and Burnaby (1996), if English has always been the unofficial state language, the recent official language act was mostly used to further repress minority rights. Similarly, Tollefson (1991) argues that language policies only benefit the dominant language speakers. In fact, he says that those who create the policies are never the ones directly affected by them. They are not the ones who must abandon their native tongue. Thus, language policies can sometimes have exploitative motives that reinforce the power asymmetry between language groups. These types of policies raise the question of language rights as a part of human's individual freedoms.

Another top-down factor that affects language decline are the language policies in education. In the 19th and 20th century, the US government implemented Native American Boarding schools in order to teach Native American children about the norms and values of the country and repress their culture. It also strictly prohibited the use of their native tongue. This example demonstrates the ability for language policies in education to extinguish a language. Additionally, in many countries, the debate on enforcing students to learn English as a Second Language (ESL) is becoming more and more prevalent. Certain scholars believe in the advantage provided by an English education or by a strict implementation of a dominant language in school to unify and standardize communication in a country, which would lead then to economic gain (Miller 2002). Others, like UNESCO, defend the use of mother tongues and promoting culture, and believe that it is an

important factor of a community's identity and self-worth. This debate is affecting the types of educational language policies implemented by governments.

Furthermore, many communities have requested aid from their government to implement language preservation programs in schools or through community centers. Herman M. Batibo (1996) did a case study on the Naro speakers of Botswana, one of the 26 languages spoken in the country. It was only spoken by 9, 000 people and was slowly going extinct until a group of NGOS, as well as the Dutch Reformed Church, founded a center to promote Naro language and culture. This entailed creating a dictionary and a more established written grammar, teaching sustainable forms of income acquisition through their cultural trade of crafts. Through these programs, they instilled pride and self-esteem in the community. The language was no longer associated with its negative stigma as a social and economic barrier. However, these programs that empower language communities require funds, outside aid, and social mobilization, which is not accessible or possible for all minority language communities. Therefore, funding, as well as governmental or non-governmental aid is needed to preserve a language. Unfortunately, in many cases, a government cannot afford to provide these programs.

Finally, Suzanne Romaine (2000) discusses the importance of a bilingual education for the maintenance of an indigenous language. Her case study on Hawaii demonstrates the state's struggle to implement a bilingual educational system, which would teach both English and Hawaiian. Eventually, they received the funding necessary to realize this project. The author explains that often the dominant language speakers will criticize these types of projects as being a waste of money. In addition,

she explains, “the retaining of ethnic identity and lack of integration are seen as their own fault and can be used as legitimation for an unequal division of power and resources” (2000:22). Hence, even if a subculture manages to preserve and revitalize its language, they will still remain in a disadvantageous position because of the false assumption that they did not make an effort to assimilate.

Although language policies are implemented to resolve problems of language diversity, they seem to actually reinforce the political status of a language. The act of choosing an official language imposes a higher status on that language. In addition, the government chooses which languages are worth investing in, what kind of educational programs should exist, and which languages should be preserved. It is important to note that a lot of these decisions are made based on the economic state of the country. This leads me to believe that the wealth of a country is associated with the changes in linguistic diversity.

Bottom-up factors

Some scholars believe that changes in linguistic diversity are attributed to the individual’s themselves. Indeed, the transmission of language to younger generations is crucial in maintaining a language, especially for endangered languages. Dixon (1991) studied the evolution of the endangered Aboriginal language, Dyirbal, over the course of 26 years. In 1963, 500 people spoke the language. The number of speakers rapidly declined, since the younger generations did not learn it and opted for a more commonly used language in the region. By 1989, only 3 people were left speaking this language and they were all born before 1920. Thus, the transmission of a

language to the younger populations is essential in the continuation of a language (Dixon 1991; Hagège 2009).

Additionally, Yamamoto et al. (2008) discuss the role of language in forming identity. They describe language as "personal and at the same time intensely social" (2008: 1). Indeed, in addition to its social and communicative characteristics, it provides the individual with a sense of self and identity. They believe that the blame for the loss of a language is rooted within the speaker. Not transmitting a language to a younger generation is a "personal decision" (2008: 61) and it impacts the rest of the community in a negative way. Consequently, they argue that it is the community's responsibility to maintain their language, driven by the sense of self that the language provides.

However, this assumption that language maintenance lies within individual's agency overlooks the societal pressures and power dynamics established through the top-down factors. In many situations, an individual will feel pressured to switch from a minority language to a dominant one because of the economic, political, or social advantages that it provides. Both Herman M. Batibo (2009) and Matthias Brenziger (2009) have written about poverty and language diversity in Africa. Almost all countries in Africa have a rich linguistic diversity. Numerous and distinct language communities cohabit in the same districts or regions. Generally, there are one or multiple dominant languages. These languages are either the ones imposed by colonizers, such as English or French (Batibo 2009) or they can be a language of trade or religion, such as Swahili in Eastern Africa or Hausa in Nigeria and Niger (Brenzinger 2009). These dominant languages are usually spoken by the upper class

and facilitate communication between different language groups (Batibo 2009). As a result, in order for an individual to obtain social mobility, he or she must learn the dominating language, which provides him or her with the possibility of a better education, job opportunities, greater communication, and a chance to climb up the social ladder (Flammia and Saunders 2007; Batibo 2009; Brezinger 2009). In addition, not knowing the dominant language is socially isolating because you become unaware of the major issues transmitted through the news and government reports on health issues (HIV/AIDS), terrorist attacks, environmental hazard...etc. (Batibo 2009). These studies on the case of poverty and language shifts in African countries exemplify ways in which a language is attributed to an economic, political, and social status and becomes more or less desirable.

Indeed, as the example above demonstrates, languages often have different sociocultural statuses. Speakers of a higher status language are more integrated in society. Claude Hagège (2009) describes the term “inegalitarian bilingualism” as “the situation in which one language exerts formidable pressure upon the other because it holds a much stronger position due to its social status or its widespread national or international use” (2009: 79). Many members of these types of communities will feel pressured to switch to the dominant language and are more willing to teach it to their children.

In order for an endangered language to be maintained or preserved, there needs to be a positive status associated with this language. This process can involve simultaneously using the minority language and the dominant language in different social situations. Each language is used in a different situation to serve a separate

purpose—the mother tongue is used for cultural expression and family situations, whereas, the dominant language is used for commerce, education, governmental documents…etc. This allotment of languages to specific situations is called diglossia and it enables both languages to be equally used. In addition, in order to preserve a minority language, it must be transcribed to a written form, if it isn't already, which allows for a better understanding and sharing of information in that language, as well as providing a historical trace. Indeed, one of the most important factors for language preservation or revitalization is for the community to have a positive relationship with its language. This means that if an individual has an emotional connection with a language, due to the cultural heritage it represents for them, then they will be more willing to speak it and to teach it to their children (Auburger 1990). Accordingly, if a language has a negative status, then an individual will be more inclined to abandon it for a more beneficial one. The decision to abandon a language in favor of another is a bottom-up mechanism. But this choice is based on the value assigned to a language, which is done through a top-down application of power. Thus, these top-down and bottom-up factors are products of each other and sustain the power dynamics between linguistic communities within a country.

II. Globalization: a major factor in language decline

Globalization has a powerful altering effect on many different aspects of society, including wealth and poverty, education, health, religion, and cultural values and perspectives (Heugh 2013). Globalization has deepened the workings of the

previously discussed economic, political, and social inequalities that lead to language decline.

Definition

It is hard to define when globalization actually started. Some believe it started during the Industrial Revolution, others say it is a much more recent phenomenon and it developed during the post-industrial era. Some argue that it does not have a specific time frame and that it has been slowly developing during the past five centuries Despite its indefinite time frame, many scholars have argued that it has been more noticeable in the past recent decades due to new technologies (Steger 2009). Indeed, it has been affecting the world on an economic, political, and social scale.

Steger (2009) defines the global economy as "the intensification and stretching of economic interrelations across the globe" (Steger 2009: 38). With the development of global economic institutions, such as the International Monetary Fund (IMF), the World Bank, and the World Trade Organization (WTO), the world's economies have been fused into a new global economy. It is controlled by central organizations, based in the western world, and specifically in the United States, which has been the principal power of the IMF and the World Bank. In addition, among the 200 biggest Transnational Corporations (TNC), which produce half of the world's industrial output, all of them have their headquarters in North America, Europe, Japan, and South Korea (Steger 2009). The geographical congregation of global economic institutions and TNCs reflects a centralization of power among the richest and most powerful nations. The global economy has reinforced and widened the existing asymmetry of power between the richest nations and the rest of the world.

Steger characterizes the global economy as the "engine behind its [globalization] rapid development" (Steger 2009: 57).

Political globalization is defined as the "intensification and expansion of political interrelations across the globe" (Steger 2009: 58). Intergovernmental institutions, such as the United Nation or the European Union, are increasingly overpowering nation-sovereignty, giving rise to a new political framework of global governance that disregards national boundaries. Steger explains that the nation-state framework is rooted in the psychological attachment of people to a nationality and culture that unites them under a single leadership. However, with globalization and the massive flow of people, capital, and technologies, these cultural boundaries, as well as state sovereignty are weakened.

Finally, as the global market continues to grow and as an increasing number of intergovernmental decisions are made, countries are now deeply intertwined and dependent on one another. It is important to state that these global political and economic decisions are based on western ideals of democracy and capitalism. A growing number of 'non-western' states are pressured to adopt these political and economic structures. Thus, western cultural values are also spreading alongside of political decisions and global markets. Steger defines culture as "the symbolic construction, articulation, and dissemination of meaning" (Steger 2009: 71). In other words, he means any form of emblematic expression, including music, language, and visual art. Some of the main factors that have enabled culture to circulate so rapidly and freely between countries are modern technologies of communication. The growing numbers of ideas and images that are exchanged through Information and

Communication Technologies have significantly impacted the way people think and behave.

There are several different perspectives on globalization, ranging from positive and encouraging sentiments to the belief that it is threatening and detrimental to society. Those who are pro-globalization, such as neo-liberals, believe it will unify the citizens of the world. They believe that working together, as opposed to being divided into nations, contributes to human progress. Others, such as neo-Marxists, believe globalization is spreading and enforcing a capitalist structure to all economies (Fairclough 2006). Regardless of the various opinions, one thing is certain—globalization is inevitable. With the constant development of modes of transportation and of information and communication technologies, information, commodities, and people will continue to reach all parts of the world. Along with this, comes the inevitable sharing and alteration of cultures, perspectives, and values.

Economic, political, and cultural globalization is driving the world towards a homogenous and borderless world. It is also imposing western ideals and setting a standard of modernization. Today, globalization is also largely attributed to Information and Communication Technologies.

Western Models of modernization

Many developing countries are aspiring towards the western model of modernity. In order to achieve this model, they must not only adopt new technological innovations, which require learning English, but they must also establish a standard language within the country. Modernization theory is based on

the assumption that ‘underdeveloped’ states must adopt the same economic and political models as those of ‘industrialized’ countries. Learning ESL (English as a Second Language) is encouraged in modernization theory for two reasons; 1) English is considered to be an important tool in modernization, 2) monolingualism (ideally in English) is regarded as advantageous for modernization, while multilingualism is associated with ‘traditional’ and ‘underdeveloped’ countries (Tollefson 1991). According to Tollefson (1991), “in Third World countries in Asia and Africa, English is seen as an essential tool for importing Western technologies and building economic ties with Europe and North America” (1991: 80). Thus, English plays an important role in the “modernization” of countries.

Japan is a good example of a country that implemented a language policy to become a monolingual state. Although this transformation happened over the course of several decades, many years before the invention of the Internet, Japan’s determination to mimic the western world led to the implementation of a single national language, a concept developed during the rise of nationalism in Europe. This removed all language barriers among Japanese people and, therefore, led to better state communication and governmental presence (Heinrich 2012).

Tollefson (1991) criticizes modernization theory and the spread of English by associating it with a form of “inequality and exploitation” (1991: 83). He refers to it as being ideological. Many critics of modernization theory have argued that if countries are perceived as ‘underdeveloped’, it is the consequence of others being labeled as ‘developed’. Thus, they argue that the spread of English is not just a tool of modernization but it is reinforcing the existing inequalities between nations by

imposing the supremacy of English-speaking countries. Furthermore, the spread of English strengthens the inequalities within the 'developing' country. Most post-colonial societies struggle with two sets of opposing institutions; one that is 'westernized' or 'developed, and the other that is 'traditional' or 'underdeveloped'. This dichotomy has resulted in a geographic, economic, and linguistic division of the population. Thus, the expansion of English or of the dominant language is uneven and only reaches those who live in urban and wealthier areas, fortifying and increasing the gap between the two institutions (Tollefson 1991).

Therefore, through this logic, globalization is in part the domination of 'western' culture on other nations and the establishment of a power dynamic. Through this perspective of globalization, we can understand how the language hierarchy is maintained through a constant reinforcement of power relations between groups within a society and between different countries. In fact, this has led to the universal assumption that monolingual policies leads to economic development, and that English is the Lingua Franca, as well as the language of modernization. Finally, these western ideals are spreading faster than they used to due to Information and Communication Technologies. Indeed, Steger (2009) argues that modern technologies, and specifically ICTs, are one the "hallmarks of contemporary globalization" (2009: 38).

1. English as the Lingua Franca

Tollefson (1991), among others (Sonntag 2003), attributes English's status as the Lingua Franca to the "economic and military power of English-speaking countries

and the expansion of the integrated global economic market which they have dominated" (1991: 82). Indeed, the spread of the English language started with the British Empire's expansion and colonization of a significant portion of the world, and then followed by the US's status as a superpower after WWII. Today, English is the universal language of business, trade, commerce, science, and technological innovations. Many multinational companies have made English the sole language of their business, whether or not the headquarters are in an English speaking country. Wardhaugh (1987) describes the language's ability to be used in all contexts and by all people around the world. He says, "Spoken almost everywhere in the world to some degree, and tied to no particular social, political, economic, or religious system, nor to a specific racial or cultural group, English belongs to everyone or to no one, or it at least is quite often regarded as having this property." (1987: 15)

Indeed, there are more non-native English speakers than there are native English speakers today (House 2003). Furthermore, Buttjes (1991) argues that we subliminally internalize cultural and social processes that are engrained in our native tongue. However, English is often a person's second language. In this case, English is decontextualized from its culture (Buttjes 1991). It has taken a specific role as a language for communication only, as opposed to a language for identification (House 2003). Buttjes (1991) explains, "most of us are even less aware of the intricate and subliminal processes by which all of us are socialized into our native first cultures" (1991: 3).

Information and Communication Technologies as driving forces of globalization

Information and Communication Technologies (ICTs) are some of the key actors behind the intensification of globalization in recent years (Steger 2009; House 2003; Heugh 2013). The Internet, for example, has completely revolutionized social interactions and methods of communication. It has enabled every person, with access to it, to share and absorb information from around the world, significantly changing cultural perspectives, political dynamics, and economic transactions. However, in order for everyone to participate in this pool of knowledge, there needs to be a universal language.

Tomasz Kamusella (2012) explains that in order for this global network to exist, there needed to be a "uniformization" of the technological language and of the units and measurements used in machinery. With the development and the commercialization of computers and the Internet, came the establishment of technological standards through the Unicode Consortium. The Unicode Consortium was founded in California in 1991. It is responsible for encoding characters of written languages for computer use. In 2011, it had encoded 93 scripts (Kamusella 2012). Having said that, the computer and the Internet were originally developed for the U.S. military and for American scientific researchers to exchange information. Thus, its standards for encoding are created with the English language in mind. Most character-based languages, such as Chinese, which has existed for centuries, are now changing their script from word characters to phonetic characters to fit the model of the keyboard (Crystal 2001). These very drastic changes in languages prove how indispensible it is to adapt to these modern technologies.

Furthermore, the American non-profit organization, Internet Corporation for Assigned Names and Numbers (ICANN), controls all domain names on the Web. More than 80 percent of websites on the Internet are in English (Steger 2009). Some countries with languages that do not use Roman letters, such as Chinese or Arabic, have complained that domain names in their native languages are being processed too slowly and hindering the spread of information in their languages and countries (Flammia and Saunders 2007). Thus, languages that are similar to English have a clear advantage. Many languages are underrepresented on the web—either because those who speak it do not have access to as many computers or because certain policies restrain the use of the language. For example, Google excludes websites written in 35 languages (Mikami et al. 2005).

Indeed, Rehm, Uszkoreit, and SpringerLink (2013) discuss the ways a language needs to be able to survive the digital world in order to survive in general. They describe the new technologies that are developing and threatening the existence of several languages, including German, a widely spoken language. Due to the U.S.' supremacy in the business world and their development of the Internet, English has become the dominant language of the web. Even as Chinese is becoming the most spoken language in the world, English still remains the most used language on the Internet (Flammia and Saunders 2007). Therefore, in order to participate in the global network of information, one must either understand English or any other dominant language of the Internet.

Contextualization and de-contextualization of languages

According to many scholars, globalization, ICTs, and the English language are leading to the uniformity of cultures. However, languages are social, allowing the speaker to be part of a group and form an identity. They also reflect the culture and history of a particular society through idioms and changes in structure and language (Hegège 2009). Thus, languages are contextualized in a unique and rich framework.

The contextualization of language is also described as the act of identifying with a certain social group and being labeled with certain social features, based off your language (Gumperz 1992). Your linguistic choices reflect a conscious decision to affiliate yourself to a certain social group and therefore to distinguish yourself from another.

Furthermore, Information and Communication Technologies has further increased the de-contextualization of languages. As it becomes the Lingua Franca, English is no longer tied to a culture. People are learning English for communication purposes, not necessarily as a carrier of culture. ICT's have further accentuated this de-contextualization phenomenon. However, this has also exposed the ways in which a language can serve as a cultural identifier. It is in fact *because* certain languages are used only for communication, that others become languages for identification. Thus the question here is whether globalization, driven by ICTs and English hegemony, lead to linguistic decline or, whether, it leads to a stronger connection between people and their native languages, which become emblems of their identity.

The following literature review reveals that there are many factors that lead to changes in a country's linguistic diversity. There are economic, political, and social

factors that act through mechanism that are either top-down, bottom-up, or a combination of both. Globalization reinforces these mechanisms by deepening the power relations between dominant languages and minority languages. It also establishes a western supremacy and encourages a model of modernization. This entails implementing language policies for countries to become monolingual and/or encouraging the acquisition of English, thus leading to language decline. Furthermore, it seems as if a country that doesn't meet the standards of development, which are founded on Western ideals, will remain dominated politically, socially, and culturally by the western world. Thus, all of these pressures, that are created in part by western supremacy onto countries and imposed through globalization processes, can lead to language decline.

As the main accelerators of globalization today, Information and Communication Technologies have heightened communication and shrunk the world into one global network. Since languages are primarily tools of communication, there seems to be an increased pressure for people to learn the languages that are the most prominently used in the cyber world. Indeed, in order to keep up with the fast and continuous spread of information through out the world, it has become necessary to learn English, today's lingua franca (language of commerce). However, many scholars have also noticed language revitalization movements, in order to counter act the linguistic homogeneity encouraged by ICTs and globalization. Although certain case studies illustrate the ways in which ICTs affect languages, it is still unclear whether or not they pose a threat to linguistic diversity or if, in fact, they are strengthening the notion of language as a cultural and identity marker and

maintaining linguistic diversity. Through a quantitative analysis, I want to understand if and how ICTs affects linguistic diversity around the world, while controlling for other possible factors. Are high levels of ICT rates in a country associated with lower linguistic diversity?

METHODOLOGY

I. Data Sources

The data that I used for my analyses came from three main sources; The World Bank, Ethnologue: Languages of the World, and Education First. The first source, the World Bank, is an international financial institution with the goal of alleviating poverty. It provides data on development in countries around the world and is often used by policy makers to make more informed decisions and evaluate improvement. It also helps researchers and scholars have a better understanding of global development issues.

The second source of information is Ethnologue: Languages of the World. Its goal is to catalogue languages from all around the world with the contribution of hundreds of linguists and researchers. This project is coordinated and published by SIL International, a non-profit organization that helps language communities around the globe by implementing language sustainability programs. Ethnologue has been recording information on languages since 1951. It categorizes languages by continent and country, and provides a wide range of information for each one. It also provides a Linguistic Diversity Index (LDI), which measures the level of language diversity in each country.

The third source is Education First. This private company, founded by Bertil Hult in 1965, offers a variety of international and educational programs, ranging from language teaching programs to cultural exchanges. Its mission is to transcend cultural, linguistic, and geographical barriers in order to help people become citizens of the world. It operates 500 schools and offices in 52 countries around the globe.

Additionally, Education First has developed the English Proficiency Index (EPI), which measures the overall level of English dexterity in 60 countries.

II. Concepts and Variables

Linguistic Diversity

One of my main concepts for these regression analyses is linguistic diversity. Linguistic diversity represents the amount of living languages in a given country. In order to measure this concept, I used the Linguistic Diversity Index (LDI), provided by Ethnologue: Languages of the world. This index, created by Joseph H. Greenberg, quantifies the level of language diversity in the country, on scale of 0 to 1. Complete linguistic diversity, represented by the highest value, 1, indicates that no two people have the same mother tongue, whereas the lowest possible value, 0, indicates a complete lack of diversity, such that everyone has the same mother tongue. This index is available for 210 countries, during three different points in time, which coincides with the publication of the last three editions of Ethnologue. These editions include Edition 15 (published in 2005), Edition 16 (published in 2009), and Edition 17 (published in 2013). The older editions of Ethnologue do not offer data on linguistic diversity. The sample for this variable consists of a total of 609 observations.

Concentration of indigenous languages and of immigrant languages

My next concept is the concentration of indigenous languages spoken in a country. An indigenous language is a language that is native to the region. However,

in certain countries, for example in France, the dominant language is also the indigenous language of the country. These types of indigenous languages are stable enough that they are susceptible to external factors. Thus, by looking at the concentration of indigenous languages in a country, as opposed to the number of speakers of an indigenous language, I tested the effects of the independent variables on smaller indigenous languages. Indeed, if a country has a high concentration of indigenous languages, this indicates that these languages divided among a population and are not majority languages. Using data, from Ethnologue, on the number of indigenous languages in each country, I created a variable that represents the percentage of the total languages spoken in a country that are indigenous. The percentage of immigrant languages was created the same way, using the number of immigrant languages and the total number of languages spoken in each country. The percentages of indigenous languages and of immigrant languages are equal to the total number of languages in the country. There 615 observation for each of these variables.

ICT consumption

In order to assess the level of ICT consumption by country, I used three variables from the World Development Index database provided by the World Bank that measure the usage and demand of these technologies by country. I choose to look at the rate of cell phone and Internet usage by country, since these two technologies are the most prominent Information and Communication Technologies in our world today. The variable on the rate of ICT imports will demonstrate the demand of these

technologies in each country. The description of each independent variable is featured below:

- Internet Users (per 100 people): The World Bank characterizes an Internet user as someone who has access to the Internet. There are 593 observations.
- Mobile Users (per 100 people): this variable looks at mobile subscription for both prepaid and post paid cellular subscriptions. The sample size consists of 595 observations.
- ICT goods import (%): This variable looks at the percentage of total Information and Technology goods imported. These ICT goods include telecommunications, audio and video, computer and related equipment; electronic components; and other information and communication technology goods. There are 439 observations for this variable.

English Proficiency

In order to measure the English proficiency in a country, I used the English Proficiency Index (EPI) created by Education First. It measures the average adult's English capacity. It was created by computing data from two Education First English tests, taken by over 750, 000 people (Education First Third Report).

Control country-level variables

The country-level variables that I controlled for are all from the World Development Index database provided by the World Bank. The description of each control variable is featured below:

- GNI per capita: The Gross National Income is converted into international dollars using Purchasing Power Parity (PPP). I divided this variable by 10, 000. There are 528 observations.
- Urban population (%): This measures the percentage of the total population living in urban areas. There are 618 observations.
- Continent dummy variable: Each country is assigned a continent. If the country is part of this continent, then the dummy variable will be equal to 1, but 0 if it is not part of it. Each category is mutually exclusive; therefore no country was assigned two continents. The continent dummy variables are Europe, Asia, Africa, North America[2], South America, and Oceania. These variables will be used to measure whether or not the geographical location of the country is driving the results. There are a total of 630 observations for these categorical dummies.

III. Method and Hypotheses

My analyses are divided into three chapters. In each chapter, I tested three hypotheses using cross-sectional regression analyses. Each observation represents a country-year combination. The three points in time represented are 2005, 2009, and 2013. Through these different hypotheses and analyses, I explored the different intricacies of language changes caused by ICTs, English proficiency, and several country-level factors.

[2] Countries that lie in Central America were placed in the North America category.

Chapter 1—ICTs and Linguistic Diversity

My first sets of analyses consisted in a series of cross-sectional regression analyses testing the association between the three ICT variables (the independent variable) and linguistic diversity (the dependent variable) for each country. Each country-year observation is matched with data from the World Bank on Information and Communication Technologies, as well as on urban population rates and on GNI. Finally, each country is assigned a categorical continent dummy variable. The sample consists of 518 observations in Table 1.1 and 401 in Table 1.2.

Before I tested the assocation of Information and Communication Technologies and linguistic diversity, I tested the effects of other country-level characteristics that could also be affecting linguistic diversity. The hypothesis that I tested are the following:

- Hypothesis 1.1: The wealth of a country is correlated with its level of linguistic diversity. More specifically, low-income countries have higher levels of linguistic diversity, while high-income countries have less linguistic diversity. (Table 1.1)
- Hypothesis 1.2: The proportion of the population that lives in urban areas is correlated with its level of linguistic diversity, in such a way that countries with large urban populations have low linguistic diversity and vice versa. In other words, urban areas are detrimental to linguistic diversity, while non-urban areas foster and maintain linguistic diversity. (Table 1.1)
- Hypothesis 1.3: The three ICT variables, Internet use, mobile use, and ICT imports, are all negatively associated with linguistic diversity. (Table 1.2)

Chapter 2— ICTs and the level of indigenous and immigrant languages

After testing the relationship of the different ICTs on linguistic diversity, I explored the effects of these same independent variables on the country's levels of indigenous and immigrant languages.

The next series of analyses consists of a comparison of two sets of regression analyses, each using identical models, but tested with two different dependent variables. These dependent variables are the percentage of indigenous languages (results in Table 2.1) and the percentage of immigrant languages in a country (results in Table 2.2). The independent variables will be the three ICT variables, as well as the country-level variables, as controls. The hypotheses are the following:

- Hypothesis 2.1: The wealth of a country has a different relationship with its level of indigenous languages than with its level of immigrant languages. More specifically, higher income level countries have lower percentages of indigenous languages and higher levels of immigrant languages spoken, while lower income countries have more indigenous languages and less immigrant languages.
- Hypothesis 2.2: The level of urban population of a country has a different relationship with its percentage of indigenous languages than with its percentage of immigrant languages. In countries with greater urban areas, there are more immigrant languages spoken than indigenous ones, while in countries with less urban areas, there are more indigenous languages spoken than immigrant ones.

- Hypothesis 2.3: The country's level of ICT consumption has a different relationship with its percentage of indigenous languages than with its percentage of immigrant languages. Greater ICT consumption is detrimental to the level of indigenous languages.

Chapter 3—English proficiency, linguistic diversity and indigenous languages

The issue of English hegemony is often cited when discussing changes in linguistic diversity. Using the English Proficiency Index (EPI), I explored its correlation with linguistic diversity, while controlling for the effects of Internet use. Then, I explored the EPI's effect on the concentration of indigenous languages in a country, since small indigenous languages are generally more vulnerable to the effects of the English language. Before I tested these two hypotheses, I wanted to verify the association between Internet user rate and English proficiency, which is frequently referenced by scholars. The hypotheses are the following:

- Hypothesis 3.1: English is the dominant language on the Internet. Thus, countries with high levels of English proficiency also have high levels of Internet presence. (Table 3.1)
- Hypothesis 3.2: English proficiency is correlated with lower linguistic diversity. (Table 3.2)
- Hypothesis 3.3: English proficiency is correlated with lower concentrations of Indigenous languages. (Table 3.3)

IV. Limitations

Due to the nature of the study, it is difficult to control for issues of spuriousness. There are many other cultural, historical, and political factors that could be leading to changes in linguistic diversity. For example, the effect of language policies, mentioned in the literature, cannot be accounted for, even though they have an important influence on linguistic diversity.

Furthermore, the data available on languages is very limited. Ethnologue is the only database on languages in the world and it still contains a lot of missing values. This missing observation might not be random and could be indicating a bias towards countries with stronger institutions and accessibility, allowing for greater amounts of statistical research. Thus, it is possible that some of the countries with the most linguistic diversity have been excluded from the data.

Due to the limited data on linguistic diversity, it was difficult to assess the trends in decline over the course of 8 years (from 2005 to 2013). Thus, I opted for a cross-sectional analysis, which does not infer causality but correlation. Based on the findings, I illustrated the ways in which there could be causal effect between the variables through various case studies.

CHAPTER 1- ICTS and Linguistic Diversity

Before I tested the relationship between Information and Communication Technologies (ICTs) and the Index of Linguistic Diversity for each country, I decided to test the validity of other factors that could be associated with different levels of linguistic diversity. These included the country's GNI per capita, its urban population rate, its geographical location, and the percentage of its total languages that are immigrant or indigenous. By doing this, I was able to test the validity of other factors that could be leading to linguistic decline.

The dependent variable for the models in Table 1.1 and Table 1.2 is the Linguistic Diversity Index (LDI). There many other factors that could account for trends in languages, but these results shed a light on the particular factors that are often associated with a country's Linguistic Diversity Index.

I. GNI per Capita and Linguistic Diversity

Hypothesis 1.1

Many of the top-down and bottom-up factors that lead to language decline can be attributed to the wealth of the country. Top-down influences on language decline, such as governmental language policies or educational language policies, are often motivated by political and social factors (Herriman and Burnaby 1996; Tollesfson 1991), but they are also implemented for economical purposes. In fact, the assumption that linguistic homogeneity leads to economic gain is one of the underlining reasoning behind the application of governmental language policies that

restrict or encourage the use of one or two languages (Ricento, 2009). This is because monolingualism is often equated with economic growth (Miller 2002, Romain 2009), while poverty is associated with multilingualism (Batibo 2009; Brezinger 2009). In addition, certain speakers of economically disadvantageous languages will teach their children a more dominant and financially beneficial language, in the hopes of social mobility (Flammia and Saunders, 2007; Batibo 2009; Brezinger 2009; Hagège 2009). Therefore, the first hypothesis is that a country's GNI per capita is negatively associated with language diversity.

Models (Table 1.1)

Model 1: Y (LDI) = X0 + X1 (GNI) + e

Model 4: Y (LDI) = X0 + X1 (GNI) + X2 (Indigenous) + e

Model 5: Y (LDI) = X0 + X1 (GNI) + X2 (Immigrant) + e

Model 7: Y (LDI) = X0 + X1 (GNI) + X2 (Urban pop) + e

Model 9: Y (LDI) = X0 + X1 (GNI) + X2 (Europe) + X3 (North America) + X4 (South America) + X5 (Asia) + X6 (Oceania) + e

Model 12: Y (LDI) = X0 + X1 (GNI) + X2 (Europe) + X3 (North America) + X4 (South America) + X5 (Asia) + X6 (Oceania) + X7 (Indigenous) + X8 (Urban pop) + e

Results

First, let's look at the bivariate correlation between GNI per capita and linguistic diversity in Model 1. GNI per capita is statistically significant with an

adjusted r-squared of .072, which means that 7.2% of the variance in linguistic diversity can be explained by the country's economic level. Furthermore, GNI per capita is negatively correlated with linguistic diversity. This suggests that in countries with a high GNI per capita, there are lower levels of language diversity, which confirms hypothesis 1.1. However, I controlled for other factors to ensure the validity of this association.

In Model 4, I added the variable on the percentage of indigenous languages in a country. GNI per capita continues to be as statistically significant and the adjusted r-squared for this model has increased by 8.3%. The percent of indigenous languages is also statistically significant, which suggests that it accounts for some of the variance in language diversity. In fact, the magnitude of the GNI coefficient has decreased by 13.6%. Furthermore, the *Indigenous* variable's coefficient is positive. Thus, the high level of indigenous languages in a country leads to more linguistic diversity. To better understand what this means, this Model was compared to Model 5, where I tested the association between GNI per capita and linguistic diversity, along with the percentage of immigrant languages in a country. In Model 5, the level of indigenous languages is negatively associated with linguistic diversity. Therefore, countries with high rates of immigrant languages have less total linguistic diversity than countries with higher rates of immigrant languages. Thus, there seems to be a higher concentration of indigenous languages in places with higher linguistic diversity. This reveals that these languages are probably smaller in number of speakers.

In Model 7, I controlled for urban population rates. Urban population rate is statistically significant, while GNI is insignificant. The r-squared has increased by 39% compared to the original bivariate model, indicating that this model is a better fit. These changes reveal that, when controlling for urban population rate, the income level of the country is spurious, while the percentage of urban population is correlated with linguistic diversity. The latter relationship will be further tested in the next hypothesis section (Hypothesis 1.2).

In Model 9, I controlled for the geographical location of the country. I used the continent dummy variables to test the relationship between continent belonging and the level of linguistic diversity. Each country was assigned one continent. The *Africa* variable, which has the highest level of Linguistic Diversity Index, was omitted from the models and acts as the reference group. All five continent dummy variables are statistically significant, signifying that there is some sort of trend in linguistic diversity for each continent, as a result of geographical, cultural, and historical factors. The adjusted r-squared (r2_a=.26) has significantly increased since Model 1. The magnitudes of Europe and North America's negative coefficients differ much more from the *Africa* dummy variable, than the magnitudes of Asia or Oceania's coefficients. In other words, if a country is part of Europe or North America, then there is less linguistic diversity than in Asia or Oceania, when compared with Africa's high level of linguistic diversity. In addition, GNI is statistically significant, but its coefficient has decreased by 58% compared to Model 1, which means that the continent to which the country belongs to accounts for a greater amount in the variance in linguistic diversity than GNI per capita does.

The overall best-fit model is Model 12, which is also the full model with an r-squared of .27. Compared to the initial bivariate model, where I was only testing GNI per capita, this model's r-squared has increase by 237.5%. The GNI per capita variable and the percent of indigenous language variables are no longer significant. However, the urban population rate and all the continent dummy variables are still statistically significant.

Through the assessment of these models, I can conclude that GNI is spurious. It had an initial association in the bivariate model (Model 1) and when controlling for the concentration of indigenous languages (Model 4). However, this association disappeared when I controlled for urban population. Indeed, when it is tested with urban population in Model 7, its coefficient is more than halved, suggesting that urban population accounts for 58% of the GNI variance. The continent variables also account for a major part of the variance in linguistic diversity. Thus the two main factors acting on linguistic diversity are the rate of urban population and the geographical location in this set of models.

Discussion

When testing for first hypothesis, I have reached three major observations. The first is that the income level of the country is not a factor that is associated with linguistic diversity, which contradicts hypothesis 1.1. In fact, if GNI was statistically significant in the bivariate model, it is probably because it was a proxy for urban population rate. This leads me to my second observation, which is that countries with high levels of urban population have lower linguistic diversity. This relationship will

be tested and discussed in more depth in the next section. The third observation is that linguistic diversity differs by continent. Therefore, each continent has a general trend in linguistic diversity, which can be explained by the similarities in the cultural and historical backgrounds of countries pertaining to the same continent.

II. Urban population rates and linguistic diversity

Hypothesis 1.2

According to certain scholars, languages are constantly evolving, especially as they come into contact with other languages and cultures (Behme 2008, Canagarajah 2007). In urban areas, languages are much more susceptible to come into contact with different cultures and languages. Therefore, I propose a second hypothesis based on the literature and on the results from the first hypothesis; urban population rates are negatively associated with language diversity.

Models (Table 1.1)

Model 6: Y (LDI)= X0 + X1 (Urban pop) + e

Model 7: Y (LDI)= X0 + X1 (Urban pop)+ X2 (GNI) + e

Model 10: Y (LDI)=X0 + X1 (GNI) + X2 (Urban pop) + X3 (Indigenous) + e

Model 11: Y (LDI)=X0 + X1 (GNI) + X2 (Urban pop) + X3 (Immigrant) + e

Model 12: Y (LDI)= X0 + X1 (GNI) + X2 (Europe) + X3 (North America) + X4 (South America) + X5 (Asia) + X6 (Oceania) + X7 (Indigenous) + X8 (Urban pop) + e

Results

The bivariate relationship between the rate of urban population and language diversity, in Model 6, shows that urban population is statistically significant with an r-squared of .10. The percent of urban population in a country has a negative relationship with language diversity, which supports my second hypothesis. Thus, simply looking at the bivariate model, it can be deduced that in countries with a larger urban population, there is less language diversity. But what happens when we control for other variables?

As already discusses in the previous section on hypothesis 1.1, in Model 7, when urban population is tested alongside GNI per capita, it remains very statistically significant, while GNI is not significant. It is clear that urban population is a much better indicator of linguistic diversity and that when GNI is significant it is because it is a proxy for urban population.

In Model 10, I controlled for the concentration of indigenous languages. The variable for urban population rate continues to be very significant, with a slight decrease of 9.5% in its coefficient. The variable *Indigenous* is not significant and the r-squared for this model has increased only by 2%. The same can be said about the percentage of immigrant languages in Model 11.

In Model 10, I added the continent variables to the bivariate model. The r-squared has increased by 177%. All of the variables are statistically significant and the magnitude of the urban population coefficient has decreased by 52%. Therefore, the location of the country accounts for a large portion of the variance in linguistic diversity.

In Model 12, the full model, all the variables are tested. The rate of urban population is statistically significant, but its magnitude has decreased to less than half of what it was in the bivariate model. The adjusted r-squared, however, has increased and is more than twice as high as it was in the bivariate model. The fact that urban population has remained significant does indicate that it is still associated with linguistic diversity. The GNI per capita variable and the percentages of indigenous/ immigrant languages in the country are no longer significant. The continent dummy variables all remain statistically significant, each experiencing a decrease in their coefficients.

The results of these models confirm that GNI per capita is spurious, when tested with urban population, which repudiates hypothesis 1.1. In addition, urban population has a negative association with linguistic diversity, confirming hypothesis 1.2.

Discussion

Both the results of the first study and the second one have demonstrated that GNI per capita acts as a proxy variable for urban population rates. There are several explanations for the spuriousness of GNI per capita. First of all, these two variables can easily be linked to one another. Countries with higher economic levels are more likely to develop significant urban areas (Quigley 1998). In addition, urban areas attract revenue. They are often intertwined with one another and grow at the same speed. The results show that it is not the country's income that is leading to lower

levels of linguistic diversity, but it has some link with the percentage of the population that is living in urban areas.

Second of all, the top-down and bottom-up factors that link a country's wealth to language decline can also be explained by the proximity and contact of linguistic communities, which is enhanced in urban areas. Many scholars have shown that in wealthier countries, there are fewer languages spoken (Batibo, 2009; Brenzinger, 2009; Flammia and Saunders, 2007). This phenomenon is explained as the result of many different mechanisms. For instance, the economic advantages associated with certain dominant languages put pressure on people to adopt different languages and, in time, lead to the economic growth of a country (Bottom-up factors). Furthermore, this adoption of an economically beneficial language is sometimes bolstered by government policies that either restrict the use of certain languages or encourage the use of others (Top-down). These top-down and bottom-up mechanisms reinstate a power dynamic between languages and, thus, between the speakers of these different linguistic community. It creates a vicious cycle that constantly reinforces stigmas attributed to different languages. Indeed, since languages have social and cultural properties, each one is associated with a particular status within the social, political, economic context of the country (Batibo 2009, Brezenger 2009, Hagège 2009, Huss and Lindgren 2011). However, these different assigned statuses would not exist if it were not for the interactions and contact between different communities. It would appear that a lot of these top-down and bottom-up pressures towards linguistic convergence are a result of the proximity of the linguistic communities, which is intensified in urban areas. Therefore, it is the exposure to different linguistic

communities and the socio-political statuses that they embody that are driving linguistic decline. Thus, it would seem that perhaps a 'survival of the fittest' principle could be applied to the decline in linguistic diversity.

Another explanation is the dependency of communities on one another that forces them to interact and communicate using a shared language. Indeed, in largely rural areas where resources are abundant and contact is limited, high linguistic diversity continues to exist. Through a cartographic study, L. J. Gorenflo, Suzanne Romaine, Russell A. Mittermeier, and Kristen Walker-Painemilla (2012) have mapped the areas in the world with high biodiversity and those that have high linguistic diversity. They have observed a correspondence between the two variables. They argue that out "of the more than 6,900 languages currently spoken on Earth, more than 4,800 occur in regions containing high biodiversity"(2012:1). Similarly, Nettle (1999) explores the spatial patterns of linguistic diversity. He notes that there are two great belts of high linguistic diversity. The first ranges from the Ivory Coast, throughout western Africa, and into the Democratic Republic of the Congo. The other belt starts in south India, runs through the Southeast Asian islands, and into New Guinea and the islands of the western Pacific. These two belts are both high in linguistic diversity and biodiversity. According to his research, these seventeen countries that lie within the two belts account for 60% of all the languages spoken in the world[3]. He explains that the "lack of face to face interaction between groups will tend to mean that their languages diverge" (1999: 68). In addition, these small and undisturbed communities are all speakers of indigenous languages.

[3] These findings are from 1999 and might have changed since.

Both studies list numerous reasons for the co-occurrence of biodiversity and language diversity. For example, in places where biodiversity has not been disturbed by urban development or growing societies, there is less contact between different linguistic communities. The need to communicate with others is unnecessary due to the abundance of resources. Therefore, it is self-sufficiency, provided by an abundance of resources that enables small linguistic groups to live independently of one another. For example, in New Guinea, there are hundreds of localized and small communities that exchange little interaction between them. The little amount of trading is not essential enough for it to lead to social dependency and linguistic acculturation (Nettle, 1999). On the other hand, Bailey et al. (1992) have studied a pygmy group in central Africa that were specialized in hunting but were not adept in agriculture. Since they could not solely consume a carnivorous diet, they had to rely on a neighboring farming community. The dependency between these two groups became so strong that the pygmies adopted the farming communities' language. These studies illustrate the ways in which communal contact, due to dependency between communities, leads to linguistic convergence.

Logically, if sustained biodiversity enables language diversity, then urban areas, where there is almost no biodiversity and a lot more interactions between language groups, would hamper linguistic heterogeneity. Languages are similar to organisms and the more isolated they are from one another, so in a region where there are less dense urban spaces, the more sustainable they are. In an urban area, where there are more interactions between different cultures and communities, there is a higher chance for language affiliation and amalgamation, reducing the number of

languages. Thus, it is the proximity and contact in urban areas could be increasing the socio-political inequalities between language communities and encouraging linguistic convergence. In addition, the dependency between language groups leads to linguistic unity.

However, I would like to stress an important component that is missing in these results. Studies on urban spaces often show that cities contain a lot of cultural and ethnic diversity and multilingualism. Therefore, it is possible to speak both a native-tongue and a dominant language. However, there is a general tendency for countries with an important urban population to have less linguistic diversity when comparing it with countries that have more rural spaces.

What happens when we look at the effect of ICTs, which defy geographical proximity, increase global dependency and communication, and reinstate asymmetries of power between nations?

III. ICTs and Linguistic Diversity

Hypothesis 3.1

ICT consumption plays a powerful role in language decline. First of all, it is today's driving force behind globalization, which is leading to cultural homogeneity (Fairclough 2006; Heugh 2013; Steger 2009). The following models are looking at the association between Information and Communication Technologies and language diversity. The three variables for ICTs are the rate of Internet users, the rate of mobile users, and the rate of ICT imports. I will be controlling for GNI, urban population rates, and geographic location.

Models (Table 1.2)

Model 14: Y (LDI) =X0 + X1 (Internet users) + e

Model 15: Y (LDI) =X0 + X1 (Internet users) + X2 (GNI) + e

Model 16: Y (LDI) =X0 + X1 (Internet users) + X2 (Urban pop)+ e

Model 17: Y (LDI) = X0 + X1 (Mobile users) + e

Model 18: Y (LDI) = X0 + X1 (Mobile users) + X2 (GNI) + e

Model 19: Y (LDI) = X0 + X1 (ICT Imports) + e

Model 20: Y (LDI) = X0 + X1 (ICT Imports) + X2 (GNI) + e

Model 21: Y (LDI) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + e

Model 22: Y (LDI) = X0 + X1 (Internet users) +X2 (Mobile users) + X3 (ICT Imports) + X4 (GNI) + e

Model 23: Y (LDI) = X1 (Internet users) + X2 (Mobile users) + X3 (ICT imports) + X4 (Urban pop) + e

Model 24: Y (LDI) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X5 (Europe) + X6 (Namerica) + X7 (Samerica) + X8 (Asia) + X9 (Oceania) + e

Model 25: Y (LDI) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X5 (Europe) + X6 (Namerica) + X7 (Samerica) + X8 (Asia) + X9 (Oceania) + X10 (GNI) + X11 (Urban pop) + e

Results

In Model 14, I tested the bivariate relationship between the rate of Internet use and the Linguistic Diversity Index. The rate of Internet use is statistically significant

and the adjusted r-squared of the model is .14. The Internet user rate of a country is negatively correlated with language diversity. This means that in a country with high levels of Internet usage, there is less language diversity. This doesn't entirely confirms the theory that Internet usage is associated with language decline since we are looking at points in time but there seems to be a connection between the levels of internet use and the levels of language diversity in a country.

In Model 15, GNI per capita is added to Model 14 as a control variable. The adjusted r-squared has slightly increased by 2%. GNI per capita is not statistically significant, while Internet usage continues to be statistically significant. In addition, the coefficient magnitude of the Internet usage rate variable increases by 25% with the GNI control. This implies that GNI per capita is a suppressor variable. So when we consider the income levels of the countries, the rate of Internet use has a greater association with linguistic diversity. GNI per capita is not statistically significant, so even if it has a suppressor effect on Internet use, its relationship with linguistic diversity is still insignificant.

In Model 16, I tested the association of Internet use with the dependent variable, while controlling for the rate of urban population. The adjusted r-squared has increased by 5% compared to the initial bivariate model. Unlike GNI per capita, which was a suppressor variable, the urban population percentage is a mediating variable, since it decreases the magnitude of the Internet coefficient by 21%. Therefore, urban population accounts for 21% of the variance in linguistic diversity. Both the rate of Internet usage and the rate of urban population are significant and

have a negative relationship with linguistic diversity, which supports both hypothesis 1.2 and hypothesis 1.3.

In Model 17, I am testing the relationship of cell phone usage on linguistic diversity. The adjusted r-squared in this model is also really low (r2_a=.0007) and the rate of mobile use is not statistically significant. Thus the rate of cell phone use doesn't seem to have any relationship with language diversity in this bivariate model. In order to test for suppressor effects, I added GNI per capita to the bivariate model. In this new model, Model 18, GNI is statistically significant, while mobile use remains insignificant. The adjusted r-squared has largely increased from .0007 in Model 17 to .08 in Model 18. The magnitude of the GNI coefficient is almost identical to its coefficient in Model 1, where we only tested the effect of GNI on language diversity. This suggests that all of the variance in linguistic diversity in this model is a result of the GNI per capita. Thus, the rate of cell phone usage has no effect on the dependent variable.

In Model 19, I tested the third ICT variable, the rate of ICT imports. The percentage of total ICT goods imported is statistically significant but has a low adjusted r-squared. Indeed, only 1.6 % of the variance can be explained by the rate of ICT imports. Although the r-squared is low, ICT import rate is still significant and has a negative association with language diversity. So, the more a country is importing ICTs, the less language diversity it has. In Model 20, when I added GNI to this model, the total percentage of ICT imports is no longer significant, where as GNI is statistically significant. The r-square has increased by 400% and is equal to the r-

squared in Model 1, where GNI was tested by itself. Thus, the rate of ICT imports is spurious.

In Model 21, all three ICT variables are tested together. Internet user rate is the only one that remains significant. This means that out of the three ICT variables, Internet accounts for the variance in language diversity. The adjusted r-squared is only 2% higher than in Model 14, where Internet use was tested by itself, and the magnitude of the Internet use coefficient has increased by 2.4%. Thus the other two variables, ICT imports and mobile use could be acting as suppressor variables. Having said that, the increase is so low, that this is pretty unlikely.

In Model 22, GNI per capita is added to model with the three ICT variables, as a control variable. Once again, Internet user rate is the only variable that is statistically significant. The r-squared value has only increased by 2.7%, which means that GNI has slightly improved the fit of the model, but not by much. However, GNI has increased the magnitude of the Internet usage rate coefficient by 30%. By increasing the magnitude of the rate of Internet usage coefficient, GNI per capita, despite being statistically insignificant, is a suppressor variable and accounts for a very low amount of the variation in Internet use, as it did in Model 15. Nonetheless, Internet usage rate has the strongest relationship with linguistic diversity compared with the three ICT variables and with the GNI per capita variable.

In Model 23, all three ICT variables are tested together, as well as the rate of urban population as a control variable. The fit of the model is higher than in the previous model and urban population rate is statistically significant. It is also negatively significant, meaning that there is less language diversity in countries with

high rates of urban population. In addition, urban areas are technological hubs, so it is possible that the ICT variables are proxies for urban population rate. However, Internet usage continues to be statistically significant, with a 25% decrease in the magnitude of its coefficient, implying that urban rate accounts for some of the variance in linguistic diversity. While adding GNI to the ICT variables led to a 30% increase in the magnitude of the Internet use rate coefficient, the urban population variable led to a 21% decrease in the magnitude of the coefficient, when comparing both models with Model 21. Thus, this reinforces the fact that GNI has a suppressor effect on Internet use, while urban population has a mediating effect.

In Model 24, I controlled for the continent dummy variables, while testing the association of the three ICT variables on linguistic diversity. The adjusted r-squared is the highest it has been so far, with an increase of 95% compared to Model 21, where I only tested the effect of the three ICT variables on linguistic diversity. Therefore, the continent dummies, which are all statistically and negatively significant, account for good amount of the variance in the model. Once again, mobile use and ICT imports remain insignificant. The Internet usage variable is significant with a decrease in the magnitude of its coefficient. Therefore, both Internet use and the geographical location of the country are associated with linguistic diversity.

Model 25 is the full model and has an adjusted r-squared of .285. In this model, Internet usage rate is still statistically significant but the magnitude of its coefficient is much lower than in Model 14 and in Model 15. The only other variables that are significant are the continent dummies. Urban population rate, which has been statistically significant so far, is now insignificant in the full model, indicating that

Internet user rate has a stronger correlation with language decline than does the rate of urban population.

Discussions

These analyses expose five main observations. The first is that the usage of cell phones does not have a significant relationship with linguistic diversity. The second is that the rate of ICT imports also doesn't have a significant association with the dependent variable, even though it was significant when tested by itself. The third observation is that Internet use is highly associated with low levels of linguistic diversity. It could be a proxy for GNI and urban population, but after controlling for these variables they don't end up significant in the final and best-fit model. The fourth observation is that while urban population does not have an association with linguistic diversity, when tested with Internet usage rate, it does have a mediating effect with Internet use since it decreases its magnitude. Therefore, I could argue that urban population does have some significance in the model. Indeed, urban areas are also technological hubs for technology, thus it affects linguistic diversity in an indirect way. The fifth is that the geographical location of the country is associated with specific linguistic trends. This can be attributed to the cultural and historical resemblance in countries that border each other.

From these observations, appear a couple of puzzling questions:

- How is the rate of Internet use associated with low linguistic diversity?
- Why aren't cell phone usage and the rate of ICT imports associated with linguistic diversity?

As was previously noted in the discussion section of hypothesis 1.2, the physical proximity of linguistic communities enhances both the top-down and the bottom-up mechanism that reduce linguistic decline, by reinforcing a hierarchy of statuses for each language. In addition, the dependency of communities increases the necessity of a mutual language. Thus, proximity and contact, enhanced in urban areas, increases power asymmetries and dependencies between communities.

The Internet has revolutionized the way the world communicates as whole. It has increased global communication and is accredited for expediting globalization today. Thus, the geographical proximity and contact between language groups that is resulting in linguistic decline, also exists on a global level. As mentioned in the literature review, ICTs are responsible for perpetuating western ideals and supremacy, thus sustaining a power dynamic between western and non-western nations. It is also pressuring nations towards a western model of linguistic homogeneity. The same mechanisms of physical proximity and contact, which are responsible for intra-national linguistic homogeneity, can be applied to the affects of ICTs, and in particular to effects of the Internet, on global linguistic uniformity.

Furthermore, the Internet is such an important part of our everyday lives that the languages that are excluded from it are perceived as having a reduced functionality. The Internet's effect on language appears to be so strong that even some languages that are stable and wildly spoken are still threatened by it. Rehm, G., Uszkoreit, H., & SpringerLink (2012) argue that although the Internet is a communication tool, it divides linguistic communities from one another within a universal network. Indeed, each language group will only access the websites and

chat-rooms that are held in their own language. In addition to the number of speakers, one of the most significant factors of a language's survival today is its digital presence. Indeed, Cavanagh (2007) describes Wilhelm's periphery-center model of the information society. In this model, each subgroup of a society is characterized by their rank in relations to their access to information. Those who are on the extreme periphery of the information society are 'immune to progress'. This proves the prominent role of technology in our world today. Indeed, "networks today, […] act as an organizing framework in which all institutions, knowledge and relationships are ordered." (2007:24)

Even languages that are not at the periphery of the information society, such as German, are faced with disadvantages due to their syntax or semiotics that doesn't fit with technological innovation. Even if German has a very important presence in the digital world, Rehm, G., Uszkoreit, H., & SpringerLink (2012) are still worried for the future. This is due to the nature of the German language, which is an impractical language when it comes to software coding and the creation of translation technologies.

In addition, the computer keyboard is not well adapted to character-based languages. Sproat (2011) explains, "writing systems with smaller character sets had a distinct advantage" (2011:152). Certain languages, such as Mandarin, have found a new way to adapt their language to fit the keyboard, demonstrating the importance of adapting to technology in modern day. The Internet's capacities to both enhance contact between linguistic groups, beyond geographical barriers, and to reduce the

functionality of a language, due to the digital dependency, is leading to linguistic decline.

The final puzzle to these results is why mobile usage is not a significant factor in changes in linguistic diversity. Demirkin and Soper argue that mobile phones are mainly used for synchronous communication, while the Internet is mostly used for asynchronous communication. Indeed, “Mobile phones are most often used for personal communication, while the Internet is more supportive of impersonal exchanges of information” (2012: 22). Thus the information exchanges on cell phones are ephemeral and do not reinforce the hierarchal statuses of languages. In addition, it does reduce the functionality of a language since it does not require the coding or writing of a language. Thus, small indigenous languages that are excluded from the Internet will not be affected by mobile communication. This reveals that there may be a difference in the ways that technology interacts with countries that have high levels of indigenous languages. In my next chapter, I will analyze the association between ICT rates and the concentrations of indigenous and immigrant languages spoken in a given country.

CHAPTER 2—The effect of ICTs on Indigenous and Immigrant languages

The dependent variables for the following models are the percentage of indigenous languages (Table 2.1) and the percentage of immigrant languages (Table 2.2) out of the total number of languages for each country. The models are the same, but the dependant variables are different. I compared each model from Table 2.1 (Models a*) with the corresponding model in Table 2.1 (Models b*). This allowed me to look at the correlation between the independant variables and the ratio of indigenous or immigrant languages to the total number of languages in a country.

Before I begin with the analyses there are a couple points that I must reitarate about the data. The first is that these two variables, the percent of indigenous languages and the percent of immigrant languages spoken add up to the total number of languages in a given country. Thus, the adjusted r-squared and the p-values are the exact same for each corresponding model, but the direction of the relationship with the dependant variables is different.

The second point is that an indigenous language is not necessarily a language that is spoken by a minority group, as it is often assumed. It is simply a language that is native to the land. However, generally, countries with a multitude of indigenous languages are countries that historically have not existed as a modern nation relative to countries that are largely monolingual. Due to colonization or other historical and socio-political factors, the diverse indigenous communities were living separately until the unification of their nation by some external factor. On the other hand, generally, largely monolingual countries have become this way due to a historical motivation toward a nationalistic ideology. Indeed, linguistic homogeneity in a nation

is a fundamental aspect of nationalist ideology (Barbour, 2000). Thus when considering these results, it is important to consider the historically established linguistic diversity of a country.

Finally, this data does not take into account multilingualism. Many indigenous communities experience a phenomenon called diglossia where their native tongue is used at home, while other more dominant languages are used at work and for business.

I. GNI per capita and percentage of indigenous and immigrant languages

Hypothesis 2.1

A country's income level reflects the amount of either indigenous or immigrant languages spoken in a country. Countries with a higher GNI per capita will have fewer indigenous languages and more indigenous ones, and vice versa.

Models (see Table 2.1 and 2.2)

Model 1a: Y (Indigenous)= X0 + X1 (GNI) + e

Model 1b: Y (Immigrant)= X0 + X1 (GNI) + e

Model 3a: Y (Indigenous)= X0 + X1 (GNI) + X2 (Urban pop) + e

Model 3b: Y (Immigrant)= X0 + X1 (GNI) + X2 (Urban pop) + e

Model 5a: Y (Indigenous)= X0+ X1 (Europe) + X2 (Namerica) + X3 (Samerica) + X4 (Asia) + X5 (Ocenia) + X6 (GNI) + e

Model 5b: Y (Immigrant)= X0+ X1 (Europe) + X2 (Namerica) + X3 (Samerica) + X4 (Asia) + X5 (Ocenia) + X6 (GNI) + e

Model 7a: Y (Indigenous)= X0+ X1 (Europe) + X2 (Namerica) + X3 (Samerica) + X4 (Asia) + X5 (Ocenia) + X6 (GNI) + X7 (Urban pop) + e

Model 7b: Y (Immigrant)= X0+ X1 (Europe) + X2 (Namerica) + X3 (Samerica) + X4 (Asia) + X5 (Ocenia) + X6 (GNI) + X7 (Urbanpop) + e

Results

In Model 1a and 1b, we are testing the bivariate relationship between GNI per capita and the amount of indigenous and immigrant languages, respectively. GNI per capita is statistically significant with an adjusted r-squared of .14, which means that 14% of the variance in the variables can be explained by the country's income level. GNI per capita is negatively correlated with the percentage of indigenous languages, but has a positive relationship with the percentage of immigrant languages. This means that wealthier countries have less indigenous languages and more immigrant languages, while low-income countries generally have more indigenous languages spoken and less immigrant ones. This supports hypothesis 2.1.

In Model 3a and 3b, the r-squared has increased by 11%, signifying that the GNI per capita and urban population rate together create a better-fitted model than the bivariate model, Model 1. GNI continues to be significant, with a decrease in its coefficient magnitude of 32%. Urban population is also statistically significant, suggesting that it accounts for the increase in the adjusted r-squared and the decrease in the GNI per capita magnitude. Both GNI per capita and the rate of urban

population are negatively associated with indigenous languages but positively associated with immigrant languages. Thus, countries with high GNI per capita and high urbanization have more speakers of immigrant languages than of indigenous ones. In countries with lower levels of urbanization and a low-income level, there are generally more indigenous languages than immigrant ones.

Finally, in Models 5a and 5b, we are testing GNI per capita and controlling for geographic location, using the continent dummy variables. The r-squared has increased by 19% since Model 1. This increases reveals that the location of a country accounts for some the variance in the results. Interestingly, while in the previous analyses in Chapter, all five continent variables were statistically significant when testing their association the levels of linguistic diversity, in these analyses, only the *Europe* variable is significant. These results suggest that only European countries have a significant trend in their ratios of indigenous to immigrant languages. European countries have the tendency to have more immigrant languages spoken than indigenous ones. This does not mean that there are more speakers of immigrant languages, but simple, when looking at the total number of languages spoken in European countries, more of those are originally from foreign countries. GNI also continues to be significant, with a decrease in the magnitude of its coefficient.

Now let's look at the full models, Model 7a and 7b. These are also the best-fit models with an r-squared of .19. GNI per capita is no longer significant, while urban population rate and the *Europe* dummy variable remain statistically significant. GNI is spurious when it is tested with urban population rate and the continent dummies *together*. But when it is tested with urban population rate and the continent dummies

individually, it continues to be significant. These results indicate that it is a combination of the rate of urban population and the location of the country that has an association with the different percentage of indigenous and immigrants in a country. In countries with high levels of urbanization, there are less indigenous languages than immigrant languages. Furthermore, in Europe, there are lot less indigenous languages and more immigrant languages spoken, when comparing it to Africa, the standard of comparison. This doesn't necessarily mean that the majority of the population speaks an immigrant language in these types of countries. In fact, the majority of the population could be speaking the one indigenous language of the country (for example, French in France), and then the rest of the population could be split between the remaining immigrant languages. But because we are looking at the number of languages and not the number of speakers, this is not accounted for in the results.

Discussion

These analyses display the following observations. Initially, GNI per capita was significant and it revealed that a country's income per capita has a different association with the percentage of indigenous languages, than with the percentage of immigrant languages in a country. Wealthier countries have a greater proportion of immigrant languages than indigenous languages. These findings seem logical, given that generally immigrant populations migrate to wealthy urban areas.

However, GNI per capita becomes spurious when it is tested with the rate of urban population and the continent dummy variables together. Although a country's GNI is often a component that is associated with lower indigenous languages and

higher immigrant languages, it is in fact the country's urban population rate and its historical workings that are correlated with the percentage of indigenous and immigrant languages. The effects of urban population rate on indigenous and immigrant languages will be further explored in the next section.

Another observation is that only the *Europe* continent dummy is significant in the final model, which means that there is a specific reasoning behind Europe's association with low indigenous rates and high immigrant language rates. The low number of indigenous languages is a result of historical and cultural mechanisms that have led to the appropriation over time of a single national language, as a way to express national pride. Hobsbawn (1990) argues that during the Victorian Era in England, the lower middle class used linguistic nationalism in the efforts to establish their national belonging, as opposed to their class belonging. The lower middle class was a class of people who, financially, were part of a low social class but who had occupations that required schooling, including artisans, clerks, and shopkeepers. They engaged in linguistic nationalism by speaking the most proper form of their national language and teaching it to their children. This was a way to justify their national belongings despite being at the bottom of the socio-economic scale. Historically, in Europe, national pride has been a very prominent cultural occurrence that was expressed through the use of a single national language. Stephen Barbour (2000) explains that nationalism acts as a significant motivator in human behavior. This zealous loyalty to a nation drove individuals to embrace the national language and engage in behaviors that lead to economical growth. In fact, "a shared first language can facilitate greater economic and political cooperation between citizens." (2000:15).

This explains why European countries have low levels of indigenous languages and also have higher-income levels.

These results exhibit that the wealth of a country does not have an effect on the concentration of indigenous or immigrant languages in a country, but serves as a proxy for both urban rates and historical, cultural, and political mechanisms. Thus, hypothesis 2.1 is not supported by these results. In addition, there is a clear inclination for European nations to have less indigenous languages and more immigrant languages.

II. Urban population and percentage of indigenous and immigrant languages

Hypothesis 2.2

The results in Chapter 1 showed that the dependency and social comparison between linguistic communities, experienced mostly in urban areas, might be affecting linguistic decline. Since the languages that are the most vulnerable to these effects are those that are smaller in numbers of speakers, my second hypothesis is that urban spaces are detrimental to the number of indigenous language rates. Furthermore, urban areas attract immigrant communities. Thus, the second part of my hypothesis is that countries with a high urban population rate also have a higher number of immigrant languages.

Models (Table 2.1 and Table 2.2)

Model 2a: Y (Indigenous)= X0 + X1 (Urban pop) + e
Model 2b: Y (Immigrant) = X0 + X1 (Urban pop) + e

Model 3a: Y (Indigenous)= X0 + X1 (Urban pop) + X2 (GNI) + e
Model 3b: Y (Immigrant) = X0 + X1 (Urban pop) + X2 (GNI) + e

Model 6a: Y (Indigenous)= X0 + X1 (Urban pop) + X2 (Europe) + X3 (North America) + X4 (South America) + X5 (Asia) + X6 (Oceania) + e
Model 6b: Y (Immigrant)= X0 + X1 (Urban pop) + X2 (Europe) + X3 (North America) + X4 (South America) + X5 (Asia) + X6 (Oceania) + e

Model 7a: Y (Indigenous)= X0+ X1 (Europe) + X2 (North America) + X3 (South America) + X4 (Asia) + X5 (Oceania) + X6 (GNI) + X7 (Urban pop) + e
Model 7b: Y (Immigrant)= X0+ X1 (Europe) + X2 (North America) + X3 (South America) + X4 (Asia) + X5 (Oceania) + X6 (GNI) + X7 (Urban pop) + e

Results

In Model 2a and Model 2b, the rate of urban population is statistically significant with each dependent variable, with an r-squared of .12. The coefficient is statistically and negatively correlated with the percent of indigenous languages, whereas it is positively correlated with the percent of immigrant languages in Model 2b. This indicates that urban areas attract and maintain a large number of immigrant languages but not indigenous ones. This is not to say that there aren't indigenous languages in urban areas, but simply that there are less of them.

If we control for GNI in both Model 3a and Model 3b, we can see that GNI is statistically significant. The adjusted r-squared is 30% higher than in Models 2a and 2b, and both coefficients have decreased, suggesting that the variables for GNI per capita and urban population together create a better fitted model than each variable in their separate bivariate models.

Next, in Model 6a and 6b, I tested the association between urban population rates and the dependent variables, controlling for the continent dummies. The r-squared is 55% higher than when urban population rate was tested by itself. Thus, the location of the country accounts for 55% of the variance in the dependent variables. Only the South America and Europe variable are significant, so this only applies to countries that are part of these two continents. Urban population remains significant with a slight decrease of 16% in the magnitude of its coefficient. The direction of its relationship with the two dependent variables is consistent with the previous models.

In the full models (Model 7a and Model 7b), which are also the best-fit models, urban population remains significant. It is negatively associated with high levels of indigenous languages, but positively associated with high levels of immigrant languages. The only other variable that is statistically significant is the *Europe* variable, which supports the previous discussion on Europe's high levels of immigrant languages and low levels of indigenous ones.

Discussion

These findings reveal that urban population is negatively associated with the percentage of indigenous languages but positively correlated with the percentage of immigrant languages. As was previously mentioned in Chapter 1, indigenous languages are more vulnerable to the contact of linguistic communities in urban areas (Gorenflo et al. 2012; Nettle 1999). However, this does not explain why there are so many immigrant languages spoken in cities.

First of all, immigration communities tend to flock towards wealthier urban areas, which offer more job opportunities. In addition, immigration is essential to

urban development (Breckner, Peukert, and Pint 2013). For example, Germany has a very high level of prosperous urban spaces and of immigration. In Munich and in Hambourg, about 30% of the population is foreign and it has been very beneficial to the development of the city. For instance, the international atmosphere, developed through patterns of migration and reflected through ethnic restaurants and shops, contribute the city's tourist attraction. Thus, cultural diversity enriches a city and contributes to its attractiveness (Breckner, Peukert, & Pinto, 2013). Not only does immigration contribute to the development of urban spaces (Breckner, Peukert, & Pinto, 2013; Su, 2010), but also immigrant communities are attracted to the many possibilities that urban areas can offer. This virtuous cycle has increased the number of immigrant communities in wealthy urban areas. This also explains why the GNI per capita variable has a mediating effect on this association.

III. ICTs and the percentages of indigenous and immigrant languages

Hypothesis 2.3

Small indigenous languages are more fragile in urban areas, where proximity to more dominant languages could be reducing their number of speakers. So my third hypothesis is that prevalent uses of ICT technologies, which exclude many minority indigenous languages, are associated with fewer indigenous languages and more immigrant languages.

Models (Table 2.1 and Table 2.2)

Model 8a: Y (Indigenous) = X0 + X1 (Internet users) +e

Model 8b: Y (Immigrant) = X0 + X1 (Internet users) +e

Model 9a: Y (Indigenous) = X0 + X1 (Internet users) + X2 (GNI) +e
Model 9b: Y (Immigrant) = X0 + X1 (Internet users) + X2 (GNI) +e

Model 10a: Y (Indigenous) = X0 + X1 (Internet users) + X2 (Urban pop) + e
Model 10b: Y (Immigrant) = X0 + X1 (Internet users) + X2 (Urban pop) + e

Model 11a: Y (Indigenous) = X0 + X1 (Mobile users) +e

Model 11b: Y (Immigrant) = X0 + X1 (Mobile users) +e

Model 12a: Y (Indigenous) = X0 + X1 (Mobile users) + X2 (GNI) + e
Model 12b: Y (Immigrant) = X0 + X1 (Mobile users) + X2 (GNI) + e

Model 13a: Y (Indigenous) = X0 + X1 (Mobile users) + X2 (Urban pop) + e
Model 13b: Y (Immigrant) = X0 + X1 (Mobile users) + X2 (Urban pop) + e

Model 14a: Y (Indigenous) = X0 + X1 (ICT Imports) + e
Model 14b: Y (Immigrant) = X0 + X1 (ICT Imports) + e

Model 15a: Y (Indigenous) = X0 + X1 (ICT Imports) + X2 (GNI) + e
Model 15b: Y (Indigenous) = X0 + X1 (ICT Imports) + X2 (GNI) + e

Model 16a: Y (Indigenous) = X0 + X1 (ICT Imports) + X2 (Urban pop) + e
Model 16b: Y (Immigrant) = X0 + X1 (ICT Imports) + X2 (Urban pop) + e

Model 17a: Y (Indigenous) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + e
Model 17b: Y (Immigrant) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + e

Model 18a: Y (Indigenous) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X4 (GNI) + e
Model 18b: Y (Immigrant) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X4 (GNI) + e

Model 19a: Y (Indigenous) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X4 (Urban pop) +e

Model 19b: Y (Immigrant) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X4 (Urban pop) +e

Model 20a: Y (Indigenous) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X4 (Europe) + X5 (North America) + X6 (South America) + X7 (Asia) + X8 (Oceania) + e

Model 20b: Y (Immigrant) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X4 (Europe) + X5 (North America) + X6 (South America) + X7 (Asia) + X8 (Oceania) + e

Model 21a: Y (Indigenous) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X4 (Europe) + X5 (North America) + X6 (South America) + X7 (Asia) + X8 (Oceania) + X9 (GNI) + X10 (Urban pop) + e

Model 21b: Y (Immigrant) = X0 + X1 (Internet users) + X2 (Mobile users) + X3 (ICT Imports) + X4 (Europe) + X5 (North America) + X6 (South America) + X7 (Asia) + X8 (Oceania) + X9 (GNI) + X10 (Urban pop) + e

Results

The first Information and Communication Technology variable that I tested was the rate of Internet use in a country. In these models, Model 8a and 8b, the rate of Internet usage is statistically significant and the adjusted r-squared value is 0.14. In Model 8a, the coefficient is negatively associated with the concentration of indigenous languages, but in Model 8b, it has a positive association with the number of immigrant languages. These results indicate that, without controlling for any other

variable, the higher the Internet usage is in a given country, the more there are immigrant languages and the less there are indigenous ones.

In Model 9a and in Model 9b, when I control for the effect of GNI, the Internet usage variable is still significant and the r-squared value has slightly increased by 7%. The magnitude of the Internet use coefficient has decreased by almost 44%, suggesting that in this model GNI is accounting for some of the variance in concentration of either indigenous or immigrant languages.

In Model 10a and 10b, I looked at the relationship between Internet usage and the dependent variables, while controlling for the rate of urban population. The r-squared has increased by 14%s since Model 8, which is twice as much as the increase from Model 8 to Model 9. This means that the model testing the effect of urban population and Internet usage is a better-fitted model than GNI per capita and Internet usage. Therefore, urban population is a stronger determinant of the variance then GNI is, when tested alongside Internet usage. Both variables are statistically significant and have a negative association with indigenous languages, suggesting that there are less of these types of languages in urban and highly connected areas. However, the percent of immigrant languages is higher in these types of countries.

The next ICT variable that I looked at was the rate of mobile usage. In Model 11a and 11b, this variable was statistically significant, however only 1% of the variance in the results can be attributed to mobile usage. Thus cell phone usage rate accounts for a very low amount of the variance of the dependent variable. Interestingly, in Model 12a and 12b, where GNI per capita is used as a control, mobile use is significant and the r-squared is now significantly higher, increasing by

1400%. This means that GNI per capita acts as a suppressor variable and increased the magnitude of mobile usage's coefficient. Thus, when we take into account the GNI of the country, mobile use is positively associated with indigenous languages and negatively associated with immigrant languages. This can be explained by the fact that mobile phones are used for local communication. Therefore, the exchanges between people could be in their indigenous tongue. However, Internet communication is used for wider communication, sometimes on a global scale, and is used for business and trade, which can explain why indigenous languages are excluded from this type of communication.

We can observe a similar effect in Model 13a and 13b, when I tested mobile usage and controlled for urban population. Both the adjusted r-squared, as well as the cell phone coefficient's magnitude, have increased. This indicates that urban population rate also acts as a suppressor variable. However, the r-squared in Model 12a/b is higher than in this model, suggesting that GNI has a greater suppressor effect on mobile use.

The last ICT variable I test for is the percentage of ICT goods imported. In Model 14a and 14b, the percent of ICT imports is not statistically significant with neither immigrant or indigenous language percentages and the adjusted r-squared is the lowest of all of the models (r2_a=.002). Interestingly, when I added GNI per capita (in Models 15a and 15b), the ICT import rate became statistically significant and the coefficients for both variables increase by more than half. The r-squared is also really high. Furthermore, when tested with urban population, in Model 16a/b, ICT import rate is also statistically significant and the r-squared is high. Therefore,

similarly to mobile usage models, urban population and GNI act as suppressor variables. High rates of ICT imports are associated with a large number of indigenous languages and low levels of immigrant languages.

In Model 17a and 17b, all three ICT variables are tested together. The r-squared is the highest it's been so far in all the models that have been compared in this chapter (r2_a=.18). All three variables are statistically significant but they have different relationships with the dependent variables. Consistent with the previous models, the Internet use rate is negatively associated with the percentage of indigenous languages and positively associated with the immigrant languages. On the other hand, the cell phone rate and ICT import rate have a positive relationship with the indigenous language variable and a negative one with the immigrant language variable. Therefore, high cell phone rates and ICT imports in the last couple of years are associated with countries that have more indigenous languages than immigrant ones. I will discuss these relationships in the discussion section below.

In Model 18a and 18b, GNI is added to the previous model as a control. The r-squared increases by 6%. All the variables are significant, but the magnitudes of the coefficients for Internet use and mobile use have slightly decreased, indicating that GNI accounts for some of the effects of these variables on the dependent variables. While GNI per capita is a mediating variable for Internet use and mobile use, it is a suppressor variable in the case of ICT imports, since its coefficient has increased.

When comparing the three variables with urban population rate, in Model 19a and 19b, mobile usage rate is no longer significant. This indicates that urban population is a better determinant of the concentration of indigenous or immigrant

languages in a country than mobile usage rate. Urban population acts as a suppressor variable for ICT imports, since the magnitude of the coefficient for this variable has increased by18%. It has a mediating effect on Internet use however, since the magnitude of its coefficient has decreased by 32%.

Now let's look at the same association, but controlling for the geographical location of the country, using the continent dummy variables. Only Europe and Asia are statistically significant. So only for countries in these continents is this association valid. The r-squared value for this model has decreased from the previous model, indicating that these continent dummies account less for the variance than urban population rates.

Finally, the best fit model, with an r-squared of .24, happens to also be the full model. Surprisingly, Internet usage rate, which has been significant in all the models, is now insignificant. It was significant with each control variable, but when all three controls are tested together, it loses its significance. Urban population is still significant, as well as the continent dummies, Asia and Europe. In addition, mobile usage rate and ICT import rate are positively associated with indigenous languages but have a negative relationship with the number of immigrant languages in a country. These are the only two ICT variables that remain significant. Therefore, they have more of an impact in determining the concentration of both indigenous and immigrant languages in a country, than Internet usage.

Discussion

These regression analyses reveal several observations on the ways that technology affects indigenous and immigrant languages. The first is that, without

controlling for all the variables together, a country with high rates of Internet usage has a lower percentage of indigenous languages, while it has a high percentage of immigrant languages.

Tracy McHenry (2002) explores the effects of the Internet on Native American culture and language. Throughout her study, she has noticed the communities' struggle between wanting to stay authentic to their heritage and the desire to be included in a larger American culture. Additionally, Kahin and Keller (1995) examined the American public's access to the Internet. They noticed that minority populations were more likely to have less access to a computer, and consequently to the Internet. Based on these findings and her own, McHenry argues that Native Americans were some of the last people to have access to the Internet because 1) they are a minority culture and 2) they attempted to stay authentic to their heritage. Another explanation to their attempt to remain authentic is due to the governmental pressures to preserved their culture and present them as a tourist attraction. Consequently, they are dependent on tourism as a source of income. Thus, in order to make a living, they must conform to this image of American heritage that they have come to represent and must maintain an authentic lifestyle. In fact, many other indigenous cultures oscillate between an aspiration to preserve an authentic culture, either for tourism, for their own sense of identity, or for a combination of both, and a desire to be included in the larger society. April R. Summitt explores this struggle encountered by the Masaii people in Kenya. While their culture depends a lot on tourism, by preserving an authentic and technology-free lifestyle, they are also excluded from modern society. Indeed, McHenry observed that once Native

American communities started to become more connected with the modern world, it began difficult to return to their authentic way of living. Technology permeated into their lives, and it became evident that this was the only way to be part of American society. As the U.S.'s virtual presence increased, and its dependency on technology became more apparent, younger generations of Native American communities frequently used technology in their everyday lives. For example, the speakers of the indigenous language Yup'ik, spoken in Bethel, Alaska, cannot escape modern technologies. Understandably, they have adopted certain communication technologies such as cable/satellite television and the Internet. The children from this community, like every other child in the U.S., spend a lot of time exposed to television programs and surfing the web, increasing their exposure to English and decreasing their proficiency in their native tongue, Yup'ik. Therefore, the rise of the Internet and other technologies pose a threat to minority indigenous languages by increasing digital dependency. Once again, linguistic diversity is affected by the power asymmetry created by technology. In areas that are high in communicative technologies, and where its population is dependent on them, like in the U.S., avoiding the use of technology is nearly impossible, and further marginalizes the groups that do not participate in the global network. Although Internet usage is no longer significant when we take into account all of the variables together, specific case studies do show that it has some changing effect on the number of the speakers of minority indigenous languages.

The second observation from these results is that the country's rate of cell phone usage and of ICT imports are associated with high levels of indigenous

languages and low levels of immigrant languages, only when controlling for GNI per capita and urban population. Mobile use was not statistically significant in my previous chapter where linguistic diversity was the dependent variable. However, in these models, GNI and urban population had a suppressor effect on mobile use, and revealed its association with the dependent variables. Thus, the discussion on cell rates and ICT imports on indigenous languages can only be understood within the country's level of economic and urban development.

Cell phones are a cheaper form of communication and are the predominant technology in lower-income countries (Horst and Miller, 2006). Cell phones are used more frequently for tasks that would otherwise be done on a computer in higher-income countries. For example, the biggest phone company in Africa, Safaricom, has developed a cell phone banking system called M-Shwari which allows the customer to take out micro loans, manage his or her account, and deposit money all through simple maneuvers feasible on any regular cell phone. These extensive cell phone usage in low-income countries have replaced certain task that would normally be done on the Internet. Additionally, cell phones require less knowledge of technology than the Internet does. Its simplicity makes it an easier technology to adopt for an isolated community or for a developing country, where the rates of Internet uses are lower. On March 10th 2005, the Economist published an article on the use of technology in low-income countries. The article refutes the assumption that the Internet will bridge the divide between rich and poor, and instead focuses on another information and communication technology; cell phones. Mobile phones have the greatest impact on development, while computers and the Internet do not. Indeed,

computers require electricity and the ability to read. Thus they are superfluous in places where literacy and electricity are low. Cell phones allow for fast, affordable, and practical communication between people and businesses. They do not necessitate as much electricity, only enough to charge, and can be used by people who are illiterate and/or who speak a language that does not have a written form or that cannot be written on a traditional keyboard. Thus, cell phone communication allows for the use of minority indigenous languages and maintains them in the process.

Finally, one of the ways to approach the question of ICT imports is to consider the kinds of countries that have the highest rates of technological imports. Jörg Meyers (2000) has observed that low-income countries have had a greater and more rapid importation of technological goods in recent years as a way to reduce the technological gap between them and more developed countries. This can be related to the pressure to move towards modernization that developing countries experience. Perhaps, if high ICT importation rates are associated with countries that have a greater amount of indigenous languages, it is in fact because these are countries that are in the process of developing and shifting their economic, political, and social model towards one that mirrors that of developed country. This explains why only once we control for GNI per capita, does this association appear. So when considering the effect of cell phones and ICT imports on indigenous languages, we must take into account the country's income level.

Going back to our discussion on biodiversity and linguistic diversity, it seems as if the greatest threat to culture is the presence and constant communication with other more dominant cultures. Cultures and languages that are fragile due to the fewer

amounts of speakers are threatened by technology. Mobile use is used in a way that maintains languages and, therefore, does not cause as much of an effect on certain minority languages.

CHAPTER 5- English Proficiency, Linguistic Diversity, and Indigenous Languages

Many scholars argue that English is today's Lingua Franca (Toffelson 1991; Steger 2009). As a hegemonic language in the world of trade, business, science, and technological research, English has taken on cultureless properties (Wardhaugh 1987; Buttjes 1991). Due to the economic benefits that it provides, many countries encourage English proficiency. It allows for better communication with other countries, and in countries with high levels of linguistic diversity, it allows for better communication within the country.

English will be the independent variables in all three sets of models. The dependent variables will be Internet user rate for the models testing hypothesis 3.1 (Table 3.1), the Linguistic Diversity Index (LDI) for the models testing hypothesis 3.2 (Table 3.2), and, finally, the percentage of indigenous languages in a country for the models testing hypothesis 3.3 (Table 3.3).

I. English Proficiency Index (EPI) and the rate of Internet users

Hypothesis 3.1

As we have seen in the literature review, the Internet, along with most computer softwares, were created with the English language in mind. In fact, according to Steger (2009), 80% of websites are in English. This "uniformization" of the web, enabled by the English language, leads us to believe that in order to use the Internet, speaking English is almost a necessity (Kamusella 2012). The dependent variable for the following models is the rate of Internet users and the independent variable is the

English Proficiency Index. My hypothesis is that the more a country is proficient in the Lingua Franca, English; the more they will use the Internet.

Models (see Table 3.1)

Model 1: Y (Internet users)= X0 + X1 (EPI) + e
Model 2: Y (Internet users)= X0 + XI (EPI) + X2 (GNI) + e
Model 3: Y (Internet users)= X0 + X1 (EPI) + X2 (Urban pop) + e
Model 4: Y (Internet users)= X0 + X1 (EPI) + X2 (Urban pop) +X3 (GNI) + e

Results

In Model 1, the bivariate model, the English aptitude of a country is highly associated with Internet rates. In fact, 55% of the variance in the rate of Internet users is associated with the level of English proficiency. These results show that if a country has a high English Proficiency Index, then it also has high Internet usage.

A study conducted by English First on the effects of English proficiency on GNI per capita showed that these two variables are highly correlated and create a virtuous cycle, where each of them increases the other (Education First 2013). Therefore, in order to ensure that English proficiency was not a proxy for income level when testing its relationship with Internet user rate, I controlled for GNI per capita in Model 2. The model's adjusted r-squared has increased by 38%, suggesting that GNI per capita has made the fit of the model much stronger. GNI per capita has a mediating effect on English proficiency and decreases the magnitude of its coefficient. Although it doesn't cause the English proficiency variable to become

spurious, GNI per capita does account for 44% of the effect of English proficiency on Internet use.

In Model 3, I tested the same bivariate model, controlling for urban population rate. The adjusted r-squared has increased by 20% compared the bivariate model (Model 1), but it isn't as high as the r-squared in Model 2, where GNI per capita is a control variable to the bivariate model. The urban population rate variable is positively correlated with Internet usage and accounts for some of the effect of English proficiency on the country's Internet user rate. Indeed, the magnitude in the EPI coefficient has decreased by 13%, but remains positive.

Finally, in the full model, all three variables are positively correlated with Internet usage. This is also the best-fit model with an adjusted r-squared of .77, meaning that a country's English proficiency, GNI per capita, and urban rate explain 77% of the variance in its Internet usage. Urban population has had a significant decrease of 65% in the magnitude of its coefficient, which suggest that although it is correlated with Internet user rate, it has a weaker association with it than the other two variables, neither of which have endured a significant decreased in their coefficients.

Discussion

The results of these models demonstrate the strong and positive association that exists between a country's English proficiency and its level of Internet consumption. The higher a country's level of English proficiency, the more it will

have Internet users. This is explained by the fact that English is one the most commonly used languages on the web and enables wider cyber communication.

The other important observation is that GNI per capita increases the fit of the bivariate model and accounts for some of the effect of English proficiency. This indicates that English proficiency also has an indirect association with the rate of Internet users, through its relationship with GNI per capita. In fact, these three variables are often connected and reinforce each other. Chinn and Fairlie (2010) demonstrate, through a quantitative analysis, that a country's income is correlated with higher Internet usage. This explains the positive association between the GNI per capita variable and the Internet variable. In addition, certain scholars believe in the power of English as the Lingua Franca as a way to increase a country's economic gain (Miller 2002, Education First report 2013). Through this logic, if English increases the country's wealth, which in turn leads to a higher access to the Internet, English proficiency has an indirect and increasing effect on Internet usage. These models display the positive direct and indirect effect of English proficiency on the rate of Internet users.

Finally, in the full model, the coefficient of urban population rate is more than halved with the presence of GNI, which could suggest that it is mainly a proxy for GNI per capita, since generally country's that have high urban rates are also wealthier. To conclude, English proficiency, as well as GNI per capita, allows for a more prominent Internet presence, confirming my first hypothesis.

II. English proficiency and linguistic diversity

Hypothesis 3.2

In Chapter 1, my results proved that prominent rates of Internet users were correlated with lower linguistic diversity, confirming hypothesis 1.3. In addition, English proficiency enables high rates of Internet usage. Based on these two sets of findings, my following models tests the relationship between English proficiency (the independent variable) and linguistic diversity (the dependent variable), while controlling for Internet user rate. This allowed me to quantify the effects of English proficiency on linguistic diversity, compared to those of Internet usage. Additionally, I controlled for urban population rate and continent dummies[4], since those were the two variables that had a significant association with linguistic diversity in the analyses in Chapter 1.

Models (See Table 3.2)

Model 5: Y (LDI)= X0 + X1 (EPI) + e
Model 6: Y (LDI)= X0 + X1 (EPI) + X2 (Urban pop) + e
Model 7: Y (LDI)= X0 + X1 (EPI) + X2 (Europe) + X3 (North America) + X4 (South America) + X5 (Asia) + e
Model 8: Y (LDI)= X0 + X1 (EPI) + X2 (Internet users) + e
Model 9: Y (LDI)= X0 + X1 (EPI) + X2 (GNI) + X3 (Urban pop) + e

Results

In Model 5, the bivariate model, the English Proficiency Index is not significant. The adjusted r-squared is also very low (r2_a=.01), suggesting that there isn't a direct association between both variables. When we control for urban population, in Model 6, the r-squared increased by 1100%, but the English

[4] Only the variables for Asia, Europe, North America, and South America were used since the EPI reports do not include any countries in Oceania. The Africa variable was omitted as the reference group.

Proficiency Index is still insignificant. Urban population has a significant and negative coefficient, as it did my analyses in Chapter 1 (see hypothesis 1.2). Since English proficiency is insignificant, urban population rate accounts for all of the variance in linguistic diversity.

In Model 7 and 8, English proficiency continues to be insignificant. Neither the continent variables, in Model 7, nor the Internet user rate variable, in Model 8, have a suppressor effect on English proficiency. The adjusted r-squared for both these models are also higher than in the bivariate model.

Finally, in the full model, the r-squared is the highest out of all five models with an adjusted r-squared of .37. The English Proficiency Index variable and the urban population rate variable are not significant. The variable measuring the country's Internet presence is still negative and significant, with a slight coefficient increase of 3%. Therefore, the more Internet user a country has, the lower the linguistic diversity index will be. The only continent variables that are significant are North America and South America. Both of these variables have a negative coefficient, which means that they both have lower linguistic diversity levels than Africa does. Interestingly, even if Internet usage and English proficiency are highly correlated with one another, they have different effects on linguistic diversity. A country with a high Internet user rate will have low linguistic diversity, which supports hypothesis 1.3. However, English proficiency is not associated with lower or higher levels of linguistic diversity.

Discussion

These results reinstate that Internet usage is negatively correlated with linguistic diversity. On the other hand, these results refute the predicted Hypothesis 3.2 by showing an absence of any sort of association between the English Proficiency Index and the Linguistic Diversity Index. I propose two reasons behind these results. The first is that these results are the product of insufficient data. Education First has only quantified the English proficiency of 60 countries. Perhaps, the small number of observations is obstructing this association. The second reasoning is that, in fact, there is no relationship between English proficiency and linguistic diversity. This can be better understood if we consider the fact that, in many countries, English is devoid of any cultural properties and, thus, is used in a way that is strictly professional (Buttjes 1991; Wardhaugh 1987). None of the countries used in these regression analyses have English as an official language. Thus, English is a second-language that may not be as present in everyday life. It has exclusively become a means of communication and not of identification. Although English aptitude is not associated with linguistic diversity, it might have an effect on languages with specific characteristics. Based on the assumption that small indigenous languages are more vulnerable to external factors, my next models will test the correlation between a country's English proficiency on their percentage of indigenous languages.

III. English proficiency and indigenous languages

Hypothesis 3.3

Indigenous language speakers are more susceptible to the pressures to adopt a language that is more frequently used and that has economic benefits, such as a language like English. Thus, my hypothesis is that English proficiency has a negative association with the percentage of indigenous languages in a country.

Models (See Table 3.3)

Model 10: Y (Indigenous) = X0 + X1 (EPI) + e

Model 11: Y (Indigenous) = X0 + X1 (EPI) + X2 (GNI) + e

Model 12: Y (Indigenous) = X0 + X1 (EPI) + X2 (Urban Pop) + e

Model 13: Y (Indigenous) = X0 + X1 (EPI) + X2 (North America) + X3 (South America) + X4 (Europe) + X5 (Asia) + e

Model 14: Y (Indigenous) = X0 + X1 (EPI) + X2 (Internet users) + e

Model 15: Y (Indigenous) = X0 + X1 (EPI) + X2 (GNI) + X3 (Internet users) + e

Model 16: Y (Indigenous) = X0 + X1 (EPI) + X2 (Urban pop) + (Internet users) + e

Model 17: Y (Indigenous) = X0 + X1 (EPI) + + X2 (North America) + X3 (South America) + X4 (Europe) + X5 (Asia) + X6 (Internet Users) + e

Model 18: Y (Indigenous) = X0 + X1 (EPI) + X2 (North America) + X3 (South America) + X4 (Europe) + X5 (Asia) + X6 (Internet Users) + X7 (GNI) + X8 (Urban pop) + e

Results

In Model 10, I tested the effect of English proficiency on the percentage of indigenous languages in a country. The adjusted r-squared is very low. In fact, only 4% of the results explain the variance in the dependent variable. The English Proficiency Index is significant and negatively correlated with the percentage of indigenous languages. This seems logical, yet the adjusted r-squared is so low that it's

unclear whether or not these results are accurate. In order to assess this relationship, I controlled for other variables.

In Model 11, GNI per capita is added as a control variable. The adjusted r-squared has greatly increased by 500%. This generous increase is attributed to the GNI per capita variable, which is statistically significant and negatively correlated to the concentration of indigenous languages. This is consistent with my previous analyses. Furthermore, English proficiency is no longer significant in this model.

In Model 12, urban population rate is used a control. We can observe the same pattern as in Model 11. The adjusted r-squared has significantly increased to .16. It still remains lower than in Model 11, indicating that GNI makes the model a better-fit model than urban population does. In addition, English proficiency is still not statistically significant, when tested with urban population.

In Model 13, the same bivariate association is tested, with the continent dummy variables as controls. Once again English proficiency is not significant, and neither are the other variables.

In my analyses in Chapter 2, we observed that Internet usage was negatively associated with high levels of indigenous languages. In addition, based on the literature and the previous findings from hypothesis 3.1, English proficiency is highly correlated with Internet usage. Therefore, in Model 14, I tested the bivariate association, controlling for Internet usage. Both variables are significant with an adjusted r-squared of .21, making it the best-fit model so far. Intriguingly though, their coefficients have opposite directions. Although they do go hand-in-hand, as we have seen in the previous model, they have different effects on the concentration of

indigenous languages. The English proficiency of a country has a positive relationship with the dependent variable, while the rate of Internet usage is negatively associated with it, as it was previously. It would appear that Internet usage has a suppressor effect on English proficiency. Thus, when testing for both variables, a country's high levels of English proficiency could be maintaining the numerous indigenous languages in a country, while Internet usage does the opposite. I will further discuss these dynamics in the discussion section.

In Model 15, I added GNI to the previous model. The fit of the model has increased by 25%. GNI per capita is significantly and negatively correlated with the amount of indigenous languages in a country. It has a stronger mediating effect on Internet use, decreasing the magnitude of its negative coefficient by 52%, than on English proficiency, reducing its coefficient only by 4%. English proficiency continues to be positively correlated with indigenous languages.

Model 17 reveals another interesting observation. When we add the continent variables to the model, even though they are insignificant, they have a suppressor relationship with the English proficiency variable, increasing its positive coefficient by 27%. On the other hand, it has a mediating relationship with Internet use by decreasing its coefficient by 13%. So, it accounts for some of the variance in Internet usage, but enhances the positive effect of English proficiency on indigenous languages.

Finally, the full model, Model 18, is the best-fit model, with an adjusted r-squared .30. Only English proficiency and urban population are significant, when controlling for all the variables. Urban population rate has a negative relationship

with the dependent variable, while the English proficiency of a country continues to remain positive. This suggests that the latter actually sustains the high levels of indigenous languages. Finally, Internet usage is no longer significant.

Discussion

The findings reveal a couple interesting observations. The first is that urban population is negatively correlated and remains significant until the last model, while GNI per capita becomes insignificant in the full model. The increased proximity and exposure to other languages in urban areas have led to top-down and bottom-up pressures towards linguistic convergence. These findings are consistent with my previous analyses in Chapter 2.

The second observation is that the English Proficiency Index becomes positively and statistically significant when it is tested with Internet usage rate. While Internet user rate has a negative effect on the percentage of indigenous languages in a country, English proficiency has a positive one. I propose an explanation based on the literature on language revitalization and maintenance. In order to better understand this argument, I will revisit the purpose of language in identity development. John Joseph (2010) explains that, “our very sense of who we are, where we belong and why, and how we relate to those around us, all have language at their centre” (2010 :9). Language is both the medium and the essence of our identity development. Furthermore, as much as language contributes to our identity, it also has another more obvious purpose, which is to interact and convey information. Thus, language has a

dual function; 1) its utility as a means for communication and 2) its cultural symbolism, which contributes to identity formation.

Furthermore, Daniel Nettle (1999) explains that people who share the same language also share what he calls ‘primary bonds’. Primary bonds, as opposed to secondary bonds, are enduring and are formed at a young age. They aren’t shaped for any specific purpose, but rather are formed because of cultural, ethnic, or religious kinship that is cemented through common ritualistic activities and celebrations, jobs, geographical closeness, and other face-to-face interactions. On the other, secondary bonds are functional relationship created to serve a particular purpose, including trading or any specialized services. These types of bonds lead to linguistic divergence. Thus, what matters most in linguistic convergence is social identification, which is only possible when people share primary social bonds (Nettle 1999). If a person feels no cultural ties to the specific linguistic culture, then he or she will not feel the need to speak it outside of a professional or specific setting. Thus, social identification is indispensable for a community to adopt another language.

Based on these theories, it seems as if the English language, in societies where it is not the official or cultural language, is used solely to communicate between communities that share secondary bonds. It has been isolated from its cultural context, exclusively incarnating its communicative functions (Wardhaugh 1987). Through this logic, dexterity in English may conceivably encourage multilingualism and maintain minority languages. By using English solely for business and commerce, speakers of indigenous languages could continue to speak their mother tongue at home. Multilingualism can function in a way where each language is

allotted to a different part of a person's life. Indeed, it is important to consider the multiple facets of a person's identity and culture (Bucholtz and Hall 2010).

Linguistic social identification appears to be so powerful, that the vast majority of languages spoken are minority indigenous languages. In fact, only 84 out of the 6,500 languages that exist are official languages of a country (Nettle 1999). The fact that only 1.3% of languages are official languages means that people are still holding on to their native tongues and are not being influenced by societal pressures.

In conclusion, English proficiency is increasing in the world, but it doesn't seem to be negatively affecting linguistic diversity or small indigenous languages. In fact, these languages continue to persist due to the cultural context that they embody. The reason that English cannot replace languages around the world, despite its prominent presence, is because it is de-contextualized. Thus, English as a second language is devoid of any cultural properties and, thus, is simply a tool for communication. In fact, Education First describes the importance of English proficiency by comparing it to literacy. Literacy was once a skill that only the elite possessed, until it later became a necessity for workforce. English proficiency is going down a similar path and is becoming indispensible in a globalized world (Education First 2013). Therefore, it is no longer just a privilege marking a higher status, but a basic skill that is being acquired by everyone in the world.

CONCLUSION

Throughout my literature review and quantitative research, I explored the different mechanisms involved in language changes. In my first chapter, the findings revealed that the rate of urban population and the rate of Internet users have the greatest association with linguistic decline. Indeed, the geographical proximity and contact of linguistic communities in urban areas has reinforced the power asymmetries between these groups, which in turn sustains the top-down and bottom-up mechanisms. It has also created a continual dependency between linguistic groups. Both the hierarchy and dependency of language groups lead to linguistic deterioration. The Internet functions similarly to the ways that urban areas affect languages, but it does so on a global scale, as opposed to a national scale. Indeed, it increases proximity and contact between nations and sustains the western hegemony on non-western nation, and propagates the model of modernity, which is based on linguistic unification.

In Chapter 2, Internet usage and cell phone usage have opposite relationships with minority indigenous languages. Internet usage is detrimental to these types of languages, since it reduces the functionality of certain languages that cannot adapt to the technological network. Cell phones, on the other hand, sustain high concentrations of indigenous languages, since they do not hamper the functionality of smaller languages.

Finally, in Chapter 3, although both of these factors are tied to globalization, English proficiency and Internet usage have different effects on linguistic diversity and on the concentration of indigenous languages in a country. While Internet usage

is associated with linguistic decline, English proficiency has no effect. Furthermore, when testing the association of these two factors on the concentration of indigenous languages, Internet usage continues to have a negative association, while English proficiency actually sustains small indigenous languages. It seems as if the existence of a language for communicative purposes only, such as the presence of English in countries where it is not the native language, has amplified the use of languages as a cultural identifier. Language reflects culture, but only once it has been decontextualized do people realize the importance of their language for cultural assimilation. This shows that the perception of a language as a marker of identity has become stronger *because* of the spread of a lingua franca, which is only used for communication purposes.

Thus, minority groups are rediscovering their roots and emphasizing the use of their native language as a form of social mobilization. As it has previously been established, language goes hand-in-hand with culture and has an important role in fortifying identities. Indeed, in the past decades, there have been numerous efforts to revitalize or maintain endangered languages (Yamamoto et al., 2008).

Broadening the topic: Language Rights

As many minority groups are being deprived of their native tongue, either because of language policies or simply because it is advantageous in our globalized world today to adopt a more main stream language, we are starting to raise the question of language rights. Only recently has this become a global issue, possibly because language is increasingly recognized as an important cultural factor of a

person's identity and self-worth. In her article, "The Right to Speak One's Own Language: Reflections on Theory and Practice" (2007), Sue Wright describes this relatively new concept, which entails that each person should have the right to practice her or his own language, without any socio-political pressure to abandon it for a more dominant language. She starts by explaining the significance of a language on a person's life:

> From 19th century romanticism to late 20th century post-modernism, some scholars have held that individuals create languages from their own individual experiences and for their personal communication needs and that each set of language practices frames reality for those who use it. (2007: 205)

Additionally, she states De Saussure's distinction between "parole", which includes all of people's idiosyncrasies and imperfections, with "langue", which is simply a system of communication. He says that "langue", the ideal and purest model of the particular linguistic system, does not exist in practice (De Saussure, 1919). Language is not a uniform structure that can simply be imposed on anyone. Yet, De Saussure's definition of "langue" is usually what language policy makers have in mind when they are implementing a standardize language in a country. Indeed, Wright explains that policy makers continue to view language as an "ideal system" as opposed to a "contextually bound performance". They, therefore, don't realize that by imposing a certain language over another or depriving a community to speak their own language, they are challenging the fundamental human rights to freedom.

Furthermore, culture is embedded in language and therefore the use of a language strengthens the culture. Language policies and language planning control the circulation of a language within a country. This has caused many regions to want

to emancipate from the state they are in. For example, through protest, organized mobilization, and continuously speak their language, Catalans in Spain are beginning to catch their government's attention. Around the world, people fight for the right of self-expression and freedom of speech. A language policy that oppresses a minority language could lead to the feeling that one is being obstructed from their right to freedom of speech. Indeed, "When large portions of the population are denied forms of self-expression, the nation's political and social foundations are weakened" (Romaine, 2000:17).

REFERENCES

Auburger, L.1990. Linguistic minority relations. *Sociolinguistica* 4: 169-190

Barbour, Stephen. 2000. “Nationalism, Language, Europe.” Pp 1-17 in *Language and Nationalism in Europe,* edited by S. Barbour and C. Carmichael. Oxford: Oxford University Press.

Baily, R., Jenike, M., Ellison, P., Bentley, B., Harrigan, A., and Peacock, N. 1992. “The ecology of birth seasonality among agriculturalists in central Africa.” *Journal of Biosocial Science*, 24: 393-412.

Batibo, Herman. 2009. “Poverty as a Factor in Language Maintenance and Language Death.” Pp 29-36 in *Language and Poverty,* edited by W. Harbert, S. McConnell-Ginet, and A. Miller. Clevedon, GBR: Multilingual Matters.

Breckner, Ingrid, Hagen Peukert, and Alexander Pinto. 2013. “The delicate search for language in spaces: Multilingualism as a resource in urban development?” Pp 209-226 in *Multilingualism and Language Diversity in Urban Areas,* edited by P. Siemund, I. Gogolin, M.E. Schulz, J. Davydova. Philadelphia, PA: John Benjamins Publishing Co.

Brenziger, Matthias. 2009. “Language Diversity and Poverty in Africa.” Pp 37-49 in *Language and Poverty,* edited by W. Harbert, S. McConnell-Ginet, and A. Miller. Clevedon, GBR: Multilingual Matters.

Buttjes, Dieter. 1991. “Mediating Languages and Cultures: The Social and Intercultural Dimension Restored.” Pp 3-16 in *Mediating Languages and Cultures*, edited by D. Buttjes and M. Byram. Clevedon, GBR: Multilingual Matters.

Cavanagh, Allison. 2007. *Sociology in the Age of the Internet.* New York, NY: Open University Press.

Chinn, Menzie D. and Robert W. Fairlie. 2010. “ICT Use in the Developing World: An Analysis of Differences in Computer and Internet Penetration.” *Review of International Economics* 18(1) 153-167.

Crystal, David. 2001. *Language and the Internet*. Cambridge, UK: Cambridge University Press.

De Saussure, Ferdinand.1986[1919]. *Course in general linguistics (3rd ed.).* Chicago: Open Court Publishing Company.

Dixon, R. M. W. 1991. "A Changing Language Situation: The Decline of Dyirbal, 1963-1989." *Language in Society* 20(2): 183-200.

Fairclough, Norman. 2006. *Language and Globalization.* New York, NY: Routledge.

Flammia, Madelyn and Carol Saunders. 2007. "Language as power on the Internet." *Journal of the American Society for Information Science and Technology* 58(12): 1899–1903.

Gorenflo, L. J., Suzanne Romaine, Russell A. Mittermeier, and Kristen Walker-Painemilla. 2012. "Co-occurrence of linguistic and biological diversity in biodiversity hotspots and high biodiversity wilderness areas" *PNAS* 109(21): 8032-8037.

Hagège, Claude. 2009. *On the Death and Life of Languages*. New Haven, CT: Yale University Press.

Herriman, Michael and Burnaby, Barbara. 1996. "Introduction." Pp 1-14 in *Language Policies in English-dominant Countries : Six Case Studies*, edited by M. Herriman, and B. Burnaby. Clevedon, GBR: Multilingual Matters.

Heugh, Kathleen. 2013. "Mobility, migration and sustainability: re-figuring languages in diversity." *International Journal of the Sociology of Language* 2013(222): 5-32.

Hobsbawn, Eris. 1990. *Nations and Nationalism since 1780: Programmes, Myth, Reality*. Cambridge: Cambridge University Press.

Horst, Heather A., and Daniel Miller. 2006. *The Cell Phone: An Anthropology of Communication.* New York, NY: Berg Press.

House, Juliane. 2003. "English as lingua franca: A threat to multilingualism?" *Journal of Sociolinguistics* 7(4): 556-578.

Huss, Leena and Anna-Riita Lindgren. 2011. “Introduction: defining language emancipation.” *International Journal of the Sociology of Language* 209(2011): 1-15.

Joseph, John E. 2010. “Identity.” Pp 9-17 in *Language and Identities*, edited by C. Llamas and D. Watt. Edinburgh: Edinburgh University Press.

Kahin, B., and J. Keller. 1995. *Public access to the Internet*. Cambridge, MA: MIT Press.

Kamusella, Tomasz. 2012. “The global regime of language recognition.” *International Journal of the Sociology of Language*, 2012(218): 59-86.

McHenry, Tracey. 2002. “Words as big as the screen: Native American languages and the Internet.” *Language Learning and Technology*, 6(2):102-115.

Miller, John J. 2002. “How do you say ‘Extinct’?” *The Wall Street Journal.* Retrieved August 18, 2013 (http://online.wsj.com/news/articles/SB10155511301334616 80).

Nettle, David. 1999. *Linguistic Diversity.* Oxford: Oxford University Press.

Quigley, John M. 1998. “Urban Diversity and Economic Growth.” The Journal of Economic Perspectives 12(2): 127-128.

Ricento, Thomas. 2009. *Ideology, Politics and Language Policies.* Philadelphia, PA: John Benjamins Publishing Company

Ridge, Stanley G. M. 1996. “Language Policy in a Democratic South Africa.” Pp 15-34 in *Language Policies in English-dominant Countries : Six Case Studies*, edited by M. Herriman, and B. Burnaby. Clevedon, GBR: Multilingual Matters.

Romain, Suzanne. 2000. “Multilingualism, conflict, and the politics of indigenous language movements.” *Estudios de Sociolinguistica* 1(1): 13-25.

Sonntag, Selma K. 2003. *The Local Politics of Global English.* Lanham, Maryland: Lexington Books.

Sproat, Richard. 2010. *Language, Technology, and Society.* Oxford: Oxford University Press.

Steger, Manfred. 2009. *Globalization: A very short introduction.* Oxford: Oxford University Press

Tollefson, James. 1991. *Planning language, planning inequality: Language policy in the community.* London, UK: Longman Press.

Vulgarakis, Evangelos and Bei Dawei. 2011. "Confronting "linguistic genocide": Language repression in Kurdistan" Pp 23-43 in *Globalization and Human Rights in the Developing World,* edited by D. M. Nault and S. L. England .

Wardhaugh, Ronald. 1987. *Languages in competition: Dominance, diversity, and decline*. Oxford: Oxford University Press.

Yamamoto, Akira, Matthias Brenzinger, and María E. Villalón. 2008. "A Place for All Languages: On Language Vitality and Revitalization." *Museum International* 60(3): 60-70.

ANNEX

Table 1: Descriptive Table for Hyp. 1.1 and Hyp 1.2

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Linguistic Diversity Index	518	0.4640541	0.3109797	0	0.99
Country-level variables					
GNI per capita (PPP)	518	1.28068	1.434338	0.018	8.467
Urban Population (%)	518	54.41098	23.16739	9.375	100
Indigenous languages (%)	518	74.0854	24.59192	9.375	100
Immigrant languages (%)	518	25.9146	24.59192	0	90.625
Continent					
Africa(omitted)	518	0.2876448	0.4531022	0	1
Asia	518	0.2432432	0.4294555	0	1
Europe	518	0.2239382	0.4172841	0	1
North America	518	0.1158301	0.3203304	0	1
South America	518	0.0656371	0.247886	0	1
Oceania	518	0.0637066	0.2444656	0	1

Table 2: Descriptive Table for Hyp 1.3

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Linguistic Diversity Index	401	0.4333682	0.3048867	0.002	0.972
Country-level variables					
GNI per capita (PPP)	401	1.410665	1.404996	0.046	6.871
Urban Population (%)	401	57.62913	22.3894	9.375	100
Continent					
Africa (omitted)	401	0.2487685	0.4328327	0	1
Asia	401	0.2216749	0.415886	0	1
Europe	401	0.2881773	0.4534729	0	1
North America	401	0.1256158	0.3318247	0	1
South America	401	0.0788177	0.2697865	0	1
Oceania	401	0.0369458	0.1888615	0	1
ICT variables					
Internet Users (%)	401	32.52879	27.77403	0.0990318	95.02
Mobile subscription (%)	401	80.6666	46.33022	0.5529338	284.3392
ICT Imports (% total goods imports)	401	7.219575	5.995322	0.3376266	45.8053

Table 3: Descriptive Table for Hyp 2.1 and Hyp 2.2

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Indigenous languages (%)	405	72.46777	23.78042	9.375	100
Immigrant languages (%)	405	27.53223	23.78042	0	90.625
Country-level variables					
GNI per capita/10, 000 (PPP)	405	1.412699	1.406135	0.046	6.871
Urban Population (%)	405	57.62913	22.3894	9.375	100
Continent					
Africa (omitted)	405	0.2493827	0.4331909	0	1
Asia	405	0.2197531	0.4145915	0	1
Europe	405	0.2888889	0.4538068	0	1
North America	405	0.1259259	0.3321762	0	1
South America	405	0.0790123	0.2700917	0	1
Oceania	405	0.037037	0.1890862	0	1
ICT variables					
Internet Users (%)	405	32.59157	27.77952	0.0990318	95.02
Mobile Users (%)	405	80.6666	46.33022	0.5529338	284.3392
ICT Imports	405	7.219575	5.995322	0.3376266	45.8053

Table 1.1

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
GNI per capita	-.05883318***			-.05105163***	-.05105163***	
Indigenous (%)		.00232535***		.0011822*		
Immigrant (%)			-.00232535***		-.0011822*	
Urban Population (%)						-.00422888***
Europe						
North America						
South America						
Asia						
Oceania						
_cons	.53940051***	.29177963***	.52431454***	.44185129***	.560071***	.69415133***
r2_a	0.07183964	0.03194161	0.03194161	0.07751794	0.07751794	0.09750659
N	518	518	518	518	518	518

Variable	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
GNI per capita	-0.02372857		-.02468589*			-0.00504472
Indigenous (%)				0.00093163		0.00079164
Immigrant (%)					-0.00093163	
Urban Population (%)	-.0032449***			-.00384875***	-.00384875***	-.00159306*
Europe		-.32876961***	-.27581585***			-.25164095***
North America		-.42107651***	-.39430661***			-.38834348***
South America		-.40488827***	-.39317422***			-.34684872***
Asia		-.11882572***	-.09504114**			-.07931412*
Oceania		-.22832499***	-.21326574***			-.21717454***
_cons	.67100116***	.65647651***	.66561841***	.60444816***	.69761098***	.65577397***
r2_a	0.10238429	0.25420114	0.26201108	0.10039776	0.10039776	0.27064082
N	518	518	518	518	518	518

Legend: *p<0.05; ** p<0.01; *** p<0.001

Table 1.2

Variable	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19
GNI per capita		0.0295435			-.06235797**	
Internet users (%)	-.00421855***	-.00549905***	-.00333677***			
Urban Population (%)			-.00170874*			
Mobile users (%)				.1909352	.2146838	
ICT Imports (%)						-.00689085**
Europe						
North America						
South America						
Asia						
Oceania						
_cons	.5688218***	.56854631***	.63856239***	.42710487***	.51425499***	.48257778***
r2_a	0.14342182	0.14646693	0.15063121	0.00069003	0.813359	0.01603557
N	401	401	401	401	401	401
Variable	Model 20	Model 21	Model 22	Model 23	Model 24	Model 25
GNI per capita	-.05846268***		0.03109422			0.02864605
Internet users (%)		-.00411775***	-.00530532***	-.00338716***	-.00227434***	-.00304426**
Urban population (%)				-.00171139*		-0.00103413
Mobile users (%)		.271857	.2830547	.260547	.275895	.2069584
ICT Imports (%)	-0.00274331	-0.00126895	-0.00167883	-0.00035632	-0.0012855	-0.00124514
Europe					-.21370652***	-.21152725***
North America					-.33211252***	-.33094928***
South America					-.34620209***	-.31376217***
Asia					-.08642759*	-.08472992*
Oceania					-.38622115***	-.38345897***
_cons	.5348657***	.57304277***	.58281631***	.63773189***	.67449422***	.7101547***
r2_a	0.07996469	0.14590104	0.1498339	0.14644055	0.28222238	0.28358954
N	401	401	401	401	401	401

Legend: *p<0.05; ** p<0.01; *** p<0.001

Table 2.1

Variable	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a
GNI per capita (PPP)	-6.375819***		-4.3272794***		-4.350251***	
Urban Population		-.37412529***	-.19233848**			-.31442864***
North America				-7.0691944	-1.7775876	-1.9737647
South America				-.5781329	1.6115158	9.9302322*
Europe				-22.017195***	-12.455068***	-12.764031***
Asia				-9.716979**	-5.7486122	-4.8699084
Oceania				-16.053923**	-10.980221	-11.069705
Internet Users (%)						
Mobile Users (%)						
ICT Imports (%)						
_cons	81.474885***	94.02829***	89.665215***	82.494096***	83.977963***	95.219514***
N	405	405	405	405	405	405
r2_a	0.14000176	0.12190029	0.15607291	0.12411103	0.16567977	0.18620218

Variable	Model 7a	Model 8a	Model 9a	Model 10a	Model 11a	Model 12a
GNI per capita (PPP)	-1.812394		-3.3363194*			-6.424428***
Urban Population (%)	-.24863412***			-.19649056**		
North America	-0.83540869					
South America	8.6435933					
Europe	-10.716509**					
Asia	-4.230869					
Oceania	-9.9988614					
Internet Users (%)		-.32461831***	-.18133927*	-.22290175***		
Mobile Users (%)					.3035*	.3269**
ICT Imports (%)						
_cons	93.174913***	83.047594***	83.091119***	91.056073***	71.597468***	80.606054***
N	405	405	405	405	405	405
r2_a	0.18894638	0.14167462	0.15049769	0.15974473	0.0106792	0.15316341

Variable	Model 13a	Model 14a	Model 15a	Model 16a	Model 17a	Model 18a
GNI per capita (PPP)			-7.3798622***			-3.9230051**
Urban Population (%)	-.377282***			-.43821356***		
North America						
South America						
Europe						
Asia						
Oceania						
Internet Users (%)					-.36987642***	-.20499495**
Mobile Users (%)	.3262526**				.252182548*	.23918859*
ICT Imports (%)		0.26499493	0.78453387***	0.77220767***	0.68009361***	0.74045632***
_cons	93.2747***	70.554622***	77.229294***	92.14663***	78.889726***	78.659389***
N	405	405	405	405	405	405
r2_a	0.13494859	0.00199304	0.17363584	0.15414688	0.18011902	0.19308689

Variable	Model 19a	Model 20a	Model 21a
GNI per capita (PPP)			-1.2964628
Urban Population (%)	-0.2427***		-.26677***
North America		-1.4415384	-0.81945671
South America		1.3280286	7.1804778
Europe		-9.9287192*	-8.5413461*
Asia		-9.96258**	-8.2643876*
Oceania		-9.57158	-8.718676
Internet Users (%)	-0.2520574	0.29296***	0.11377139
Mobile Users (%)	0.235365	0.3069587*	0.2770254*
ICT Imports (%)	.80908598***	0.71431***	0.84135***
_cons	88.152122***	81.46734***	90.654809***
N	405	405	405
r2_a	0.2081066	0.20367439	0.23895114

Legend: *p<0.05; ** p<0.01; *** p<0.001

Table 2.2

Variable	Model 1b	Model 2b	Model 3b	Model 4b	Model 5b	Model 6b
GNI per capita (PPP)	6.375819***		4.3272794***		4.350251***	
Urban Population		.37412529***	.19233848**			.31442864***
North America				7.0691944	1.7775876	1.9737647
South America				.5781329	-1.6115158	-9.9302322*
Europe				22.017195***	12.455068***	12.764031***
Asia				9.716979**	5.7486122	4.8699084
Oceania				16.053923**	10.980221	11.069705
Internet Users (%)						
Mobile Users (%)						
ICT Imports (%)						
_cons	18.525115***	5.9717106	10.334785**	17.505904***	16.022037***	4.7804859
N	405	405	405	405	405	405
r2_a	0.14000176	0.12190029	0.15607291	0.12411103	0.16567977	0.18620218

Variable	Model 7b	Model 8b	Model 9b	Model 10b	Model 11b	Model 12b
GNI per capita (PPP)	1.812394		3.3363194*			6.4244278***
Urban Population (%)	.24863412***			.19649056**		
North America	0.83540869					
South America	-8.6435933					
Europe	10.716509**					
Asia	4.230869					
Oceania	9.9988614					
Internet Users (%)		.32461831***	.18133927*	.22290175***		
Mobile Users (%)					-.3035*	-.3269**
ICT Imports (%)						
_cons	6.8250874*	16.952406***	16.908881***	8.9439274**	28.402532***	19.393946***
N	405	405	405	405	405	405
r2_a	0.18894638	0.14167462	0.15049769	0.15974473	0.0106792	0.15316341

Variable	Model 13b	Model 14b	Model 15b	Model 16b	Model 17b	Model 18b
GNI per capita (PPP)			7.3798622***			3.9230051**
Urban Population (%)	.377282***			.43821356***		
North America						
South America						
Europe						
Asia						
Oceania						
Internet Users (%)					.36987642***	.20499495**
Mobile Users (%)	-.3262526**				-.252182548*	-.23918859*
ICT Imports (%)		-0.26499493	-0.784534***	-0.77220767***	-0.680094***	-0.74045632***
_cons	6.7252072*	29.445378***	22.770706***	7.8533698*	21.110274***	21.340611***
N	405	405	405	405	405	405
r2_a	0.13494859	0.00199304	0.17363584	0.15414688	0.18011902	0.19308689

Variable	Model 19b	Model 20b	Model 21b
GNI per capita (PPP)			1.2964628
Urban Population (%)	0.2427***		.26677***
North America		1.4415384	0.81945671
South America		-1.3280286	-7.1804778
Europe		9.9287192*	8.5413461*
Asia		9.96258**	8.2643876*
Oceania		9.57158	-8.718676
Internet Users (%)	0.2520574	0.29296***	0.11377139
Mobile Users (%)	-0.235365	-0.3069587*	-0.2770254*
ICT Imports (%)	-.80908598***	-0.71431***	-0.84135***
_cons	11.847878***	18.53266***	9.3451911***
N	405	405	405
r2_a	0.2081066	0.20367439	0.23895114

Legend: *p<0.05; ** p<0.01; *** p<0.001

Table 3.1

Variable	Model 1	Model 2	Model 3	Model 4
English Proficiency Index	2.7849373***	1.5541427***	2.4325581***	1.604482***
GNI per capita		10.003057***		8.6151454***
Urban Population (%)			.52083333***	.178004*
_cons	-102.12262***	-57.611917***	-120.01872***	-69.904031***
N	154	154	154	154
r2_a	0.55410068	0.75876815	0.6614333	0.76610415

Table 3.2

Variable	Model 5	Model 6	Model 7	Model 8	Model 9
English Proficiency Index	-0.00476816	-0.0010795	-0.00474434	0.0063223	0.00357929
Urban Population (%)		-.00545203***			-0.00032614
Europe			-0.08909507		0.01219272
North America			-.21912708*		-.21165593*
South America			-.29287022**		-.25273826**
Asia			0.13111375		0.16282574
Internet Users (%)				-.0039823***	-.00408663**
_cons	.60295379***	.79028826***	.66004714***	0.1962705	0.37096192
N	154	154	154	154	154
r2_a	0.0098998	0.12450718	0.30852268	0.07471189	0.37092365

Table 3.3

Variable	Model 10	Model 11	Model 12	Model 13	Model 14
English Proficiency Index	-.71683239**	0.3572645	-0.38859049	0.10210364	.79466364*
GNI per capita		-8.7295249***			
Urban Population (%)			-.4851573***		
North America				10.482623	
South America				8.8935378	
Europe				-12.814903	
Asia				4.0737273	
Internet Users (%)					-.5427397***
_cons	107.1304***	68.286544***	123.80066***	65.259952***	51.704398**
N	154	154	154	154	154
r2_a	0.04134972	0.23901601	0.15847136	0.11447923	0.20559467

Variable	Model 15	Model 16	Model 17	Model 18
English Proficiency Index	.75942354*	0.62360487	1.0344733**	.86175677*
GNI per capita	-6.1410755**			-2.729848
Urban Population (%)		-.26843685*		-.4294526**
North America			11.010452	15.338358
South America			12.567569	21.800871*
Europe			-1.4679503	-1.5389745
Asia			7.5904927	9.1555871
Internet Users (%)	-.25876585*	-.41610326***	-.47146594***	-0.11989685
_cons	53.378547***	73.860474***	31.094772	57.862656**
N	154	154	154	154
r2_a	0.25492302	0.22900037	0.22050114	0.30148847