

BOLIVIAN IMMIGRANTS IN SÃO PAULO:  
A SOCIOLINGUISTIC STUDY OF  
LANGUAGE CONTACT IN THE CITY

*“Se você não aprender falar direitinho você não vai se dar bem com São Paulo.*

*Agora se você falar bem você se dá bem com São Paulo.”*

DISSERTATION

Zur Erlangung des Grades eines Doktors der Philosophie  
dem Fachbereich Philosophie und Geisteswissenschaften

Freie Universität Berlin

Vorgelegt von Stephanie Niehoff

Berlin, 2011

Erstgutachter: Prof. Dr. Uli Reich (Freie Universität Berlin)  
Zweitgutachterin: Prof. Dra. Verena Kewitz (Universidade de São Paulo)

Tag der Disputation: 20. Januar 2012

*This thesis is dedicated to my parents,*

*Dorle and Harald*

## Acknowledgements

---

I would like to express my gratitude to all the people who supported me with conducting and writing this research project.

My special thanks go to:

My Bolivian interview partners in São Paulo, for sharing their life stories with me.

The *Paulistanos* who participated in the quantitative study, for giving me their honest opinions.

My friends from Santa Fé, for their input as native speakers, their hospitality and kindness.

My colleagues and friends Anne, Gordon, Martin, Silca, Susanne, and Victor, for their strict eyes, helpful commentaries and wise corrections.

The *Duvidosos* of the *FFCHL, Universidade de São Paulo*, for their instructive objections and open discussions.

Michael Thomas, for technical advice and IT-support.

Prof. Verena Kewitz, *Universidade de São Paulo*, for opening my eyes about São Paulo, for sharing her knowledge and linguistic insights, and for her never-ending patience and support.

Prof. Uli Reich, *Freie Universität Berlin*, who introduced me to Linguistics and taught me to listen carefully, for his assistance and guidance and his stimulating enthusiasm about Sociolinguistics and Brazil.

The *German Academic Exchange Service (DAAD)* and the *Friedrich-Ebert-Foundation* for financing this project.

In addition, I would like to thank my family and my boyfriend for their constant support, their patience, their motivating words and their wise advice from the beginning until the end of this thesis. Dorle, Harald, Ulli, Christian and Micky – I could not have done it without you.

## Contents

---

Contents.....	i
List of Figures and Tables .....	vi
List of Abbreviations .....	viii
1. Introduction.....	1
1.1 Migration to São Paulo .....	2
1.2 Open questions and aims of the study.....	3
1.3 Outline of the study .....	5
2. The Bolivians in São Paulo.....	7
2.1 The city .....	7
2.1.1 The historical context: immigration.....	9
2.1.2 The socioeconomic context: segregation.....	12
2.2 The Bolivians .....	27
2.2.1 The historical context.....	27
2.2.2 The socioeconomic context.....	29
2.3 Summing up .....	34
3. State of the Art.....	36
3.1 (Urban) Sociolinguistics .....	36
3.1.1 Structural variation.....	38
3.1.2 Stylistic variation .....	41
3.1.3 Perception.....	44
3.1.4 Urban sociolinguistics in Brazil.....	46
3.1.5 Urban sociolinguistics in Bolivia.....	49
3.2 Language Contact and Second Language Acquisition .....	51
3.2.1 Process oriented approaches .....	53
3.2.1.1 Borrowing and shift-induced interference .....	54
3.2.1.2 Borrowing and imposition .....	56
3.2.1.3 Communication accommodation .....	59
3.2.1.4 Koine, koineization and dialect leveling .....	61
3.2.1.5 Code mixing.....	64

3.2.2 Classroom contact situations between Spanish and Portuguese .....	66
3.2.3 Natural contact situations between Spanish and Portuguese.....	68
3.3 Holistic approaches.....	70
3.3.1 <i>Migrationslinguistik</i> .....	70
3.3.2 <i>Urbane Prozesse</i> .....	72
3.4 Summing up.....	73
4. A theoretical approach to language contact caused by migration.....	75
4.1 The sociological basis.....	76
4.2 The linguistic processes in language contact situations and their cause.....	77
4.2.1 Accommodation toward social networks – by social indexicality.....	78
4.2.2 Further processes: L1-transfer and simplification.....	80
4.3 Bolivian immigrants in São Paulo.....	81
4.4 Methodology.....	82
4.5 Summing up.....	84
5. Methodology: a mixed methods approach.....	86
5.1 The sequential exploratory design (QUAL → quan).....	86
5.2 The qualitative part of the study.....	88
5.2.1 Ethics.....	89
5.2.2 Participants.....	90
5.2.2.1 Making contact.....	90
5.2.2.2 The sampling strategy.....	91
5.2.2.3 Selecting the informants.....	92
5.2.3 Sources of data.....	98
5.2.3.1 Observation.....	98
5.2.3.2 The sociolinguistic interview.....	99
5.2.3 Transcription and analysis.....	102
5.3 Connecting the two parts.....	103
5.3.1 Participants.....	104
5.3.2 Sources of data.....	104
5.3.3 Selecting the examples.....	105
5.4 The quantitative part of the study.....	106
5.4.1 Participants.....	106

5.4.2	The questionnaire .....	111
5.4.3	Procedure.....	112
5.4.4	Data analysis .....	113
5.5	Summing up .....	114
6.	Characteristics of the Bolivian's Portuguese in São Paulo .....	116
6.1	The varieties in contact.....	117
6.2	The analysis of the corpus .....	120
6.2.1	Semi-open vowels $\epsilon$ and $\circ$ .....	121
6.2.2	Diphthongization before final /s/ .....	123
6.2.3	Nasal vowels .....	125
6.2.4	Reduction vs. variation of o/u and e/i.....	127
6.2.5	Palatization of -ti/di.....	130
6.2.6	Vibrants .....	133
6.2.6.1	Vibrants at syllable onset.....	134
6.2.6.2	Vibrants in syllable coda.....	136
6.2.6.3	Rhotacism.....	137
6.2.7	Sibilants inside the word.....	137
6.2.8	Elision of final /s/ and absence of number agreement .....	139
6.2.9	Gender attribution and absence of gender agreement .....	144
6.2.10	Personal pronouns and verbal paradigm.....	149
6.2.10.1	Personal pronouns .....	153
6.2.10.2	The absence of verbal agreement .....	156
6.2.11	Prepositions .....	159
6.2.12	Word order .....	165
6.2.13	Dicourse marker .....	166
6.2.14	Common Expressions .....	170
6.2.15	Code-Mixing .....	172
6.3	Summing up .....	176
7.	The evaluation of the Bolivians' speech.....	183
7.1	The linguistic perception of the Bolivians.....	186
7.1.1	Linguistic features.....	187
7.1.2	Linguistic classification of the speech.....	191

7.1.2.1 Labeling the language .....	191
7.1.2.2 The speakers' place of origin .....	193
7.1.2.3 The speakers' style.....	196
7.2 Social evaluation.....	201
7.2.1 The speakers' place of residence .....	201
7.2.2 The ascription of social characteristics.....	204
7.3. Personal evaluation .....	207
7.4 Summing up .....	214
8. The Portuguese of Bolivian immigrants.....	217
8.1 The not assimilated workers (NAW).....	218
8.1.1 The social network.....	219
8.1.2 Linguistic characterization.....	222
8.1.2.1 Accommodation .....	223
8.1.2.2 L1-transfer.....	224
8.1.2.3 Morphological simplification.....	226
8.1.3 Evaluation by <i>Paulistanos</i> .....	226
8.1.4 Resume.....	227
8.2 The socially assimilated workers (SAW) .....	228
8.2.1 The social network.....	228
8.2.2 Linguistic characterization.....	232
8.2.2.1 Accommodation .....	232
8.2.2.2 L1-transfer.....	234
8.2.2.3 Morphological simplification.....	234
8.2.3 Evaluation by <i>Paulistanos</i> .....	234
8.2.4 Resume.....	235
8.3 The structurally assimilated academics (StAA) .....	236
8.3.1 Social network.....	236
8.3.2 Linguistic characterization.....	239
8.3.2.1 Accommodation .....	239
8.3.2.2 L1-transfer.....	240
8.3.2.3 Morphological simplification.....	242
8.3.3 Evaluation by <i>Paulistanos</i> .....	242

8.3.4 Resume.....	243
8.4 Summing up.....	243
9. Conclusion and outlook.....	247
10. References.....	251
Annex I: Overview of the informants.....	I
Annex II: Transcription protocol.....	IV
Annex III: The questionnaire.....	V

## List of Figures and Tables

---

Figure 2.1: São Paulo according to districts, <i>subprefeituras</i> and zones.....	8
Figure 2.2: São Paulo's Human development index in 2000, according to districts.....	13
Figure 2.3: Analphabetism and functional analphabetism in 2000, according to districts.....	14
Figure 2.4: Average salary of heads of households in SM per circles, city of São Paulo in 1991 and 2000.....	15
Figure 2.5: Average income of employed in Reais (April 2006) according to zones.....	16
Figure 2.6: Rate of heads of households with monthly income of less than three SM (2004) according to districts.....	17
Figure 2.7: Inhabitants living in favelas (2000), in relation to districts' total population in percent..	17
Figure 2.8: Formal working opportunities in São Paulo 2004 according to district.....	18
Figure 2.9: Gated communities with usable area under 150m <sup>2</sup> in São Paulo.....	21
Figure 2.10: Gated communities with usable area above 150 m <sup>2</sup> in São Paulo.....	21
Figure 2.11: Frequented places during leisure time by Paulistanos according to district – comparison between inhabitants of Moema/Consolação and Marsilac.....	22
Figure 2.12: Homicide as cause of death (2005) according to districts.....	24
Figure 2.13: Risk of registered occurrence of homicide (2005) according to districts.....	24
Figure 2.14: Risk of registered occurrence of robbery according to districts.....	25
Figure 5.1: Sequential Exploratory Design applied to this study.....	87
Figure 5.2: All interviews before selection according to research site (N = 70).....	91
Figure 5.3: Number of interviews (NI) according to profession after selection (N=28).....	93
Figure 5.4: NI according to profession and research site (N = 28).....	94
Figure 5.5: NI according to profession and duration of stay in years (N = 28).....	94
Figure 5.6: NI according to year of arrival in São Paulo (N = 28).....	95
Figure 5.7: NI according to age at time of the arrival in São Paulo (N = 28).....	95
Figure 5.8: NI according to age at time of the interview (N=28).....	95
Figure 5.9: NI according to place of origin (N = 28).....	97
Figure 5.10: NI according to gender (N = 28).....	97
Figure 5.11: NI according to gender (N=100).....	108
Figure 5.12: NI according to age in years (N = 100).....	108
Figure 5.13: NI according to place of residence (N = 100).....	108
Figure 5.14: NI according to profession (N = 100).....	109
Figure 5.15: NI according to foreign language competence (N = 100).....	109
Figure 5.16: NI according to residential neighborhood and profession (N = 100).....	110
Figure 5.17: NI according to residential neighborhood and foreign language proficiency.....	110
Table 6.1: Castilian vowel system.....	121
Table 6.2: Brazilian Portuguese vowel system.....	121
Table 6.3: Pronominal system of Bolivian Castilian.....	149

Table 6.4: Pronominal systems of Brazilian Portuguese (formal and informal).....	150
Table 6.5: Verbal paradigm of simple present in ancient PC, current formal and informal Brazilian Portuguese.....	152
Table 6.6: Distribution of investigated features and corresponding linguistic processes.....	178
Figure 7.1: Responses on linguistic perception by Paulistanos according to language level.....	187
Table 7.1: Excerpt from the responses: description of linguistic markers.....	188
Figure 7.2: Number of particular linguistic markers according to speaker.....	190
Figure 7.3: Classification of the language.....	192
Figure 7.4: Origin of speakers / Brazil and foreign countries.....	194
Figure 7.5: Origin of speakers outside Brazil.....	195
Figure 7.6: Origin of speakers inside Brazil.....	196
Table 7.2: Excerpt from the responses: Stylistic characterization of the three speakers.....	197
Figure 7.7: Distribution of the label Portuguese according to PP and PC.....	200
Figure 7.8: Place of residence according to speakers, differences inside São Paulo excluded.....	202
Figure 7.9: Place of residence according to speakers, differences outside São Paulo excluded.....	202
Table 7.3: S II: Cross table Residence of informants according to Bolivian presence * residence in SP (social status).....	203
Figure 7.10: Evaluation of speech examples: Semantic differential about prosperity.....	204
Figure 7.11: Evaluation of speech examples: semantic differential about knowledge.....	205
Figure 7.12: Evaluation of speech examples: semantic differential about effort.....	205
Table 7.4: Correlation between three semantic differentials according to speaker.....	206
Figure 7.13: Evaluation of speech examples according to likeability scale.....	208
Figure 7.14: Evaluation of speech examples according to sociability scale.....	209
Figure 7.15: Evaluation of speech examples according to attractiveness scale.....	209
Table 7.5: One-way descriptives: attractiveness of S I according to informants' profession.....	211
Table 7.6: One-way descriptives: S III: personal characteristics according to informants' age.....	212
Table 7.7: Two-way analysis: Dependent Variable: S II: likeability.....	213
Table 7.8: Tests of Between-Subjects Effects: Dependent Variable: S II: likeability.....	213
Figure 8.1: Prototypical network of one of the NAW (EGO-centered).....	221
Figure 8.2: Prototypical network of one of the SAW (EGO-centered).....	231
Figure 8.3 Prototypical network of one of the StAA (EGO-centered).....	238
Table I.1: Overview of the informants' social characteristics and links to the audio files.....	I
Table I.2: Speaker's attributes indicated by citation in the text.....	III
Table II.1: Explanation of the signs used in the transcription of the interviews.....	IV
Figure III.1: Picture story.....	V
Figure III.2: The online questionnaire (test of language regard).....	V

## List of Abbreviations

---

ADRB	<i>Asociación de los Residentes Bolivianos</i>
CAT	Communication accommodation theory
CC	<i>Castellano Culto</i>
CoP	Community of Practice
CP	<i>Castellano Popular</i>
DPU	<i>Dialectos portugueses del Uruguay</i>
L1	First language
L2	Second language
MMR	Mixed methods research
NAW	Not assimilated workers
NI	Number of informants
NURC	<i>Norma Urbana Culta</i>
PC	<i>Português Culto</i> (São Paulo)
PP	<i>Português Popular</i> (São Paulo)
PU	<i>Português Uruguayo</i>
S I	Speaker I
S II	Speaker II
S III	Speaker III
SAW	Socially assimilated workers
SLA	Second Language Acquisition
SLM	Speech Learning Model
SLU	Second Language Use
SM	<i>Salário mínimo</i>
sn	Stephanie Niehoff, the author
SP	São Paulo
SPA	Socially privileged area
SUP	Socially unprivileged area
StAA	Structurally assimilated academics
ZL	<i>Zona Leste</i>
ZN	<i>Zona Norte</i>
ZO	<i>Zona Oeste</i>
ZS	<i>Zona Sul</i>

## 1. Introduction

---

In the 21<sup>st</sup> century, urbanity dominates the lives of most people worldwide and in particular in Brazil where São Paulo constitutes Brazil's and South America's largest city. Thus, from a sociolinguistic perspective, the social characteristics of cities, in particular their size, density and heterogeneity stemming primarily from migration, must be included into linguistic research. Applied to São Paulo, the goal should be a thorough description and analysis of the migration waves, because the entire history of this megacity is marked by migration resulting in language contact situations, particularly inner-Romance language contact. The linguistic outcome of this process has been little studied. Thus, I aim for closing this gap of investigation by studying the example of the Bolivians' Portuguese, the currently largest immigration group. The questions arising in this context are outlined and will be answered in this study by empirical research that includes the immigrants as well as the receiving society by qualitative and quantitative methodology.

---

We are living in the urban millennium. The majority of the world's population lives in cities<sup>1</sup>; since 2005, there are twenty megacities with more than ten million inhabitants worldwide. In Brazil, the urbanization rate adds up to 84.4% (cf. Portal Brasil 2010), in the federal state of São Paulo the latest census in 2010 even showed that 95.9% of the population live in urban conglomerates (*Instituto Brasileiro de Geografia e Estatística* [IBGE] 2010: 24) and more than 11 million people in the city of São Paulo.<sup>2</sup>

These cities owe their size not to natural growth but to migration. "Meeting the challenges of migration", under this title the United Nations Population Fund (UNFPA) and the International Migration Policy Programme (IMP) published an extensive report to develop policy recommendations, observing the increasing importance of migration: In 2004, 175 million persons are estimated to be living outside their country of origin (cf. Jenny and Obaid 2004: 2).

Due to globalization, migration waves across the border of national states gradually gain in importance. Thus, scientific research that intends to take into account the central questions of our societies' present and future development must focus on the cities and particularly the international migration that leads to the immense rise in the population of the cities. Many sociologists even postulate that today's sociology has to be urban sociology (cf. Ernst 2007), focusing on the life conditions, the

---

<sup>1</sup> The definition of *cities* according to their size differs widely: In Latin America, settlements with more than 2000 inhabitants are considered as cities; in China only settlements with more than 100,000 inhabitants (cf. Korff 2007: 1).

<sup>2</sup> For a more detailed description of the city and the metropolitan region of São Paulo, see Chapter 2.

ways of living together and the integration of immigrants in cities – be it international or national migration. The consequences of migration challenge both the migrants themselves as they leave their original country behind and are confronted with the new environment, facing e.g. cultural and linguistic barriers, and the receiving society, as the newcomers might constitute economic competition, e.g. depressing wages at the labor market or even imposing a threat to the identity of the original society.

For the investigation of the consequences of migration from a sociological as well as a linguistic angle, both sides have to be observed: Do the migrants integrate into the new society, do they preserve their own culture or do they become transnational? Does the receiving society expect assimilation, does it follow a multi-culturalist approach or does it simply ignore the immigrants? In linguistic terms, Coulmas (2005) expresses the challenges of migration this way:

*Immigrant groups find themselves in a new social and linguistic environment to which they adjust in many ways and which adjusts to them. Sociolinguistic evidence reveals a wide range of patterns of mutual adjustment with host communities insisting on assimilation to various degrees and immigrant communities showing variable inclination to yield to assimilation pressure (Coulmas 2005: 147).*

By analogy, sociolinguistics should concentrate on the language contact in cities – intra- as well as transnational, i.e. intra- as well as translingual – to describe, analyze and explain the whole range of possibilities on both sides, e.g. complete assimilation versus total isolation by the immigrant group or high pressure versus ignorance by the receiving society.

## 1.1 Migration to São Paulo

---

The social system of São Paulo is not only defined by the size of its population, but also by the geographical dimension, the population density and the inner differentiation (cf. Hofmeister 1999: 232 ff.), i.e. the heterogeneity of the population regarding lifestyles, social status, cultures and origins (cf. Siebel 1997: 30). Consequently, for the investigation of the migration to the city, it first must be deciphered how these dimensions influence the transnational network. The next step is then to study the results of cognitive assimilation, i.e. for example the learning of the receiving society's language. The city of São Paulo is strongly dominated by national and international immigration: mainly Italian, Portuguese and Spanish at the end of the nineteenth century, Jewish, Lebanese and Japanese immigrants following in the first half of the twentieth century, and, more recently, the Brazilians coming from São Paulo's countryside or Brazil's poor northeast. They all came and brought their

culture, their languages and varieties to turn the city into what it is today, the economic, financial and cultural center of Brazil and South America (cf. Ernst 2007).

From a linguistic perspective, it seems astonishing that people in São Paulo speak almost exclusively Portuguese despite this multitude of different languages having come into the city during the last 120 years. Naturally, the contact with other languages, particularly with other Romance languages such as Italian and Spanish dialects – more than one 800,000 Italians and almost 400,000 Spanish arrived to São Paulo between 1872 and 1952 (Freitas 1999: 44) – influenced São Paulo’s Portuguese, but unfortunately, there are no linguistic studies investigating the language contact situations of that time.

In 2005, I tried to find traces of this inner-Romance language contact in the typical Italian quarters of São Paulo, *Mooca* and *Bom Retiro*, aiming for the description and analysis of the Portuguese of Italians having come to São Paulo in the first half of the twentieth century (Niehoff 2006). Interestingly, I located very few Italian immigrants of the first generation, but new immigrants from Brazil’s Latin-American neighbor countries having come to exactly the same places as the Italian immigrants before. Akin to the immigration before, the outcome of this language contact stays completely unattended by linguistic scholars, only some anthropologists have yet discovered this group that already contains more than 200,000 people in São Paulo.

In order to close this gap of investigation and in an attempt to “use the present to explain the past” (Labov 1974), I decided to study the Portuguese of the current immigrants – native speakers of another Romance language, in this case the Castilian of Bolivia.

## 1.2 Open questions and aims of the study

---

Having the sociological basis in mind various questions stand at the beginning of this study: Do the Bolivians actually learn Portuguese? Are they successful? Does their environment require that they learn Portuguese? What is the target variety of the Bolivians? Do they all speak the same, is there a typically “Bolivian accent”? If so, is it recognized by the *Paulistanos*?<sup>3</sup> If not, do commonalities and differences concern phenomena from all linguistic areas? Which linguistic processes influence the speech of the Bolivians? Which factors influence the linguistic behavior of the Bolivians? How are their social networks organized? Which factors influence the quality of these networks? Are catego-

---

<sup>3</sup> Originally determining all inhabitants of the city of São Paulo “natural ou habitante da cidade de São Paulo” (Dicionário da língua portuguesa 2002: 1257), in this study, the meaning was limited to people who regard themselves as *Paulistanos* and are not Bolivian immigrants themselves.

ries like heterogeneity, size and density important? Do individual categories like the duration of stay or the age at arrival make any difference? How do the *Paulistanos* as the receiving society come into view? Can the Bolivian's linguistic behavior be explained by the perception of the receiving society? Does the close relatedness between Spanish and Portuguese influence the Portuguese of the Bolivians?

In short: **Who speaks Portuguese? In which way? Why?**

In order to answer these questions, the first aim of this study is the thorough recording<sup>4</sup> and description of the Portuguese spoken by the Bolivian immigrants. Since a community of 200,000 persons must be heterogeneous, I want to consider as many different social subgroups as possible, focusing first on the individual and then reintegrating the single speaker “into the overall matrix of the speech community” (Sankoff 2008: 659). This approach is necessary for an exact comparison of the influence of the varying social factors; with any other approach the model would become too imprecise to explain the complexity of linguistic reality.

In addition, I will analyze how the linguistic outcome of this language contact can be interpreted regarding the linguistic processes underlying it. Already existing theoretical approaches on language contact should be applied and if necessary complemented. The focus lies on the social factors possibly influencing the speech of the Bolivian immigrants, in particular on the social networks of the Bolivian immigrants in São Paulo and the interplay between migrant and receiving society. Thus, I also investigate how the Bolivians' speech is perceived and evaluated by the *Paulistanos*. This way, I hope to add to the clarification of the linguistic dimension of international migration – not only of Bolivians to São Paulo, but in general.

As I focus on the outcome of language learning, i.e. the Portuguese of the Bolivians, I do not investigate the Spanish of the speakers. Hence, I cannot take into account individual style shifting in different communication situations in Spanish. Thorough (sociolinguistic) studies of the Bolivians' Castilian remain a strong desideratum, especially with regard to probable Aymara or Quechua structures influencing not only the speech of bilinguals but also of monolingual speakers of Bolivian Castilian.

---

<sup>4</sup> The transcriptions of forty Bolivians speaking Portuguese (1<sup>st</sup> generation) are accessible (cf. Table I Annex I).

Additionally, as my focus lies on the social factors influencing the speech of Bolivians and the linguistic processes influencing it, not every phenomenon will be discussed in every detail, but only as far as it is necessary for the objectives of this study. I hope, however, that my conclusions motivate further scholars to use my data for a more comprehensive discussion of single features.

### 1.3 Outline of the study

---

My objectives require the following outline of the study: In Chapter 2, the historical, socio-geographical, economic and linguistic environment of the Bolivians in São Paulo is illustrated, with regard to the situation in their home country, but first and foremost their life conditions of their new home, the city of São Paulo. Special focus is placed on the inner differentiation of the city that leads to a sociogeographic segregation resulting in almost completely different lifestyles of their inhabitants according to their social status.

In Chapter 3, the state of the art regarding urban sociolinguistics as well as regarding language contact and second language acquisition theories is discussed. Special attention is paid to studies investigating Spanish-Portuguese language contact and holistic approaches that explicitly focus on the social causes and linguistic consequences of migration. Additionally, studies focusing on the varieties of São Paulo and the varieties of the native regions of the Bolivian immigrants as well as the few existing anthropological studies about the Bolivian immigrants in São Paulo will be presented.

In Chapter 4, I resort to the theoretical approaches presented in Chapter 3 to formulate my hypotheses regarding the social factors influencing the Bolivians' Portuguese and on the other hand the linguistic processes that are triggered by these factors. The methodology valid for verifying or falsifying this approach – and any dealing with socio-indexicality – is also discussed.

In Chapter 5, I hence present the methodological design: In order to reach the aims of this study, a sequential exploratory design was chosen, consisting of the main qualitative part focusing on the speech of Bolivian immigrants and a further quantitative part that concerns the perception and evaluation of their speech by *Paulistanos*, the original inhabitants of São Paulo.

In Chapters 6 and 7, the analyses of the two empirical parts are presented. The qualitative analysis of the speech corpus resorts to a great amount of speech examples from the corpus of 28 Bolivian speakers. For the analysis, the linguistic environment is considered as well as the social distribution of single features among the speakers. The quantitative part consists of the analysis of 100 question-

naires by descriptive and inferential statistics, showing the *Paulistanos*' attitudes towards the speech of three Bolivians that were found to be prototypical for the linguistic behavior of the entire Bolivian community.

The following discussion in Chapter 8 picks up the analyses of both empirical parts, relating them to the hypotheses I had formulated before, i.e. the interplay between the qualities of the social networks and the linguistic processes which influence the language and thus lead to different outcomes of the various subgroups' speech. These subgroups are characterized by their social situation – their social network, in particular – by their use of linguistic features, and by their evaluation through the *Paulistanos*.

In the concluding Chapter 9, I sum up the results, answer the questions that had been asked in Chapter 1 and give an outlook on further research.

## 2. The Bolivians in São Paulo

---

In this chapter the historical and socioeconomic background of the Bolivian immigration to São Paulo is illuminated. First I present the city's social characteristics before demonstrating how the Bolivian immigration fits into this bigger picture, because they have significant impact on the linguistic environment of the Bolivians as will be shown in the further chapters.

São Paulo's socioeconomics are characterized by sociospatial segregation regarding the socioeconomic status of its inhabitants. This status is conditioned by the level of education and subsequently the profession, but is also reflected by geographical distribution and participation in urban culture. The different social groups lead completely different lives; the occasions where they could meet are rare. Thus, one could argue that the distinction of social classes is still relevant for the society of São Paulo.

Considering the pull and push factors determining the Bolivian immigration, two different migration lines are classified: Since the 1970s, Bolivian academics have come to São Paulo, often because of political reasons. The more recent mass migration, though, is due to economic differences between sending and receiving country. These disparities inside the Bolivian community are even increased by the social differences inside São Paulo's society.

---

As base for the further analyses of the speech of Bolivian immigrants in São Paulo and the linguistic processes underlying it, their social environment must be understood. São Paulo's heterogeneous society is divided into smaller units that facilitate the classification of the Bolivian immigrants in their new environment. Therefore, not only socioeconomic categories, but the geographical dimension must be taken into account. Considering the historical development the subsequently differing socioeconomic quality of the districts and regions, geographic mobility and social mobility in São Paulo, I get valuable information about the social networks of the city's inhabitants, because these factors even influence the daily routine of the *Paulistanos* – a fact that has to be taken into account for São Paulo's linguistic classification, too.

### 2.1 The city

---

The city of São Paulo extends over an area of more than 1500 km<sup>2</sup>, placed in the metropolitan region of São Paulo with an extension of more than 8000 km<sup>2</sup> (EMPLASA. *Região Metropolitana de São Paulo*). Because of these dimensions, two main aspects need to be addressed: The heterogeneity of the various districts and quarters as historically shaped by immigration and the transport system of the city enabling the inhabitants to cross the extreme distances inside the city. The city is administrated by 31 *subprefeituras* (Prefeitura de São Paulo 2011). In public perception, these administrative units do not play an important role though: People mostly refer to their quarters and to

the superior entity of the *zonas*, differentiating between the *Centro* (the central region), the *Zonas Norte* (the northern zone), *Leste* (the eastern zone), *Sul* (the southern zone) and *Oeste* (the western zone) (cf. Figure 2.1).



Figure 2.1: São Paulo according to districts, *subprefeituras* and zones (Município em Mapas: Panorama. Político Administrativo:1)

The eldest areas are located in the center, the west and northwest of São Paulo:

*A planta da cidade de São Paulo de 1881 já mostrava a expansão da cidade nas direções Oeste e Noroeste, além do núcleo histórico inicial, configurando os bairros burgueses (Verás 1999: 88).*

But shortly afterwards, proletarian quarters in the eastern zone come into existence because of the beginning industrialization of the city. In this era, the mass migration to São Paulo begins: In 1870 it

was a little city of less than 30,000 inhabitants, yet growing rapidly, reaching more than one million inhabitants in 1930 (cf. Vasconcelos 2006: 53)<sup>5</sup>. Today it is the largest and most important urban center of Brazil and South America (Kohlhepp 1997: 137).

### 2.1.1 The historical context: immigration

---

Following Lee's theory of migration (Lee 1966) I differentiate between factors at destination (pull factors) and factors at origin (push factors) that make the individual take the reasonable costs of leaving its native country (or region)<sup>6</sup> behind and going to another, often unknown place.

São Paulo's most important pull factor has always been the economy: After the prohibition of slave trade 1850 and the final abolition of slavery in Brazil 1888, São Paulo's booming coffee plantations were in need of cheap workers (cf. Ianni 1979: 17). The coffee boom created enough capital to promote investments in the industrial development of the region. This starting industrialization of São Paulo led to even more demand of manpower and consequently to the first wave of mass migration to São Paulo during the end of the nineteenth century. According to the *Memorial do Imigrante* in São Paulo, more than 2.3 million European immigrants arrived at the harbor of Santos between 1870 and 1907, among them 1.2 million Italians, but also Portuguese, Spanish, Germans and others (*Memorial do Imigrante* 2011).

As the flow of European immigrants slowed down before the outbreak of the 1<sup>st</sup> World War, governmental policy successfully tried to attract Japanese immigrants: More than 180,000 Japanese came to São Paulo between 1908 and 1940 (cf. *Governo de São Paulo* 2008). Apart from the European and the Japanese, immigrants stemming from the Ottoman Empire, mostly Lebanese and Syrian, as well as eastern European Jewish immigrants came to São Paulo. They settled in the quarter of *Bom Retiro*, which therefore developed into a center for eastern European Jews and the Arab immigrants (cf. Pasternak Taschner 2003: 44), who worked mainly in the textile industry, producing and selling clothes for the growing urban population:

*Quickly, the Syrians and Libanese started to produce clothes they sold, and around 1940, [...] they had strong investments in the ready-made-clothing industry [...]. Another important contribution was done by the jews [...] who had success when they dedicated themselves to the production and commercialization of ready-made-clothes. Their*

---

<sup>5</sup> For an excellent visualization of São Paulo's expansion (city and metropolitan region) between 1882 to 2002, see *Município em Mapas: Expansão Urbana*: 3.

<sup>6</sup> Mostly referring to international migration processes, I will further speak of different countries, but the observations are also valid for intranational migration.

*commercial advancement provided for their good reputation, stimulating the installation of little factories concentrated in São Paulo, in particular in the quarter Bom Retiro (Campos Castro 2009: 21 f., own translation, sn).*

The Italian immigrants, though, had settled near the factories and the railway, as they arrived from São Paulo's harbour of Santos. They formed quarters like *Brás* and *Mooca* in the east of the city center. One witness who came to São Paulo 1949 describes the situation this way:

*Formerly, almost the whole Italian community was gathered here in the quarter Brás. When I came to Brazil, they told me the capital of Italy was the Brás. (cited in Niehoff 2006: 10, own translation, sn).*

In analogy to the ethnic separation, the districts soon got separated according to socioeconomic classes, too:

*São Paulo came to be organized in two different parts along the rivers that constituted actual dividing bands: In the East, the rather working-class and mixed quarters were located, with the homes of the workers, industry and commerce. Its pioneer was the Brás. The West was formed by quarters of the rural aristocracy and the industrial Bourgeoisie in Campos Eliséos, Vila Buarque, Higienópolis e Avenida Paulista (Véras 1999: 88, own translation, sn).*

Thus, even the earliest phases of São Paulo's development show characteristics of a socially divided city, where the distribution of urban space heavily depends on the social status of its inhabitants:

This process was aggravated in further decades, when the international immigrants were substituted by a new migrational wave, this time consisting of national migrants particularly from São Paulo's hinterland and Brazil's poor northeast. They further accelerated São Paulo's rapid growth: From 1950 on, São Paulo's population increased by 600 percent in a period of 30 years reaching more than 12 million in 1980 (Villa and Rodriguez 1996: 29). The *Nordestinos*<sup>7</sup> were pushed to São Paulo by their miserable life conditions caused by drought, bad harvest and the price decline of the agricultural products traditionally produced there, like sugar or coffee (cf. Baptista 1998: 106). Like the European immigrants before them, they hoped to find work in the industrial sector and thus better life conditions in São Paulo, at least for their children.

Interestingly, they also settled in the traditionally Italian quarters like *Brás*, where the majority of the now well established descendants of the European immigrants had already left to districts with better living situations, as described by Véras 1999 (87):

*[The Brás] successful immigrants leave to live in exclusively residential quarters that are less crowded and polluted. New occupants arrive: now national migrants, generally Nordestinos, who settle in the cortiços, the 'cabeças de porco' (pigheads; here: popular*

---

<sup>7</sup> Brazilian expression for: migrants coming from the Northeast of Brazil.

*term for favela, sn) which were adopted for collective use, dense, and as precarious and promiscuitive as at the beginning of the 20th century (own translation, sn).*

But the old proletarian quarters in the center of the city did not offer enough space to house the *Nordestinos* and their children. Thus, the city area expanded towards more distant areas: Following the expansion of the railway and the road system<sup>8</sup>, the big factories left the city center. Many of the mostly unskilled new workers also tried to occupy cheaper space in the peripheral areas, building *favelas* and living in miserable conditions:

*The precarious situation amplified: In 1975, it was estimated that 450.000 peripheral housing units were self-constructed. [...] In 1980, the population of the city reached 8.587.665 inhabitants. 25% of them were analphabets and 8.5% gained less than one minimal wage. The 3.500 clandestine housing parcels already occupied 34.000 hectares of the 60.000 hectares of the urban area. [...] The 1.086 favelas accommodated 166.040 barracks where 813.596 favelados<sup>9</sup> lived (Vasconcelos 2006: 56; own translation, sn).*

Thus, the city develops into a metropolitan region, where people with little income live in the periphery in poor living conditions, in so called *favelas* and *cortiços*. From 1973 until 1991, the number of people living in *favelas* multiplied, reaching more than 9% of the entire population of the metropolitan area, i.e. more than 900.000 people (cf. Saraiva and Marques 2007: 16). In the *Zona Leste*, the peripherization of the population in the 80s and 90s leads to a population explosion with growth rates of 101.9% between 1980 and 1991. In the year 2000, the population of the peripheral areas represents 47% of the total population. At the same time, the area of high concentration of upper income classes loses inhabitants. In absolute numbers, this signifies the additional amount of one million new inhabitants in the peripheral areas vis-à-vis a reduction of 265,000 people in the central areas (cf. Bógus and Taschner 2004: 73).

In the last decades, the influence of rural-urban migration on the growth of Brazilian cities is declining; between 1990 and 2000 net migration only accounted for 35% (cf. Rodríguez and Busso 2009: 120).

Despite all problems, though, São Paulo's economic opportunities still attract people from other countries; the international immigration became more important again. Since the 1960s, Korean immigrants had come to São Paulo, currently constituting a community of 250,000 people. Most of them work in the textile industry. Initially, the Koreans had been contracted as illegal and hence

---

<sup>8</sup> The building of the big roads following the *Plano de Avenidas* in the 1930s (cf. Bógus and Vêras 2000: 85), had resulted in the city consisting of concentric rings and its radial roads, the *radiais*.

<sup>9</sup> Brazilian term for inhabitants of a *favela*, i.e.: housing areas that were built illegally and therefore have little or no infrastructure etc.

cheap labor for the textile industry then dominated by Arabs and Jews (cf. Cacciamali and Gomes Azevedo 2006: 135). As they themselves started to overtake the little factories and sweatshops in *Bom Retiro*, they soon required further cheap labor, first *Nordestinos*. Yet, protected by Brazilian work legislation, these workers did not accept the hard working conditions and thus became too expensive (cf. Cacciamali and Gomes Azevedo 2006: 137). Therefore, the Koreans started to employ illegal immigrants from Hispano-America, consequently accelerating and amplifying the Bolivian mass migration to São Paulo.<sup>10</sup>

### 2.1.2 The socioeconomic context: segregation

---

The historical development of São Paulo led to ethnic and geographic segregation that soon also reflected socioeconomic separation, “transforming the segregation regarding housing in one of the principal characteristics of the current urban-metropolitan order” (Ribeiro 2008: 13, own translation, sn).

According to Rolnik and Frúgoli (2001: 46), the sociospatial distribution “reflects the history of territorial exclusion that took place in the city of São Paulo [...]” (own translation, sn). They refer to various processes separating the urban space: the deindustrialization of the city<sup>11</sup>, the expulsion of people with less income to more and more distant peripheries and the homogenization of the districts of upper income classes in the southeastern zone, dislocating also the economic city center of a now tertiarized city from the *Avenida Paulista*<sup>12</sup> to the region of *Avenida Brigadeiro*, *Faria Lima*, *Marginal do Rio Pinheiros* and *Avenida Luís Carlos Berrini* (cf. e.g. Villaça and Zioni 2005: 29).

Correspondingly, Figure 2.2 shows the enormous differences with regard to the Human development index between districts in the periphery and the districts in the southwestern zone, the area of high concentration of upper income classes:

---

<sup>10</sup> For more detailed information about the Bolivian immigration, see 2.2.

<sup>11</sup> Between 1980 and 1990, the urban industrial sector decreased from 36% to 22% (Rolnik and Frúgoli 2001: 47); only sectors depending on the huge consumer’s market of São Paulo like food or textile industry stay in the more central regions (l.c.: 48).

<sup>12</sup> Already in the 50s, the economical center had moved from the historical city center to *Avenida Paulista*.

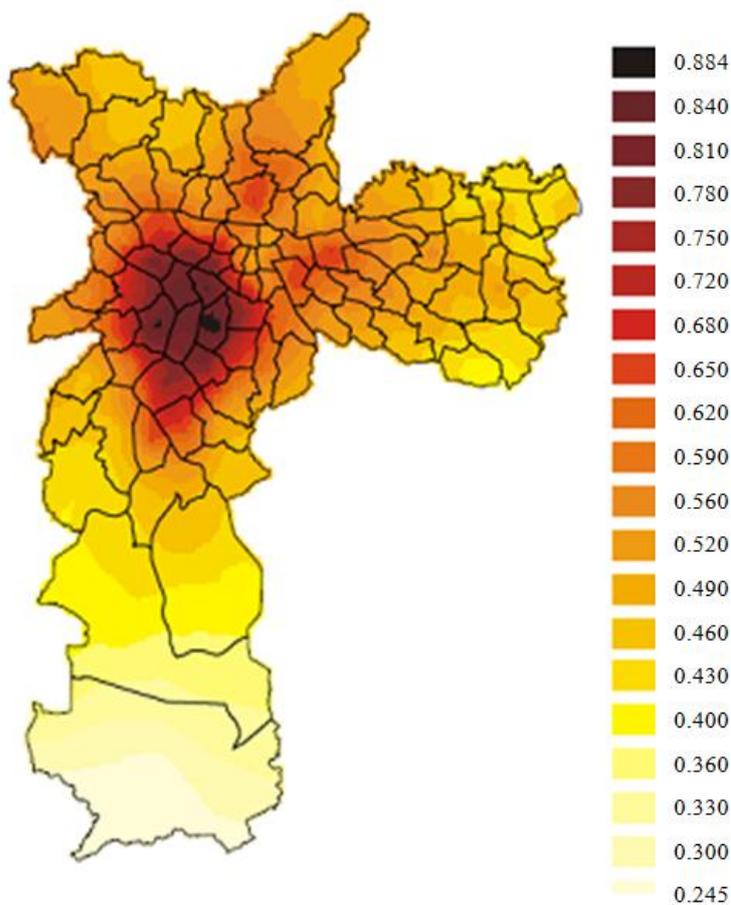


Figure 2.2: São Paulo's Human development index<sup>13</sup> in 2000, according to districts (Município em Mapas: Índices sociais. Índice de Desenvolvimento Humano IDH: 2).

Taking into account data about the economic situation, educational level and health of the *Paulistanos*, the districts in the extreme south have an index of less than 0.300, indicating a situation worse than e.g. in Afghanistan in the year 2011. At the same time, in *Moema*, São Paulo's highest value of 0.884 equals countries like Italy. Central districts like *Sé*, *Brás* and *Parí* though, display rather low indices between 0.400 and 0.500 despite their central placement.

Thus, the districts where the European workers and *Nordestinos* settled in the past and where the Bolivian immigrants live nowadays, display markedly lower development levels than the southwest. These data are confirmed by Figure 2.3, showing the concentration of analphabets and functional analphabets in the city. Similar to Figure 2.2 the number of analphabets in most peripheral regions is more than four times higher than in the southwestern districts. Central districts like *Brás*, *Sé*, *Cam-*

<sup>13</sup> For the calculation of this index, the authors combined various indices: the head of households' salary, the child mortality rate and the head of household's alphabetization rate in combination with average years of schooling (cf. Município em Mapas: Índices sociais. Índice de Desenvolvimento Humano IDH: 2).

*bucí* and *Belém*, however, also display considerably high numbers of analphabets and particularly functional analphabets.

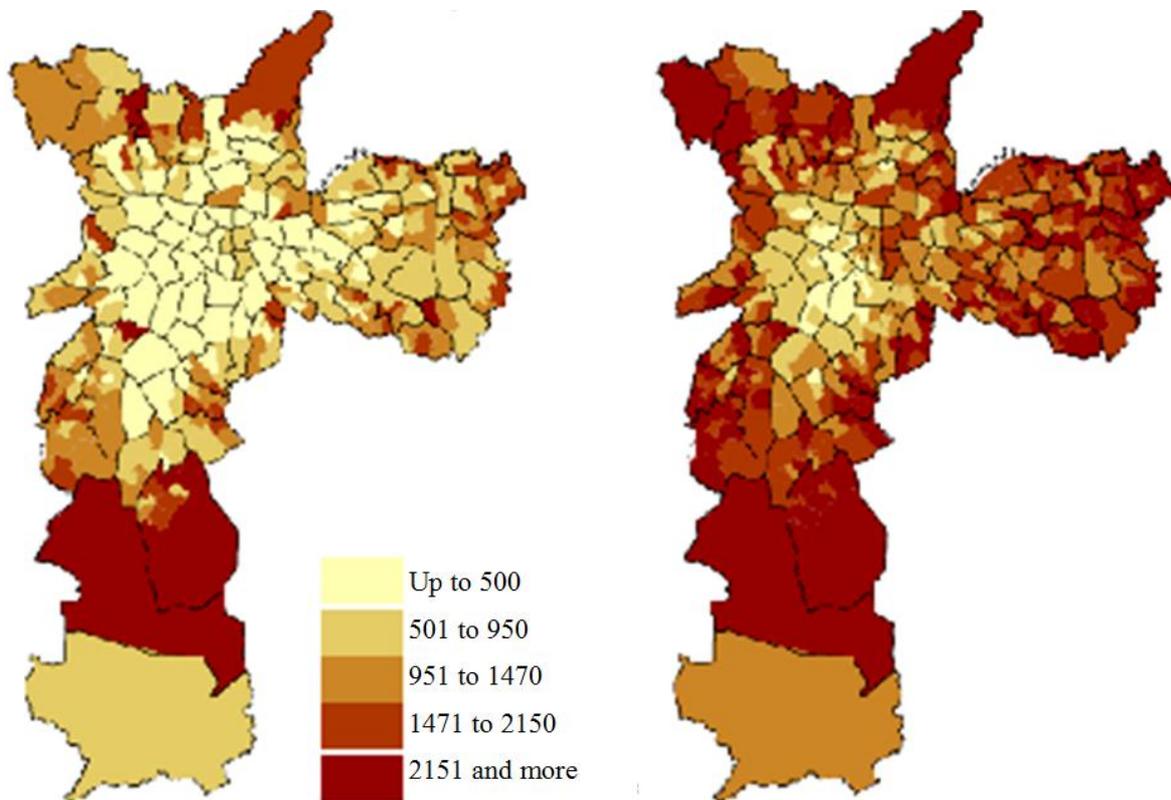


Figure 2.3: Analphabetism (left) and functional analphabetism (right) in 2000, according to districts (cf. *Município em Mapas: Analfabetismo. Analfabetismo no Município: 4*).

The gap between central and peripheral areas is further underlined by the method of Bogus and Taschner (2004) that divides the city into circles: Comparing the average salaries of heads of households in 1991 and 2000 as done in Figure 2.4, there is the clear division between the inner (central, interior and intermediary) and the outer circles (exterior and peripheral). In 1991, the average salary in the interior circle was more than two times higher than in the peripheral circle. As the average salary decreases in the peripheral region until the year 2000, but raises in the inner circles, this gap widens during this period reaching a quotient of 2.79.

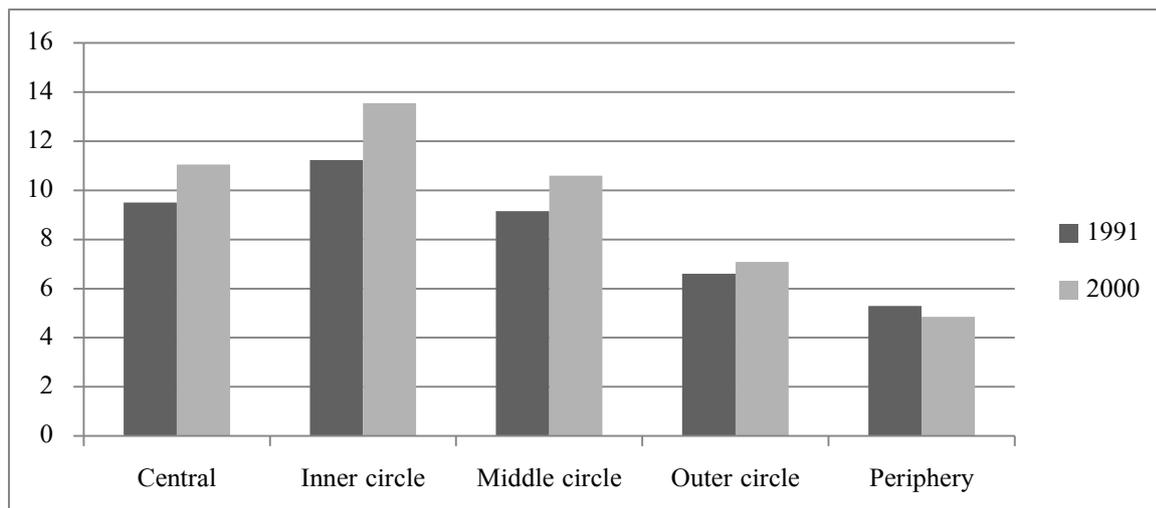


Figure 2.4: Average salary of heads of households in salários mínimos (SM) per circles, city of São Paulo in 1991 and 2000 (cf. Bogús and Taschner 2004: 56)

However, the unequal social distribution between center and periphery does not tell the entire truth. It must be considered that e.g. the inner circle contains socioeconomically such different quarters like *Sé* and *Jardim Paulista*. Therefore, more detailed classification about districts and regions is needed such as in Figure 2.5 and Figure 2.6: Here, the authors divide the city in eight zones<sup>14</sup>, reflecting not only the geographic *Zonas* presented before, but also differences inside the zones.

<sup>14</sup>*Leste 1: Água Rasa, Aricanduva, Artur Alvim, Brás, Belém, Cangaíba, Carrão, Mooca, Pari, Penha, São Lucas, Sapopemba, Tatuapé, Vila Formosa, Vila Matilde, Vila Prudente.*

*Leste 2: Cidade Líder, Cidade Tiradentes, Ermelino Matarazzo, Guaianases, Iguatemi, Itaim Paulista, Itaquera, Jardim Helena, José Bonifácio, Lajeado, Parque do Carmo, Ponte Rasa, São Mateus, São Miguel, São Rafael, Vila Curuçá, Vila Jacuí.*

*Sul 1: Campo Belo, Campo Grande, Cursino, Ipiranga, Jabaquara, Moema, Sacomã, Santo Amaro, Saúde, Vila Mariana.*

*Sul 2: Campo Limpo, Capão Redondo, Capela do Socorro, Cidade Ademar, Cidade Dutra, Grajaú, Jardim Ângela, Jardim São Luís, Marsilac, Parelheiros, Pedreira, Vila Andrade.*

*Oeste: Alto de Pinheiros, Barra Funda, Butantã, Itaim Bibi, Jaguará, Jaguaré, Jardim Paulista, Lapa, Morumbi, Perdizes, Pinheiros, Raposo Tavares, Rio Pequeno, Vila Leopoldina, Vila Sônia.*

*Norte 1: Jaçanã, Mandaqui, Santana, Tremembé, Tucuruvi, Vila Guilherme, Vila Maria, Vila Medeiros.*

*Norte 2: Anhangüera, Brasilândia, Cachoeirinha, Casa Verde, Freguesia do Ó, Jaraguá, Limão, Perus, Pirituba, São Domingos.*

*Centro: Bela Vista, Bom Retiro, Cambuci, Consolação, Liberdade, República, Santa Cecília e Sé.*

Figure 2.5: Average income of employed in Reais (April 2006) according to zones  
(Município em Mapas: Economia Urbana. Mercado de Trabalho: 5).

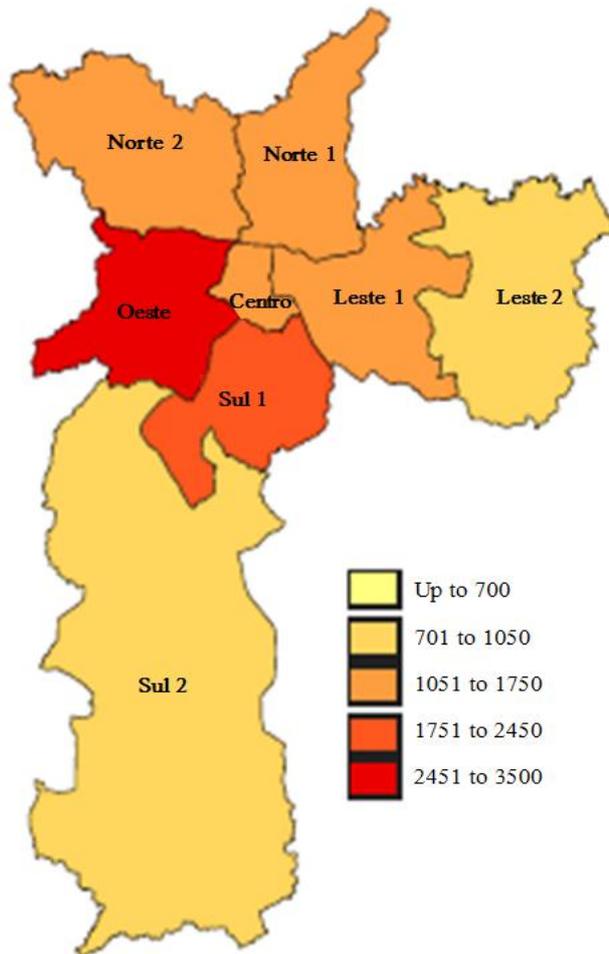


Figure 2.5 confirms the average income differences between central and peripheral regions as seen above. People living in the *Zona Oeste* earn approximately three times more in comparison to people living in *Zona Sul 2* and *Leste 2*. Moreover, the average salaries in the *Centro* itself and in *Zona Leste 1*, including districts of the former inner circles, are only half of those in the *Zona Oeste*, although these data only concern people actually being employed. Figure 2.6 includes also the unemployed, confirming the results above and showing at the same time that the zones are actually quite homogenous; only the districts *Santana* and *Tucuruvi*, belonging to *Zona Norte 1* in Figure 2.5, are clearly more heterogeneous: Here, only 20% to 30% of the districts' population earn less than three SM, but as a zone they get the same rating as *Zona Norte 2*, although the rate of head of households with little income is considerably higher there.

Figure 2.6: Rate of heads of households with monthly income of less than three SM (2004) according to districts (cf. Município em Mapas: Índices sociais. Mapa da vulnerabilidade social: 4).

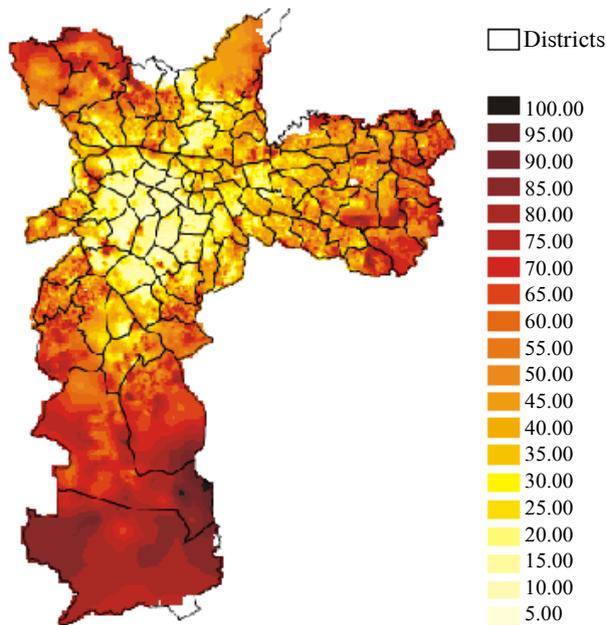
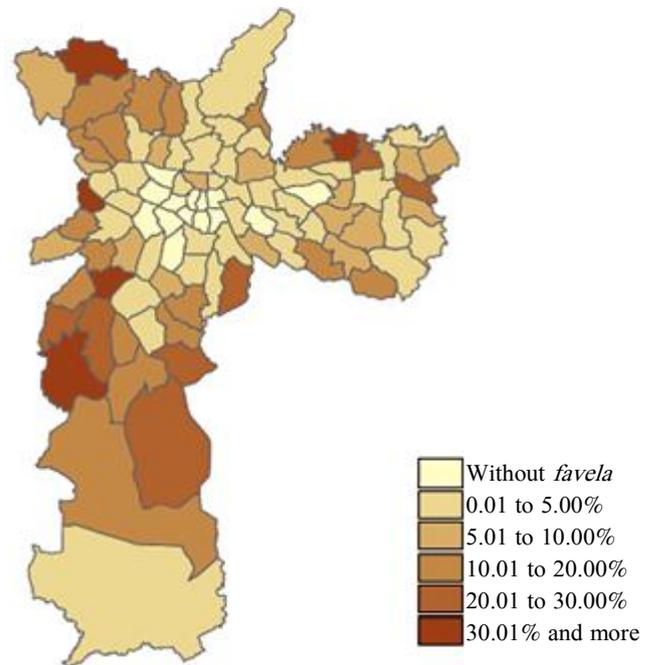


Figure 2.7: Inhabitants living in favelas (2000), in relation to districts' total population in percent (cf. Olhar São Paulo: Contrastes Urbanos. Território e Desigualdades Sociais: 55).



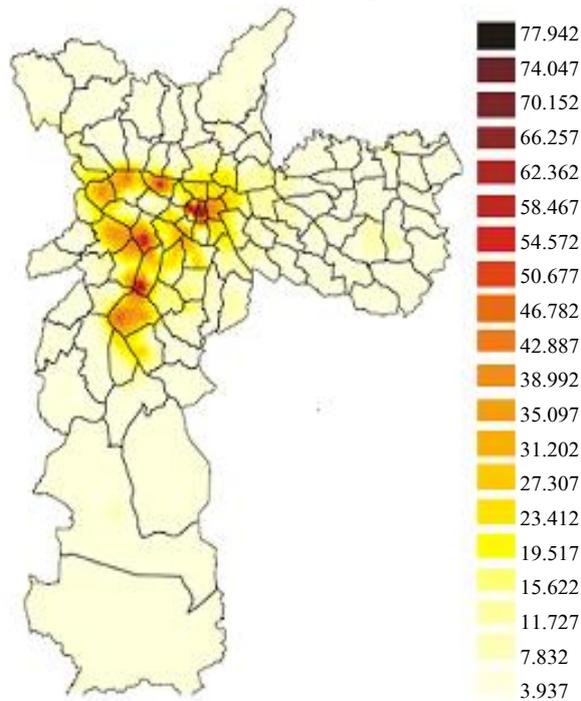
The economic situation and the salary differences are also reflected in the housing situation of São Paulo's inhabitants. Taschner and Bógus (1999: 38) claim that São Paulo is a city "[...] where almost half of its inhabitants live in a clandestine or precarious way". This means, people live in *favelas* or *cortiços*, in illegal and irregular conditions on little space narrowly together, and the houses are constructed on small, propitious parcels of land (cf. Ribeiro 2008: 13). Official data assume that in 2003, 3.4 million inhabitants of São Paulo live in precarious situations (Olhar São Paulo: Contrastes Urbanos. Território e Desigualdades Sociais: 55). Interestingly, only the most ancient central districts including rather poor quarters like *Sé* and *Brás* have no *favelas*, but in rather rich southwestern neighborhoods like *Morumbi* there are the districts of *Vila Andrade* and *Vila Sônia*, that have high rates of favelization.<sup>15</sup>

Apart from income differences, Figures 2.5 and 2.6 also reflect the job opportunities of the different zones. Economically, São Paulo's profile is characterized as "multisectoral economy dominated by service" (*Município em Mapas: Economia Urbana. Distribuição Espacial da Atividade Econômica: I*, own translation, sn), because 46.3% of the companies belong to the tertiary sector, 39.4% to com-

<sup>15</sup> The close proximity to the rich districts is obviously one of the reasons for the development of *favelas* here, because people living there work e.g. as nannies in the quarters like *Morumbi*.

merce, and only 11.9% form the primary sector, previously the economic motor of the city. In analogy, the central districts offer most working opportunities, particularly *Sé* and *Republica*, and along the *Marginais Pinheiros* and *Tietê*, i.e. the districts of *Pinheiros*, *Itaim* and *Santo Amaro* in the southwest and *Lapa*, *Barra Funda* and *Vila Leopoldina* rather in the northwest (cf. l.c.: 3).

Figure 2.8: Formal working opportunities in São Paulo 2004 according to district (cf. Município em Mapas: Economia Urbana. Distribuição Espacial da Atividade Economica: 3).



The clear areas in the wealthiest districts, though, indicate pure residential areas. They are located in close neighborhood to the most productive centers, e.g. *Jardins* between *Pinheiros* and *Itaim*, but nonetheless are strictly separated:

*[...] [T]hese spaces of production (concentration of employment in the productive sector) and consumption (residencies of the middle and high class) by prosperity seem to lie together considering the entire city; however, they do not mix (Município em Mapas: Economia Urbana. Distribuição Espacial da Atividade Economica: 3, own translation, sn).*

The eastern zone on the other hand experiences industrial deconcentration and reconversion, and the dispersion of little industries in different districts accompanied by strong precarization of working opportunities (Rolnik and Frúgoli 2001: 56).

The textile industry with many small sweatshops working with less than ten sewing machines and ten workers would be an example of the little industries that still can be found in this area. Thus, in the *Zona Leste 1*, the textile industry constitutes one of the five most important economic branches of the area, as 8.2% of the official workers work here (cf. Município em Mapas: Panorama. Economia

Urbana).<sup>16</sup> In total, “mesmo se for considerado o total de empregos, o Vetor Leste tem 3.6 habitantes por emprego, contra 1,9 no Vetor Sudoeste” (Villaça and Zioni 2005: 11).

Naturally, this converse distribution of job opportunities and most populous districts signifies a major challenge for São Paulo’s transport system, particularly because the heads of households’ salaries are so low in the poorer zones (see Figure 2.5 and 3.6) that on average two or three persons have to work to earn the family’s living (Villaça and Zioni 2005: 11 and 37).

In accordance, approximately 75% of all journeys by public transportation are made by persons whose average income consists of less than 20 SM (i.e.: 62). Thus, São Paulo’s transport system is chronically overcrowded. Politics privileging automobile traffic (cf. Vêras 1999: 75) have led to rapidly growing numbers of cars on São Paulo’s streets and to daily traffic chaos, reaching more than 300 km of traffic jam inside the city every day (cf. Estadão 2009, June 10). Public transportation, though, also relies mainly on buses, because the railway system (trains and subway) is only little developed. The subway, for example, consists of little more than 60 kilometers and is thus the most jammed subway worldwide; transporting an average of 11.5 million passengers per kilometer (cf. Dantas 2011).

The two lines connecting the populous regions in the periphery with the central regions are particularly crowded. In 2009, the *Linha 1 Azul* connecting north and south was used by more than 900,000 people every working day; the *Linha 3 Vermelha* between *Zona Leste* and *Zona Oeste* was used by more than one million people per working day, corresponding to 8.4 resp. 9.8 passengers/m<sup>2</sup> in 2010 (Amaral 2011), a number easily surpassing the recommendations of the World Health Organization of six passengers/m<sup>2</sup>. Additionally, public transport is expensive: In 2011, a single ticket for the subway cost 2.90 *Reais*, and the bus fare with 3.00 *Reais* was even more expensive. Therefore, despite the traffic jam, most *Paulistanos* use the car. In the metropolitan region, individual rides represent 52% of all daily journeys (cf. Villaça and Zioni 2005: 43). Further 26% of the households with an income of up to two SM cannot afford public transport and have no car, therefore “trocaram o ônibus pelo par de tênis” (Ribeiro 2008: 12), i.e. they simply have to walk. Another 13% switched to the bicycle as a means of transportation. Knowing the territorial expansion of São Paulo, this means that these people are excluded from geographic mobility and hence from urban life.

---

<sup>16</sup> Due to the illegal status of most sweatshops, these data only show the concentration of the legal textile industry there and not necessarily the absence of illegal sweat shops in other regions.

Therefore, traffic and public transportation constitute factors that must be considered as influential on the consistence of the individual networks: Taking into account the time *Paulistanos* spend in transit getting to and from work or school – in 2010 the average *Paulistano* spent 109 minutes every day only for this journey, and even 162 minutes for all daily journeys (cf. Macedo 2010) – the possibilities to meet people and maintain personal contacts outside the working place are limited to the weekend.

Additionally, the costs caused by the long distances between residential and working areas are enormous if the time spent in traffic was counted as productive:

*In the metropolises of São Paulo and Rio de Janeiro, where the distances between the peripheral quarters and the center are the biggest, this potential loss corresponds [...] to 26% of the total wages owned by the workers (Ribeiro 2008: 13, own translation, sn).*

To live near the working place is thus often a privilege of wealthy people in the southwestern areas. Additionally, people of lower income classes were attracted by the possibility to live near the wealthy houses and started to build *favelas* not only in the periphery but also in central areas where they could find open space (Bógus and Taschner 2004: 26). In reaction to the precarization, though, the verticalization in the inner districts of the city and the boom of the gated communities started in the 1990s. Born first and foremost as reaction to crime rates and a feeling of insecurity, the number of gated communities increased steadily, reaching 1053 units in the metropolitan area in 2004. As Figures 2.9 and 2.10 show, apart from the *Zona Sul 2* (cf. Figure 2.2), gated communities are found in every area but the historical center. In the richest quarters of the southwest, though, there is a high density of gated communities which have a usable area is bigger than 150 m<sup>2</sup>, i.e. one more indicator for the wealth of its population and their need for security.

Figure 2.9: Gated communities with usable area under 150m<sup>2</sup> in SP (Loffredo D'Ottaviano 2006: 11).



Figure 2.10: Gated communities with usable area above 150 m<sup>2</sup> in SP (Loffredo D'Ottaviano 2006: 11).



At the same time, this voluntary separation of middle and higher classes signifies a new segregation in the city, the (semi)-privatization of public space and the fragmentation of the city:

*The old duality between centro-periferia is dissolving, to give place to a new one: safe against secure places (Rolnik 2008: 11, own translation, sn).*

Thus, although the contrast between periphery and center might become less striking, the separation between the different classes continues. The space for contact between the classes as it would be expected for an urban conglomeration is more and more reduced:

Superposing the centro-periferia pattern, the recent transformations create space in which the different social groups are often very close, but separated by walls and security technology. They do rather not move or interact in the same circles or areas (Caldeira 2000: 211, own translation, sn). Apart from the socioeconomic status, space and geographic mobility, factors like leisure time and cultural traditions also play an important role in forming the *Paulistanos'* Community of Practice(CoP)<sup>17</sup> and social networks. Sketching a typical day of an employee in São Paulo, she has to go to work in the morning, spending an average of almost 1.5 hours to get to work and another 1.5

---

<sup>17</sup> For more information on this concept, please see 3.1.1.

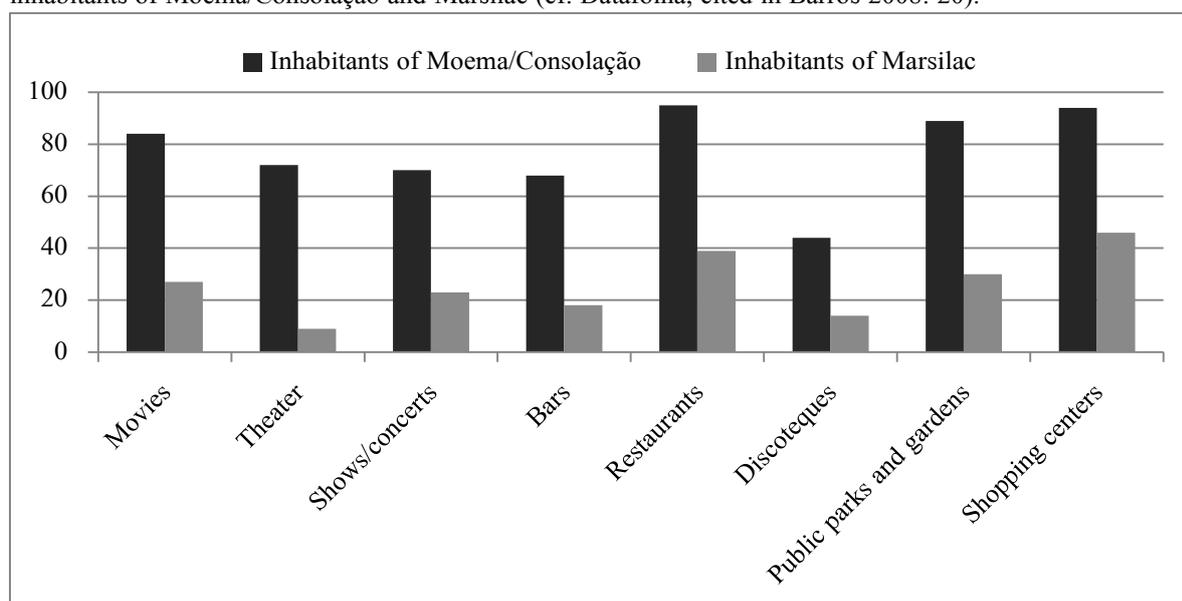
hours to go home (Macedo 2010) in the car or on public transport, which both are spaces where little communication happens. Working approximately 41.9 hours a week, this means that the average leisure time is actually restricted to the evenings and the weekends, where many prefer staying at home:

*There is an increased number of persons who stay at home in their leisure time because of the difficulty of locomotion, the exhaustion caused by work and by the new technologies of electronic diversion like videogames and internet (Barros 2008: 21, own translation, sn).*

While the latter factors cannot be applied to the lower classes as only 43.9% of the *Paulistanos* have access to the internet, clearly excluding people with low level of education (cf. Instituto Brasileiro de Geografia e Estatística 2008), the exhaustion because of work and traffic is valid for most *Paulistanos*.

The difference between the socioeconomic classes can also be seen with regard to their cultural participation: The higher classes can participate in São Paulo's cultural life, because they live near cultural institutions (cf. Barros 2008: 19). As Figure 2.11 shows, the cost of most cultural events and of transportation as well as the differences in regard of cultural education cause additional separation considering the *Paulistanos'* activities during leisure time. Here, the customs of the inhabitants of one of the wealthiest districts in the southwestern zone are compared with those of inhabitants living in the extreme south of São Paulo:

Figure 2.11: Frequented places during leisure time by Paulistanos according to district – comparison between inhabitants of Moema/Consolação and Marsilac (cf. Datafolha, cited in Barros 2008: 20).



We see that inhabitants living in *Marsilac* do not participate in urban cultural life to the same extent as the inhabitants of *Moema* do. Neither do they go to the movies or the theater, nor to bars and restaurants. Public parks and gardens are also mainly frequented by inhabitants of the rich district. The only place where almost half of the people of *Marsilac* go, too, is the shopping center. But as Rolnik and Frúgoli (2001: 49 ff.) point out, the shopping centers are also socially divided. Contact between the various social groups is not to be expected here. Thus, urban culture is only for the higher classes, the lower socioeconomic classes cannot participate here:

*In a city divided between the part that is rich, legal and infra-structured and the part that is poor, illegal and precarious, the population that is in the disfavorable situation ends up having very little access to the economic and cultural opportunities the urban environment offers. The access to the territories with the best conditions of urbanity is exclusive for those who are already part of it (Rolnik 2008: 10, own translation, sn).*

This means also that people of different socioeconomic classes live different lives in the city of São Paulo. Apart from contacts at the working place – where communication with hierarchically lower-ranked colleagues is mostly rather limited –, in the streets, where people sell goods in a traffic jam, or in restaurants where lower class people work as waitresses, there is almost no place and time, where people from higher classes come into contact with lower classes and communicate with them. This segregation is accepted and even promoted by the higher classes, mostly because of their sense of insecurity and endangerment, while the lower classes do not have the possibilities for geographic or social mobility: “The conglomeration among those on equal terms which are the same for the rich and the poor is voluntary for the rich and involuntary for the poor” (Loffredo D'Ottaviano 2006: 3).

This means that crime and violence play an important role not only for the *Paulistanos*' choice of housing but also for their and the Bolivians' social networks and living conditions.

Considering the official crime statistics of São Paulo's administration, the wish of many *Paulistanos* to live behind fenced walls becomes comprehensible. According to the statistics, despite a decrease of 69.4% between 1997 and 2007 (cf. Waiselfisz 2010: 26), there were still 1927 murder victims in the capital in 2007, corresponding to a rate of 28.3 resp. 17.4<sup>18</sup>/100,000 (cf. l.c.: 24). In 2009, 2635 persons died because of violence inflicted by others<sup>19</sup> (cf. Benites 2010). Considering the distribu-

---

<sup>18</sup> In comparison to other federal capitals of Brazil, São Paulo has the second lowest homicide rate only beaten by Palmas, the capital of the rather little populated Tocantins (12.8) (cf. Waiselfisz 2010: 25).

<sup>19</sup> Technically, death because of violence inflicted by others does not equate homicide but constitutes a further category of crime statistics.

tion of homicides in the city, Figures 2.12 and 2.13 show that in 2005 homicide was one of the three most common causes in the peripheral regions, *Zona Sul 2*, *Zona Leste 2* and *Zona Norte 2*, but also in the central districts *Sé* and *Brás*; in the extreme south and the districts *Vila Andrade* and *Campo Limpo*, both dominated by *favelas*, it even constitutes the most common cause of death. Figure 2.13 underlines that again particularly in the north, the extreme east, the area around *Vila Andrade* and also in the old center of São Paulo, the risk of becoming registered as a victim of homicide is extraordinarily high.<sup>20</sup>

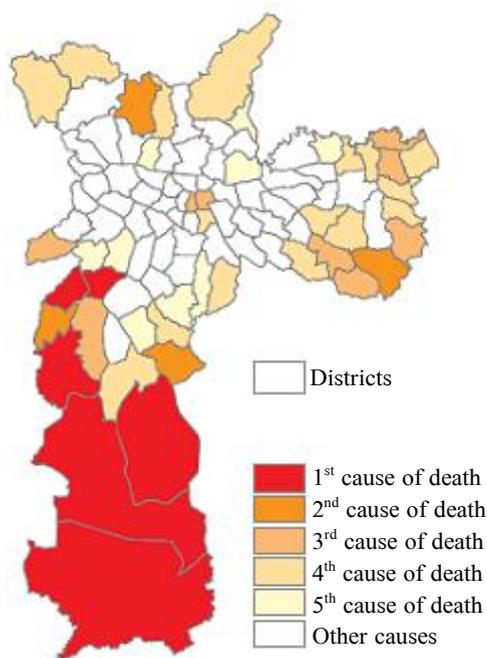


Figure 2.12: Homicide as cause of death (2005) according to districts (Olhar São Paulo: *Violência e Criminalidade*: 58).

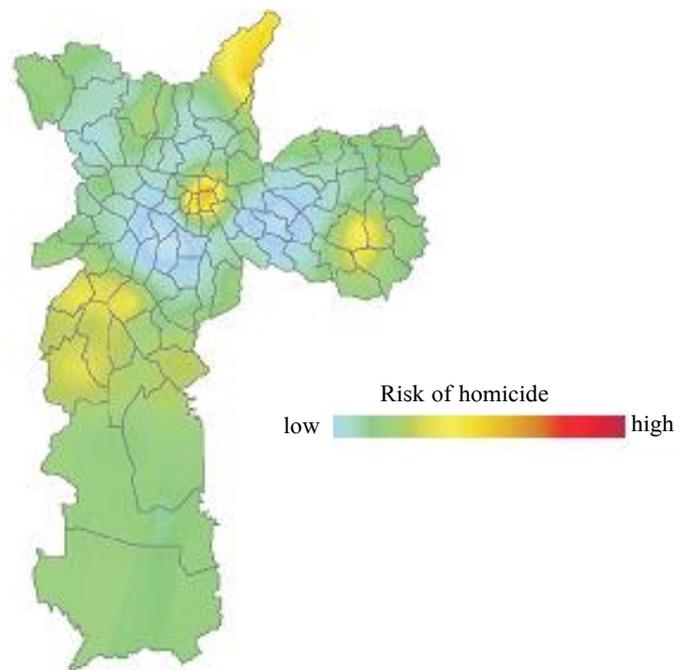


Figure 2.13: Risk of registered occurrence of homicide (2005) according to districts (Olhar São Paulo: *Violência e Criminalidade*: 12).

In 2009, the district of *Brás*, where many Bolivians live, headed the crime statistics considering the most targeted group, men between 15 and 44 years old with 140 violent deaths (cf. Benites 2010). Scholars explain this fact by the economic and social situation of the central quarters:

*In the central areas of the city (Pari, Bom Retiro, Sé, Belém), the highest risk of homicide seems to be associated with the explosive mixture of economic activities (formal and informal, legal and illegal), combined with the concentration of collective housing units*

<sup>20</sup> The reason for the rather low risks of the extreme south and northwest, standing in contrast to Figure 5.9, are explicable by the fact that these numbers show not only high rates of homicide but also the probability that the murder is registered by the police, and in these regions, the police is either not present or not trusted enough to be called in the case of murder (or robbery, see Figure 2.14) (*Olhar São Paulo: Violência e Criminalidade*: 15).

*with high rates of subnormality. These contexts foster conflicts which can have fatal results (Olhar São Paulo: Violência e Criminalidade: 12, own translation, sn).*

In the wealthy districts like *Alto de Pinheiros* or *Moema*, however, where many gated communities are situated, the risk of homicide is extremely low. But studies showed that inhabitants there feel more insecure (21% say that they are afraid of violent crimes), although the risk of violence is rather low (9%), while *Paulistanos* living for example in *Capão Redondo*, having the second highest risk of violence (30%), only 20% feel insecure (cf. Bedinelli 2008).

Considering crimes like robbery or theft, though, the distribution is different as Figure 2.14 demonstrates.

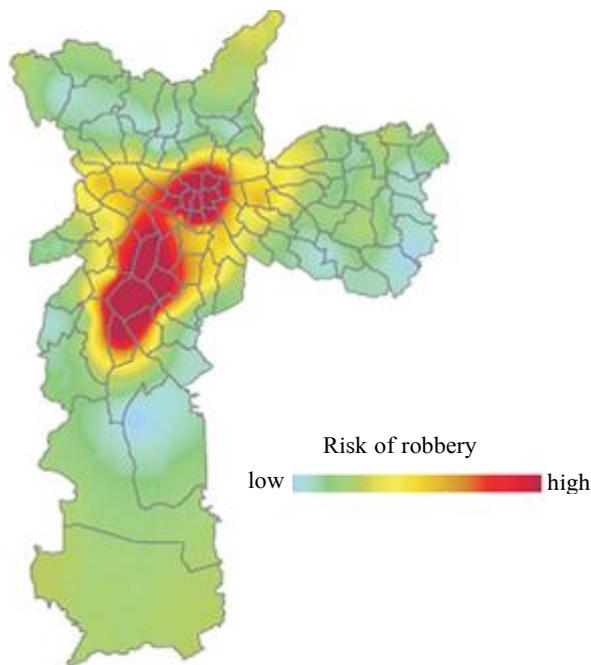


Figure 2.14: Risk of registered occurrence of robbery according to districts (Olhar São Paulo: Violência e Criminalidade: 15).

In correspondence to the wealth of its population, the risk of getting robbed in the southwestern area is particularly high while it is quite low in the peripheral regions with rather poor population. In the old center, however, including districts like *Brás*, *Parí* and *Bom Retiro*, the risk of robbery is also extremely high despite the low income of most people there – this fact again is due to these districts being the commercial center of São Paulo:

*The immense flow of persons in the central commercial area also increases the opportunity for committing crimes against property, even though the income in the centro is not that elevated (Olhar São Paulo: Violência e Criminalidade: 16, own translation, sn).*

According to the scholars, differences between homicide and robbery statistics are due to the motivation of the different crimes: While crimes against persons are poverty driven, crimes against property are opportunity driven (cf. Olhar São Paulo: Violência e Criminalidade: 16).

Thus, important aspects for the qualities of social networks and the importance of socioeconomic class can be deduced from the crime rates in São Paulo: Violence or even the feeling of insecurity in the public reduces the possibilities of contact in the city, because the streets and quarters dominated by violence are not used as public space any longer. Particularly the population of the middle and higher classes are too afraid to spend time there. Since people prefer semi-private places like gated communities or shopping centers, where they feel more secure, this means that the social contacts are again restricted, resulting in homogeneous close-knit individual networks.

## 2.2 The Bolivians

---

Though Bolivia is rich regarding natural resources like fossil fuels<sup>21</sup>, the country's human development index is 0.643, i.e. it takes the 95th rank out of 160 countries with comparable data. In comparison with the other countries in South-America, Bolivia is placed below the regional average (current average Human Development Index: 0.706 (cf. UNDP International Human Development Indicators). 65% of the population still lives below the national poverty line (Pampuch and Echalar Ascarrunz 2009: 186), having a per capita income of around 880 US-Dollars per year. Thus, the economic situation plays a decisive role in the decision of young Bolivians to leave their country behind and help their family by remittances.<sup>22</sup> In addition, young academics also come to Brazil to work as doctors or engineers, constituting a different group of immigrants. The characteristics of the two groups will be elaborated in the next chapters regarding historical, socioeconomic and linguistic aspects.

### 2.2.1 The historical context

---

Before the mass migration that only began at the end of the twentieth century, Bolivian immigration to São Paulo had taken place since the 1960s – on a smaller scale than today.

*Mais ou menos da época das sessenta/setenta aqui era um paraíso para os bolivianos para os poucos bolivianos que tinha. É que não tinha muitos bolivianos em primeiro lugar... então tinha mais campo para eles trabalharem na área da confecção. A partir dos anos 90 começou a encher mais<sup>23</sup> (B7 (40) m. 1986 J. Ce: 1.46).<sup>24</sup>*

At this time, Bolivia's political situation was particularly complex, including military coups, governmental overthrows and civil war-like conflicts in the whole country (Pampuch and Echalar Ascarrunz, 2009: 66 ff.). Thus, some politically involved families fled to Brazil and stayed in São Paulo where they belonged to the middle or upper class. In addition, many of Bolivia's economically well

---

<sup>21</sup> According to Pampuch and Echalar Ascarrunz (2009: 150), Bolivia has the second largest certificated reserves in Latin America.

<sup>22</sup> In 2008, the remittances to Bolivia added up to more than 1,000 million dollars (cf. *Instituto Boliviano de Comercio Exterior* 2008: 6).

<sup>23</sup>In order to ensure that the original quotes of the Bolivians are preserved, I will not translate or change the quotes of my interview partners in the text. But, in some cases, I will offer a translation for better understanding: “More or less during the sixties/seventies, here, it was a paradise for the Bolivians, for the few Bolivians who were here. Not many Bolivians were here at first. At this time, there was more space for them to work in the area of ready-made-clothes. From the nineties on, it started to fill up more” (my translation, sn).

<sup>24</sup> The abbreviations offer information about the informant with regard to whether she is included in the linguistic analysis (B=not included), age (40), sex: m.=male, year of arrival to São Paulo: 1986, profession: J=Journalist and place of residence: Ce = Centro. For a detailed explanation and an overview of the speakers, see Annex I, Table I.1 and Table I.2.

situated families<sup>25</sup> used and still use to send their sons and daughters to study abroad, either to Argentina, to the United States or to Brazil. Those coming to Brazil mostly settled in the cities of the Southeast, São Paulo and Rio de Janeiro, to visit Brazilian universities. Many of these academics, often physicians or engineers, stayed in these cities afterwards, got married to Brazilians and formed part of São Paulo's and Rio de Janeiro's middle and upper class finding better economic opportunities in Brazil than in Bolivia.

The profile of the immigrant coming to São Paulo to work in the textile industry towards the end of the twentieth century differs considerably from the academics described above: The majority is male and in the best working age between 15 and 35 years (Silva 2005: 16).<sup>26</sup> According to e.g. Silva (2005: 7 f.), most of them come from Bolivia's urban centers like El Alto/La Paz, Potosí, Santa Cruz and Cochabamba.<sup>27</sup> Ethnically, most of them are indigenous people, mainly Quechua or Aymara, meaning that they display the typical physical features associated with the inhabitants of the Andes. As there is “uma linha divisória entre a minoria branca mais rica, com ascendência hispânica, e a maioria indígena e, portanto, mais pobre” (Silva 2005: 12), this signifies also a socioeconomic division between the first and second group of immigrants. The majority of this later group of Bolivian immigrants finished second grade at school, corresponding more or less to ten years of schooling, but has no academic grade. Until the 1990s, most of them aimed for the US, Europe or Argentina, where a better economic situation and strong currencies attracted the Bolivians to come and work e.g. in hotels or construction sites. Since the end of the twentieth century Brazil and São Paulo, in particular, have become one of the most important destinations of Bolivian immigration, because the Koreans started to employ Bolivians in the textile industry and Argentina was shaken by a heavy economic crisis at the end of the 1990s culminating in the devaluation of the peso in 2002 (cf. e.g. Cortés Conde 2003), losing hence its attractiveness for the Bolivians.<sup>28</sup>

---

<sup>25</sup> Ethnically, they are often descendants of European families constituting the urban oligarchies in Santa Cruz, Sucre, La Paz and Cochabamba.

<sup>26</sup> According to the Bolivian *Instituto Nacional de Estadística*, in 2010, the medium age of Bolivia's population is 21.90 years (cf. *Instituto Nacional de Estadística*: 1).

<sup>27</sup> In 2010, 66% (= 6,922,107) of the Bolivian population (= 10,426,154) live in cities (*Instituto Nacional de Estadística*: 1).

<sup>28</sup> The economic and commercial liberalization of Brazil and Bolivia, progressing since the 1980s was a necessary precondition, facilitating the (illegal) migration between the countries (cf. Cacciamali and Gomes Azevedo 2006: 5).

As many of the immigrants come to São Paulo illegally, the Bolivians do not stand under the protection of the Brazilian laws or trade unions and work long hours without complaint. Thus the Koreans started to actively seek Bolivian workers for their sweatshops:

*The textile industry is highly competitive, nationally as much as in the scenario of international commerce. It is characterized by few barriers for the market entry and the infinitely elastic offer of manpower, considering the various flows of emigration from the peripheral countries. The reduction of costs through the decrease of salaries is frequently employed (Cacciamali and Azevedo 2006: 133, own translation, sn).*

After the first Bolivians were established there, more and more Bolivians came to São Paulo, attracted by the possibility to work, the rising transnational networks and the active propaganda that is distributed via radio or newspaper (Silva 2005: 19).

*The biggest percentage of those coming from Hispano-America is represented by the Bolivians [...]. Those being born in Bolivia increased by 65% in the nineties; in 1991, their share of the Hispano-American immigrants was 15,38% (Taschner and Bógus 1999: 49, own translation, sn).*

The negative articles and reports about slave-like working conditions in São Paulo's textile industry that appeared in the media<sup>29</sup> had no directly negative effect on the immigration rate of Bolivians. On the contrary, the vice-consul of Bolivia in São Paulo estimates that in 2010 more than 200,000 Bolivians live in the city (personal communication 3.2.2010). Due to the high rate of illegality of the immigrants, there are still no reliable statistical data about their number available: Silva (2008: 20) cites estimations of the NGO *Pastoral dos Migrantes*, the last census and the federal police, revealing a difference of almost 50,000 persons between the police's statistics and the NGO's estimations.

### 2.2.2 The socioeconomic context<sup>30</sup>

As shown before, the Bolivians coming to São Paulo can be divided into two groups: There are the academics, particularly doctors and dentists, who come to Brazil to finish their studies and stay to improve their already good living conditions: "Eu estava muito bem lá só achava que ia ser melhor aqui" (B8 (48) f. 1989 P. Os: 1:12). This immigration of people belonging to the urban middle/high class of Bolivia has a long tradition and is continued until today. As they achieve a high level of education at Bolivia's and Brazil's universities, they also form part of the middle class in São Paulo.

---

<sup>29</sup>Cf. e.g. *Senzalas bolivianas* (Barros 2000: Bolivian slave settlements, my translation, sn), *Imigrantes bolivianos vivem como escravos em São Paulo* (Varella 2004: Bolivian immigrants live like slaves in São Paulo, own translation, sn) or *A dura vida dos imigrantes bolivianos em São Paulo* (dos Santos et al. 2004: The hard life of the Bolivian immigrants in São Paulo, own translation, sn)

<sup>30</sup> If not indicated otherwise, the facts presented here rely on my own observations during the field studies in São Paulo (cf. Chapter 5), because of the limited amount of literature available.

Correspondingly, they do not live in poor districts of the *Zona Leste* or the periphery but rather in the southwestern and central quarters or in peripheral yet middle class regions like *Guarulhos* or *Osasco*. In regard of spatial mobility, the immigrants with an academic background use public transport as well as their own cars in order to get to work or to make journeys inside Brazil. But these immigrants represent only a minority compared to the masses of Bolivians that come to Brazil to work in the textile industry.

The majority of the latter group live in *Bom Retiro*, the traditional textile center of São Paulo and its neighboring districts *Brás* and *Parí*. Particularly *Brás* is a typical region for the new immigrants, because many of them arrive at the train stations of *Brás* or *Luz*. Here one can find many cheap guest houses and split apartments that offer accommodation to the newcomers for the first nights in São Paulo. Thus, these quarters signify a certain center for the community. In some places, the Bolivians even outnumber the Brazilians, particularly at weekends when the Bolivian markets at *Rua Coimbra (Brás)* and *Praça Kantuta (Parí)* form a gathering point for the Bolivian community. To escape raids of the federal police which happen regularly in these areas,

*the immigrants without documents adopt the strategy to leave the quarters of high concentration of Hispano-Americans [...]. This strategy, together with other economic factors such as the renting prices, has led to a spatial dis-concentration, in particular of those who work in the textile industry (da Silva 2008: 24, own translation, sn).*

Thus, many Bolivians live in peripheral districts like *Patriarca* in the extreme east or *Vila Andrade* in the extreme south. As shown before, these districts are dominated by lower classes that themselves had migrated there in the decades before. There, the Bolivian population is not as concentrated as around *Bom Retiro* and *Parí*, but I myself observed football championships of various teams consisting only of Bolivian participants held in *Patriarca*.

The few statistics available about Bolivians in São Paulo also confirm their concentration in the central region and inner circle: Between 1991 and 2000, a decade when the Bolivian mass migration only started, (legal) Bolivian immigration increased by 65% (Bógus and Taschner 2004: 49), reaching 22.37% of all immigrants into São Paulo in 2001.

The majority of the Bolivians employed in the textile industry lives and works at the same place, i.e. in the crowded sweatshops. Since they are paid per unit, they have to work very long hours. Because of their illegal status, their working hours of up to 16 hours a day and their little financial resources, they show only “escassa presença no espaço publico” (Cymbalista and Xavier 2007: 130) and are

almost invisible in the city (Soraia Vilela 2010). The governmental amnesty of 2009 that allowed the legalization of ca. 30,000 Bolivian immigrants (cf. Giannini 2009) improved at least the legal situation for many of them. Other aspects such as the lack of leisure time or money continue to limit their spatial mobility:

*[...] não dá tempo pra ir a viajar não dá pra conhecer muito assim. Com o tempo que nós temos aqui só fim de semana só no domingo (B6 (30) m. 2003 S. ZL: 7.45)*

Only Bolivians who have their own sweatshop can afford a car. The others take bus, train and the subway, if they can afford it:

*Para mim é um privilégio ir à praça porque passagem é todo para mim um privilégio (A4 (21) m. 2006 S. ZL: 12.00).*

Thus, some Bolivians have only been to the districts where they work and to the Bolivian markets but never to São Paulo's most prominent street, the *Avenida Paulista*, after having spent several years in the city. One of the speakers states that after having been living in the city for 20 years: "Yo São Paulo no conozco tanto casi" (A25 (60) f. 1989 S/SA ZL: 14.40).

Accordingly, many *Paulistanos* know the immigrants only because of journal articles that formed the term of the *trabalhador escravo Boliviano* (cf. e.g. Moraes 2004 and O Globo 2010) which has a huge impact on public opinion until today.

As a matter of fact, the working and living conditions of the Bolivian seamstresses are very hard, particularly when they arrive to São Paulo. Additionally, it is mostly the *dona*<sup>31</sup> who pays for the (illegal) journey from Bolivia to Brazil. Until they have paid back this "loan", they do not get any salary:

*Como você ganhe segundo o avanço e quase não avança porque não sabe assim [...] e aí que não ganha não. Además do que faz dois mês que tem que trabalhar de graça para pagar o passagem (A4 (21) m. 2006 S. ZL: 18.06).*

Originally, Koreans were the *donas* of the sweatshops exploiting the Bolivian workers:

*Los Coreanos mucho nos exploraron [...] nós trabalhávamos, un centavito pagavan [...] no dava para trabalhar assim (A25 (60) f. 1989 S/SA ZL: 7.35).*

Today, though, the *donas* are often Bolivians themselves, having substituted the Koreans that now only work as contractors assigning work to the Bolivian sweatshops (cf. Cacciamali and Gomes Azevedo 2006: 6). However, this substitution did not necessarily lead to better working and living conditions for the new immigrants.

---

<sup>31</sup> I use the Brazilian word for the bosses of sweatshops.

The food and the sleeping place offered by the *dona* are also deducted from the already meager salary:

*Uma amiga minha [...] ela costura para fábricas. Ela contratava eu me lembro que ela comentou que tinha um galpão grande tipo garagem coberta e dava aí para que eles durmam e dava comida e pagava 50 por mês (B8 (48) f. 1989 P. Os: 11.03).*

Even if this might be an extremely negative exception, it is further confirmed that at least as long as the Bolivians are illegal, the salary is very low:

*Agora se você não tem documento você vai ganhar só uns 250 reais - esse é o problema (A9 (42) m. 1994 D. ZL: 21.14).*

Even worse, often several workers – men, women and even children – only share one room, apart from the actual sewing factory, where they sleep, eat and live in cramped conditions – a situation that provokes social problems inside the group and also the increase of sexual diseases. Bad nutrition, the dust of cloth and small sticky rooms additionally promote diseases like tuberculosis (cf. Patussi, cited in Cacciamali and Gomes Azevedo 2006: Appendix III, 6). This situation is aggravated by their illegal status, because many immigrants are too afraid to visit a physician.

Fortunately, there are several organizations that help the immigrants in these situations. The *Centro Pastoral do Migrante*, for example, is an organization that offers help like gratis consultation with a lawyer of Bolivian origin (Cacciamali and Gomes Azevedo 2006: Appendix I, 3) and support with the bureaucracy in connection with their legalization (cf. Camacho Kadluba 2007: 30). At the *Asociación de los Residentes Bolivianos (ADRB)*<sup>32</sup>, for example, the Bolivians get free medical treatment and advice by Bolivian physicians.

Additionally, there are numerous folkloric associations, some of them conjoined in the *Associação Cultural de Grupos e Conjuntos Folclóricos Bolívia Brasil*, who foster the folkloric dances of Bolivia, presenting them at feasts held on Bolivian public holidays, e.g. at the *Memorial da América Latina* (cf. e.g. Badilho 2007). The most popular leisure activity of the Bolivians, though, is to play soccer:

*Grande parte dos bolivianos opta [...] quase que exclusivamente, como atividade de lazer no Brasil, pela prática do futebol (Alves 2011: 80).*

Because of their working hours, the seamstresses' leisure time is restricted to Saturday afternoon and Sunday. Then they gather at the *Rua Coimbra* or the *Praça Kantuta* to play soccer, eat typical Bolivi-

---

<sup>32</sup> The association was founded by Bolivian doctors and dentists who came to Brasil in the late 60s, with the purpose of “ajudar os mais jovens [...] para ser bem encaminhados” (A21 (62) m. 1964 E. Ce: 5.20), i.e. “a promoção cultural, recreativa e social dos seus compatriotas” (Silva 2005: 37).

an food and buy CDs, clothing from their homeland and cheap telephone cards to call their family in Bolivia.

Other points of attraction for the immigrants, that are often mentioned, are the zoo of São Paulo, situated in the Southern Zone, and the parks, for example the *Parque Ibirapuera*, situated in the south, too. As these places are mainly frequented by Brazilians, these could be places where Brazilians and Bolivians meet and communicate in Portuguese.

On the micro level, this means that Bolivians working in the textile industry can or even must live in São Paulo almost without leaving the Bolivian sphere, because the extreme working conditions and their illegality heavily limit their opportunities to participate in urban life. On the macro level it can be observed that the Bolivian community has already built its own space in the city where contact between Brazilians and Bolivians is rare.

The immigrants with academic background, though, are rarely involved with Bolivian associations in their leisure time. They attend the *Praça Kantuta* to eat Bolivian food, but according to the sales assistants there, one can observe that the wealthier Bolivians arriving by car and spending more money, often accompanied by Brazilian family members or friends come around noon, while the seamstresses only eat later.

This segregation is further promoted by the experiences of discrimination and violence of many Bolivians. They are often stigmatized as criminals, because the Brazilians think that most of the Hispano-Americans are involved in drug trafficking as they use Coca (cf. e.g. Silva 2008: 40):

*Cuando uno entra nun shopping, como soy así, medio morena, cholita, ellos piensan que yo vengo a roubar (cited in Silva 2008: 38).*

Formerly, this situation had been better, but because of neighborly conflicts with drunk and noisy Bolivians, one of the speakers states that even in *Bom Retiro* and *Parí* the prejudices have increased a lot:

*Cuando teníamos llegado recién en 1989 gente brasileira eran muchas finas. [...] agora tudo discriminación tudo discriminación. Ya no nos olhan os brasileiros fican olhando con raiva [...] como unas quantas personas han venido bebado o mal visto [...].”(A25 (60) f. 1989 S/SA ZL: 9.37).*

Experiences like being threatened or even being beaten up are common for the Bolivians that meet the phenotypic image which the Brazilians have of them:

*Yo estuve en la rua, no, los brasileiros gritan pra nos: ‘Boliviano, vai vai vai pra su país [...] vai embora vaya na Bolívia.’”(B9 (24) m. 2005 S. ZL: 8.04).*

The ones working in the textile industry are also often victims of violence and crime:

*Pouco tempo após a sua chegada a São Paulo, foi assaltado e brutalmente espancado. [...] traumatizado com infortúnio, sofreu seqüelas por causa dos golpes desferidos (Barros 2000: 4).*

As most of them live in districts of São Paulo with high criminality rates like *Brás* or the periphery (cf. Figure 2.11 to 3.13), many of them narrate about crimes like robbery and theft that happened to them and are perceived as further discriminations:

*Una vez me venian aqui en la rua levaram toda mi mercadoria brasileiro tudo. Levaram tudo porque só boliviano a brasileiro deixou aí [...] só de mim levaram (A3 (34) f. 2004 SA ZL: 3.58).*

One of the *ADRB* even mentions the death of Bolivians in the dangerous districts of the *Zona Leste*: “Muitas vezes são assassinados aqui na Zona Leste e ninguém fica sabendo” (A21 (62) m. 1964 E. Ce: 12.14). Therefore, many Bolivians are too afraid to go out and leave the known places where they feel protected by their community:

*Es diferente allá como en Bolivia porque allá en Bolivia ... se puede salir allá en Bolivia salir a pasear mientras aquí no, no se puedes (B11 (29) f. 2008 S. Pe: 0.41).*

As the other group of Bolivian immigrants, the academics, stem from the Bolivian upper class and mostly have European ancestry, they do not look like the indigenous Bolivians and therefore suffer less discrimination. Thus, the primary reason for segregation is not nationality but social status, i.e. the life of the Bolivian academics belonging to São Paulo’s middle class differs in every aspect from the lives of the seamstresses. Correspondingly, the socioeconomic segregation of São Paulo’s society is continued inside the Bolivian community:

*O problema é nosso porque nós somos muito como se diz separados nós não estamos juntos [...] então é eu vejo em outras comunidades assim os coreanos os árabes [...] eles trabalham juntos, chega um e ajudam ele. Aqui não. Somos muito desunidos (A16 (32) f. 2004 P. ZL: 7.00).*

## 2.3 Summing up

---

In this chapter I characterized São Paulo as a segregated city. The socioeconomic segregation between high/middle class and lower class manifests itself in the geographic distribution in the city, in the means of transport, in crime and in cultural participation in urban life, resulting in a high degree of homogenization inside each social group but a strong disparity between the different classes.

Thus, the category of socioeconomic classes is still highly valid in São Paulo, because they influence every part of the individual’s life. Correspondingly, time and space where people from different classes could meet, is restricted to the working place or to casual encounters.

The social stratification of the community of the Bolivian immigrants seems to be similar as the one established for São Paulo: On the one hand, the immigrants with academic background originating from Bolivia's urban upper and middle class participate in urban life just as the upper and middle class *Paulistanos* do. On the other hand, the Bolivians working in the textile industry of São Paulo live in the poor regions of the city, they are victims of crime and discrimination, and have little leisure time, little geographic mobility and thus few chances to take part in the social life of São Paulo. Therefore, they have created their own Bolivian urban space where they live separated from São Paulo's Brazilian and also Bolivian upper class population.

### 3. State of the Art

---

This chapter gives a critical overview of the literature relevant for the comprehension of this study. The various areas of investigation influencing this thesis, urban sociolinguistics and language contact, are presented by moving from general theories to specific applications, from general urban sociolinguistic approaches to those concerning cities in Brazil and Bolivia, from unspecific language-contact theories to Spanish-Portuguese language contact situations. These aspects have already been explored broadly. However, due to the limited space available here, I limit my elaborations to studies that illuminate best the state of the art important for the theoretical and methodological design of this study: I discuss the sociolinguistic tradition of urban studies in modern linguistics and show examples for its application to Bolivian Castilian and Brazilian Portuguese. With regard to language contact and second language acquisition I include psycholinguistic and sociolinguistic approaches, as well as studies that explicitly concern narrowly related languages and the role of migration.

---

Bolivian immigrants in São Paulo themselves have never been the center of a linguistic study so far, however, my inquiry concerns various aspects that have already been investigated and hence constitute the basis for the theoretical and methodological approach.

As the title of this study implies, there are two main aspects of research to be considered: In order to elaborate the influence of certain social characteristics of São Paulo on the Bolivian's Portuguese by appropriate methodology I resort to theoretical and methodological findings of urban sociolinguistics. In order to establish the linguistic process underlying language contact situations, studies focusing on this aspect of language contact and second language acquisition must be considered, in particular between two languages as closely related as Spanish and Portuguese. As literature about both fields is rather ample, I concentrate on studies of modern linguistics that later serve as the foundation for my own approach.

#### 3.1 (Urban) Sociolinguistics

---

Today's sociolinguistics combines various research interests, theoretical approaches and methodological principles. A common ground for the different branches of sociolinguistics is the view of language as intrinsically social and thus the "common orientation to language data, believing that analyses of linguistic behavior must be based on empirical data" (Milroy and Gordon 2006: 2).

These assumptions might seem rather obvious, but as Coupland (2007: 4) points out, they came into being in the 1960s as a reaction to the common belief held by the public and also many scholars that "what matters in language is linguistic uniformity and 'standardness'", constituting thus a real innovation and challenge to "the belief that 'standard' language is more orderly and more worthwhile

than ‘non-standard’ language”. Until now, the variability and variation observed in language use is often dismissed by theoretical linguistics<sup>33</sup>, while sociolinguistics sees it as “the essential property of language” (Milroy and Gordon 2006: 4). Therefore, Coupland’s statement that “sociolinguists need to make a case for observing language as it is used in everyday life and for not relying on intuited or fabricated instances of language” (Coupland 2007: 4) is still valuable today.

Investigating language variation in urban societies constitutes only one part of the sociolinguistic discipline. Particularly in the USA, however, Labov’s seminal sociolinguistic studies that actually founded the discipline were conducted in cities, e.g. about New York’s English (cf. e.g. Labov 1966 and 1972).

Early urban studies conditioned by dialectological traditions of their time, still selected certain speech groups as research subjects that seemed to constitute the stereotypical urban speaker<sup>34</sup> speaking “rough Cockney” or “the ‘genuine’ speech of Larne” (Sivertsen 1960 and Gregg 1964, cited in Milroy and Gordon 2006: 17). But soon it was recognized that

*[c]ities challenge the view that one discrete social style (e.g. a dialect) is associated with one place [...]. It has become the norm to consider cities as sociolinguistic systems that organise linguistic variation in complex ways (Coupland 2007: 2).*

To describe and analyze this complexity and the dynamics of language variation in urban societies, there are various theoretical and methodological approaches considered valid today. In an extensive overview of “three waves of analytic practice” (Eckert in review: 2), the author distinguishes first the survey era associated predominantly with the research of Labov (cf. e.g. 1966, 1972, 1974, 1991 and 2004)<sup>35</sup>. The survey era focuses on the structural association of linguistic variation with social variables such as class, ethnic group and gender. The second wave is characterized by the ethnographic approach that includes long-term studies such as those of Gal (1979) and Milroy (1980) which sought to discover “locally salient social categories [...] in virtue of their place in local social practice” (Eckert in review: 7). The third wave, the study of practice and stylistic perspective, sees variation not as a reflection of social classes but as the combination of indexical units in styles that

---

<sup>33</sup> The dichotomic distinction between language structure (competence) and language use (performance) as advocated by generative linguistics, for example, leads to the perception of variation as peripheral: “In this tradition variability is treated as a methodological complication: It introduces a kind of noise which obscures the important underlying invariance” (Milroy and Gordon 2006: 4).

<sup>34</sup> According to Milroy and Gordon (2006), these selections were conducted to get the urban counterpart of the then preferred speaker of dialectology, the famous NORM (non-mobile, old, rural men) (cf. l.c.: 17).

<sup>35</sup> The enumeration of the authors does not imply any exclusiveness. On the contrary, the same authors conducted studies that are associated with other waves (cf. e.g. Labov 1963).

individuals use for constructing social meaning and their identity in social practice (cf. l.c.: 14 ff.) as shown by e.g. Coupland (2001 and 2007) and Eckert (1989, 2000 and 2008). Although the three approaches differ greatly in terms of research interest, analytic practice and methodology, Eckert (in review) underlines that the waves do not supersede and replace the former (cf. l.c.: 27). On the contrary, “each wave adds to the preceding” (l.c.).

Following Eckert’s classification, structural variation concerns the first two waves that “related the meaning of variables quite directly to the social categories with which they correlate” (Eckert (in review): 2), while stylistic variation sees “variables as having more basic meanings that combine stylistically to construct the kinds of personae that populate social categories“ (l.c.).

For this study, both structural and stylistic variation, have to be considered, as well as studies about linguistic perception and evaluation. I will thus shortly discuss the main characteristics of these analytic practices, focusing particularly on their methodical implications, before I sketch the state of urban sociolinguistics in Brazil and Bolivia.

### 3.1.1 Structural variation

---

Structural variation constitutes an approach that “primarily addresses the properties of variable linguistic systems” (Milroy and Gordon 2006: 8) and not the individual speaker. Accordingly, Labov (2006: 5) states that “the language of individuals cannot be understood without knowledge of the community of which they are members”; thus his research focuses on homogenous speech communities of native speakers, even if this means excluding for example immigrants who live in the social environment researched.<sup>36</sup>

Assuming that “language varieties carry the social status of their speakers” (Eckert in review: 3), Labov establishes a “continuum of linguistic prestige” (l.c.) that reflects social stratification insofar as the standard variety of highest prestige is assigned to the upper end of the class hierarchy, while stigmatized low prestige varieties are relegated to the lower end. As people try to move into higher socioeconomic classes, they will also aim for the standard linguistically.<sup>37</sup>

---

<sup>36</sup> As Labov (1966: 3 and 4) himself states that the speech community of a city like NYC is very complex, with no fixed boundaries, mobility and “the line which divides the native speaker from the foreigner is broken by many doubtful cases”, this means also that he deliberately excludes the important factor of continuing migration from his surveys.

<sup>37</sup> Labov’s own studies (1972 and 2001) showed that these correlations are more complex: Not only those at the lowest level of the socioeconomic hierarchy use the lowest prestige varieties but also members of the upper working and lower middle class. At the same time, some lower middle class speakers display a wider range of variation, particularly towards more standard-close variants. The estimation that the lower classes speak more

The variety that Labov (1972) considers “the most systematic form of language” (l.c.: xvii) situated at the center of his interests as the source of language change, is not the standard but the vernacular; i.e. “the language first acquired by the language learner, controlled perfectly, and used primarily among intimate friends and family members” (Labov 2006: 86), in short the individual’s least reflective speech. Speakers who pay attention to their speech will tend to approximate the standard. To avoid this “worst plague of the sociolinguist” (Mollica and Braga 1992: 102, own translation, sn) and be able to record the vernacular, the researcher must overcome the observer’s paradox<sup>38</sup> by, for example, hiding her true research interest or choosing certain emotional topics like childhood games that create an atmosphere in which the speaker will use the vernacular (Labov 1972, cited l.c.).

In order to make the empirical data comparable and quantitatively analyzable, the typical method of data collection is the sociolinguistic interview. The (mostly phonological) variation is then quantified and analyzed statistically depending on the social variables established previously. Naturally, this also means that the social variables taken into account are theoretically unlimited, but in practice, only the correlation of linguistic features with rather broad categories can be captured in this way. Significant individual characteristics and locally specific social dynamics underlying these categories, however, cannot be determined on the basis of this kind of survey data.

*To understand the correlations between language and these global social categories, we need procedures which allow us to examine the specifics of local practice and local conditions, and which are sensitive to the local social categories and locally contracted ties with which speakers operate in their everyday lives (Milroy and Gordon 2006: 116).*

Thus, researchers conduct ethnographic studies, i.e. they

*focus on smaller communities for relatively long periods of time, and aim to discover, rather than to assume, locally salient social categories (Eckert in review: 7).*

According to Eckert (l.c.: 5), these ethnographic studies constitute the second wave of (urban) sociolinguistics, bridging the differences between the first wave and the third wave just described, the study of “practice and stylistic perspective” (l.c.: 14). On one hand, the general structural correlation between linguistic features and social categories is maintained, even if finer local categories now

---

naturally, while the higher social classes display a higher degree of conventionalization, has been criticized as too static (cf. e.g. Auer 1989: 39).

<sup>38</sup> The observer’s paradox describes the paradoxical situation in which, although the communication is observed and even recorded, the speaker should talk as if she were not being observed to obtain the vernacular (cf. e.g. Labov 2006: 86).

substitute the primary categories.<sup>39</sup> On the other hand, the long-term involvement of researchers in the community enabled them to discover the categories “in virtue of their place in local social practice” (Eckert in review: 7). In this way, the primary focus is often modified from abstract language patterns to speakers as social actors (cf. Milroy and Gordon 2006: 3), while the researcher mostly takes on an observing role inside the community; i.e. the recordings are only one part of the data collected, the other part consists of e.g. observations, pictures or notes that together constitute a thorough and multifaceted picture of the researched people. Nonetheless, the ethnographic approach can also be combined with quantitative methodology as Milroy (1980) demonstrates, although a qualitative study certainly corresponds better to the requirements of ethnographic involvement.

One category successfully introduced to substitute the socioeconomic classes at a local level was the social network. According to Dubois and Horvath (1998),

*an individual's social network is the aggregate of relationships contracted with others, a boundless web of ties which reaches out through social and geographical space (cited in Milroy and Gordon 2006: 117).*

Adopting the sociological concept of ego-centered networks (cf. e.g. Friedrichs 1995), Milroy (1980) showed in her study of working class people in Belfast the linguistic variation subject to the different networks of individual speakers. Depending on the quality of the individual networks and the quality of the ties – distinguishing density and complexity as well as first-order network ties constituted by weak and strong ties from second-order ties<sup>40</sup> – the use of linguistic forms varies considerably. Roughly said, men with close-knit networks in particular maintained socially rather stigmatized linguistic forms while young women with – due to their work across the city – rather loose networks tended to change their language in the direction of standard features (cf. Milroy 1980: 139 ff.). Thus, Milroy (1980) deduces that the

*two effects of the processes of urbanization and industrialization will be to disperse traditional close-knit networks, and to accelerate linguistic standardization (l.c.: 185).*

In later studies, however, she admits that the type of network often associated with rural societies, also has to be considered for urban environments:

---

<sup>39</sup> In Milroy and Milroy (1992), the authors give an extended example for combining the category of social class with social networks, referring for example to the maintenance of the vernacular in Belfast's close-knit social networks (l.c.: 4).

<sup>40</sup> The density of the network depends on the extent to which the single ties (= contacts) are linked to each other; if they are many-stranded, i.e. the contact concerns various relations, e.g. work and family, the network is multiplex (vs. simplex). First-order network ties refer to a person's direct contacts; second-order ties are those to whom the link is indirect; strong ties are “roughly ties that connect friends or kin as opposed to those that connect acquaintances [= weak ties]” (cf. Milroy 2006: 550, Milroy and Gordon 2006: 117).

*Neighborhoods involving close kinship and personal ties seem still to be created rather than discouraged by city life, since those who form urban ethnic communities gravitate to form ties with, and often to live with, others from similar linguistic or ethnic backgrounds (Giddens 1989, cited in Milroy 2006: 559).*

A concept related to networks that is also employed to describe smaller social categories is Community of Practice (CoP), “[...] an aggregate of people who come together around mutual engagement in an endeavor. [...] practices emerge in the course of this mutual endeavor” (Paiva and Duarte 2003: 464). Meyerhoff (2006) points to the important difference, that the membership in a dense network can be due to chance or circumstance<sup>41</sup>, while it is conscious in regard of CoP. In addition, network theory deals primarily with the structure and the properties of the ties linking the ego with other contacts (Milroy 2006: 552) and thus still belongs to structural variation. CoP, on the other hand, focuses on practice, the interaction going on in the community, i.e. this part of the ego’s social network. Thus, it is closely connected with stylistic variation and will be discussed in the next chapter.

### 3.1.2 Stylistic variation

---

As was shown above, in quantitative variation surveys, stylistic variation is reduced to “the mechanistic view that style is simply the result of degrees of attention paid to speech” (Bell 1991: 76). This hypothesis was heavily criticized. Coupland (2007: 38 f.), for example, claims that “stylistic 'strata' are [...] artifacts of the empirical methods used”, because the application of quantitative methodology that

*predetermines categories [...] and which then examines the aggregated scores [...] asks only very limited questions about stylistic variation (Coupland 2001: 191).*

Rickford (2001) also underlines that he prefers Bell’s model of audience design (cf. Bell 1984, 1991, 2001 and 2009) as responsible for stylistic variation:

*While attention to speech does not strike me as ultimately likely to be the ‘right’ unidimensional model, the unidimensional audience design approach of Allan Bell [...] still strikes me as very promising (Rickford 2001: 222).*

Resorting to speech accommodation theory, which will be discussed in 2.2.1.3, Bell’s audience design was a reaction to Labov’s characterization of stylistic variation. Bell (1991: 76) hypothesized “that the audience is the primary factor in influencing speakers to shift their style”. As he further suggests a model that integrates qualitative and quantitative research that “will be partly in standard

---

<sup>41</sup> That the membership of a network can be entirely assigned to chance or circumstance seems rather improbable, at least in democratic cultures where the individual herself chooses friends, work, peer groups etc.

variationist terms” (Bell 2001: 168), his approach seems to represent a compromise between Labov’s and Coupland’s approach.

For the latter, however, the study of stylistic variation is something completely different from first wave sociolinguistics, because it focuses on practice, i.e. “it attends to social 'doing' in place of structural 'being'” (Coupland 2007: 49). Thus it “reworks the assumptions underlying the structural variationist model” (l.c.). This view is fiercely endorsed by Eckert (cf. e.g. 2001), who emphasizes that the social meaning of variables must be considered beyond the socioeconomic categories and the social stratification of prestige and stigma. Particularly today, as

*social life is more obviously amenable to being socially constructed in late-modernity [...], [s]tudying social meaning through sociolinguistic styling gives us a way of understanding social identities and social relationships with sufficient flexibility and dynamism to capture some of the qualities of late-modern social life” (Coupland 2007: 30).*

Based on Bourdieu’s habitus concept<sup>42</sup>, the study of stylistic variation in practice focuses on

*the day-to-day activity in which human-beings make sense of their lives and move their projects along in the face of constraints imposed by social structure and the power relations that keep that structure in place (Eckert in review: 14).*

In a linguistic sense, it is assumed that “variables carry complex indexical meanings well beyond place in the macro-sociological matrix” (l.c.: 27). Indexical meaning, as introduced by Silverstein (e.g. 1976 and 2003), is used here claiming that “‘indexical order’ is the concept necessary to [show] us how to relate the micro-social to the macro-social frames of analysis of any sociolinguistic phenomenon” (Silverstein 2003: 193), distinguishing n-th and n+1st indexical order: “[F]irst-order-indexicality entails the association by social actors of a linguistic form or variety [...] with some meaningful social group such as female, Asian, Spanish, working class, aristocratic and so forth. [...] second-order indexicality is a metapragmatic concept, describing the noticing, discussion, and rationalization of first-order-indexicality. It is these second-order indexical processes that emerge as ideologies” (Milroy 2004: 167, original emphasis). It must be observed, however, that semiotically, the concept of *symbol* based on convention and custom seems to be more adequate for this relation than the relation of indexes based on contiguity. Silverstein himself explicitly mentions the importance of convention for the indexical interpretation: “Indexical significance [...] can be ‘read’ in relation to conventional norms” (Silverstein 2003: 196 f.).

---

<sup>42</sup> “The habitus is the lasting set of dispositions – patterns of behavior, affect and thought – that one develops in the course of life in a particular social position. The habitus mediates between social structure and individual behavior, shaping, but not determining, individual action” (Eckert in review: 14).

However, maintaining the original denomination of Silverstein and Eckert, communication represents a dynamic model where social indexes of linguistic forms are continually interpreted by speaker and hearer according to language ideology:

*The difference between the notion of marker as used in variation studies and the index of Silverstein's treatment is in the ideological embedding of the process by which the link between form and meaning is made and remade. Participation in discourse involves a continual interpretation of forms in context, an in-the-moment assigning of indexical values to linguistic forms. A form with an indexical value, what Silverstein calls an  $n$ th order usage, is always available for reinterpretation - for the acquisition of an  $n + 1$ st value (Eckert 2008: 463).*

Thus, the variables studied by quantitative variationists are the basis for the attribution of indexical value to the feature, because the latter is “constructed locally, but around ideological issues that link everyday interactions to the political economy” (Eckert in review: 19). Therefore, there is no simple and fixed correlation between linguistic features and the social dimension. Instead, it depends on the perspective of the hearer, the style in which it is embedded (cf. e.g. Eckert 2008: 465), the relation between speaker and hearer or the practice in which they are involved together, and which meaning the hearer associates (and the speaker implies) with the linguistic feature.<sup>43</sup> In a CoP, for example, the regular contact and the common endeavor form a certain style, i.e. the indexical meaning of linguistic features is unambiguous in this CoP, although it might be different for outsiders, and also outside the CoP.

Thus, research on style must focus on

*the ways in which the same speakers talk differently on different occasions rather than the ways in which different speakers talk differently from each other (Bell 2009: 265).*

Methodologically, researchers should only focus on a few speakers in their CoP or even on single case studies, as Coupland (2007) suggests:

*Single-case analyses are more likely to allow an adequate sensitivity to context and contextualisation, where we can come to understand what the styling of variation can achieve. There is the possibility of generalising to what is stylistically possible, rather than to 'what people typically do' (Coupland 2007: 28).*

As linguistic areas to be studied, they suggest not concentrating on phonology, but also taking morphosyntactic and lexical features into account (cf. Eckert in review: 27). Suggestions like Preston's (2001), to employ quantitative methodology, are strictly rejected as “methodological convenience

---

<sup>43</sup> Eckert (2008: 453) calls the quantity of potential meanings the features' *indexical field*. But as she herself does not seem to be sure of this term (cf. l.c.: 473), I will not use it here.

[that] has curtailed theoretical elaboration” (Coupland 2001: 192), because in their view the data have to be collected inside the realm of discourse between speaker and hearer.

### 3.1.3 Perception

---

In linguistics, the hearer’s perspective seems to play a subordinated role, although it has a decisive part at least in stylistic variation. Stimulated by social psychological research, the study of language attitudes, perception and evaluation of speech, achieved deep insights into the connection between style and perception. Perception is understood as not only the sensory input but rather the “filter through which sensory data are strained, and it is obvious that the establishment and maintenance of this filter are culturally specific and – within social groupings – individualized to greater or lesser extent” (Edwards 2009: 154). However, I will not adopt the distinction between *perception*<sup>44</sup>, *representation*<sup>45</sup> and *perceptive based representation*<sup>46</sup> as suggested by Krefeld and Pustka (2010), because in my understanding, the linguistic features and the attitudes or social stereotypes triggered by them belong inseparably together, as do the indexicality orders I just discussed for stylistic variation.<sup>47</sup>; Studies like Niedzilski (1999) show that “representation” overcomes “perception”, even if the hearers refer to actual communication situations, perceiving linguistic features that are objectively nonexistent. Preston (2010: 4) subsumes the elements *language attitudes*, *perception* and *evaluation of speech* under the umbrella term *language regard*, referring to “what nonlinguists believe about languages and language varieties [...] as well as how they evaluate them”. As I agree with him on the strong connection between language attitude and language ideology, I adopt this term for this study. With regard to urban sociolinguistics it gains importance, because they trigger stereotypical social group attitudes:

*Social perceptions of speech occur in nonrandom ways. [...] [Through them] we can view social structures and [...] confirm what we observe from other perspectives.” (Edwards 1999: 102)*

---

<sup>44</sup> Perception: Referring to perception in an actual communication situation (cf. Krefeld and Pustka 2010, own translation, sn)

<sup>45</sup> Repräsentation: Referring to knowledge and attitudes about languages independent of an actual communication situation (cf. Krefeld and Pustka 2010, own translation, sn)

<sup>46</sup> Perzeptiv basierte Repräsentation: Representations based on perception in an actual communication situation (cf. Krefeld and Pustka 2010, own translation, sn)

<sup>47</sup> The authors themselves state that representation and perception are closely related, regularly influenced and modified by each other: “current perceptions are filtered by already existing representations” (Krefeld and Pustka 2010: 14, own translation, sn)

This means, the speakers are evaluated according to the characteristics that reflect how the group is perceived to which they are conventionally associated with, and not intrinsic or aesthetic qualities. In addition, Cargile et al. (1994: 223) hypothesize that

*attitudes triggered by various linguistic features are most likely to affect recipients' behaviours towards senders in contexts of low familiarity,*

because there is “less uncertainty to be reduced” (l.c.) between two familiar persons. Thus, in an urban environment, where the direct contact and communication with unfamiliar persons is just more probable – be it a taxi driver, the lady in the subway or the pizza delivery boy – language regard must be considered an influential factor on speech outcome.

Considering the three waves mentioned above, the study of language regard was conducted in studies of every wave. Labov (1966), for example, investigated the “general attitudes towards the speech of New York City”, revealing “one extreme of a continuum of linguistic insecurity” (cited in Labov 2006: 324), particularly with regard to the speech of working-class people, confirming thus his estimation about the stigmatized vernacular.<sup>48</sup>

Nevertheless, most of the studies were conducted by scholars coming from a social psychological background. Their research revealed general psychological patterns like the tendency to evaluate speech examples according to two dimensions: social status and solidarity (Milroy and Preston 1999: 4-5). Unfortunately,

*speech examples submitted for judgment have differed broadly, and rarely has the inquiry extended to consider which linguistic elements (or which combinations or frequencies of elements) were chiefly responsible for the judgments elicited (Milroy and Preston 1999: 5).*

Since then, few studies like those of Niedzilski (1999) and Campbell-Kibler (2008 and 2009) showed that by experimental tests, the investigation of single linguistic features is also possible. The latter employed the matched-guise technique (cf. Lambert et al. 1960) to reveal subconscious attitudes about the variable *-ing* in American English: The author manipulated this feature artificially and revealed that speakers indeed were evaluated differently according to their pronunciation of *-ing*. She deduced that

---

<sup>48</sup> Later, the research of language regard even developed into the branch of perceptual dialectology, also called folk linguistics (cf. Preston 1999), that – comparable to a dialectological atlas – investigates the attitudes towards varieties employing for example maps where the respondents were asked to “draw boundaries [...] around speech areas” (Preston 2010: 1). This practice is mostly applied to rural areas, although e.g. Pustka (2008) successfully showed its relevance for Parisian varieties.

*rather than meaning one particular thing, (ING) is tied to a network of related concepts. [...] Which one it is used to mean (or ends up meaning) is different based on a number of contextual factors (Campbell-Kibler 2009: 149),*

confirming hence Eckert's hypotheses about the potential field of social meaning. Preston (2011: 2) also states that "general transfers of speech facts to groups [...] in which linguistic facts are related to nonlinguistic characteristics of a group [...]" are referred to as Silverstein's second-order indexicalities, "for they index not just a group identity [...] but provide a new linkage between some aspect of a group's stereotypical behavior" (l.c.: 2-3).

Thus, there is obviously a certain affinity between the study of stylistic variation and language regard. Concerning L2 speakers, Carvalho (2008) conducted a *perception test* about the language regard of *Cariocas*, the inhabitants of Rio de Janeiro, toward native Spanish speakers having acquired Portuguese on different proficiency levels that will serve as base for the perception test conducted for this study (cf. Chapter 7).

Methodologically, various implicit and explicit methods are applied for the study of language regard, since it contains affective as well as cognitive aspects that are conscious and subconsciously called forward,: Apart from the mostly quantitative matched-guise technique aiming for subconscious attitudes, qualitative methods like hand-drawn maps or the evaluation of caricatures are also applied to explore rather explicit stereotypes.

### 3.1.4 Urban sociolinguistics in Brazil

---

In Brazil, studies in urban sociolinguistics have been conducted regularly for the last forty years. Starting with the project *Norma Urbana Culta* (NURC), which was the first to collect oral data for sociolinguistic studies in six of Brazil's biggest cities<sup>49</sup> – Porto Alegre, Recife, Rio de Janeiro, Salvador and São Paulo – 600 hours of conversations between Brazilian informants with an academic grade have been collected and transcribed (cf. Vandresen 2000: 25). From these data followed various publications, for example Callou (1992) and Callou and Lopes (1993) with regard to Rio de Janeiro and Castilho and Preti (1987), Castilho (1989) and Preti (2003) with regard to São Paulo. Further studies comparing data from all the cities were published later in the series *Gramática do Português Falado* that led to the publication of ten books uniting numerous studies on phonology and syntax, but also discourse and textual coherence (cf. e.g. Abaurre 2002, Castilho 2002, Ilari 2002). The NURC data collection was conducted in the tradition of the orthodox Labovian method de-

---

<sup>49</sup> NURC was inspired by the Latin America-wide *Proyecto de Estudio de la Norma Urbana Culta* founded by the Mexican linguist Lope Blanch (cf. e.g. Lope Blanch 1986).

scribed above. Probably due to this anchoring in structural variation, Brazilian linguistics distinguishes studies concerning the Brazilian standard (called *norma culta*) and the Brazilian substandard (called *português popular*). These varieties are associated with different socioeconomic classes, claiming that people from a higher social class with higher level of education speak the *Norma Culta*, while people of a lower social class with lower education degree speak *Português Popular*:

[...] postulamos que essa variedade de língua popular é utilizada por falantes não escolarizados ou de baixo nível de escolaridade (até, no máximo, quatro anos de escolarização), moradores dos complexos centros urbanos brasileiros (Rodrigues 2009: 152).

Thus, a clear-cut polarization between the prestige variety on one hand and *Português Popular* on the other hand is established:

[...] postulamos [...] uma variedade lingüística popular ou dialeto social popular, cujas características podem ser construídas em contraste com as do dialeto culto, enquanto variedade de referência. Trata-se de uma variedade não codificada, porque não estabilizada nem unificada, e não normalizada porque não padronizada; uma variedade de menor prestígio [...] (Rodrigues 2009: 152).

The absence of number agreement in the nominal syntagm or the absence of agreement between subject and verb, for example, are highly salient features of *Português Popular* (cf. Rodrigues 2000) and are often referred to in literature and in public.

The opinion that speakers of *Português Popular* aim for the prestige variety *Português Culto* but are not able to master it, is far spread. This is shown by the heated discussion around the introduction of a new didactic book which postulates e.g. that the absence of number agreement as feature of *Português Popular* should not be considered as an error (cf. e.g. Ramos et al. 2011 and Bahé and Lucena 2011).

Rodrigues (2009) explains the unsuccessful attempts of the *Português Popular* speakers as result of missing familiarity with the *Norma Culta* because of the lack of institutional education:

[o] falante popular não chega a dominar as normas da língua padrão, dado seu distanciamento da escola e dos veículos de comunicação verbal escritos; com ela não se familiariza por não ter oportunidade de ouvi-la diuturnamente, apesar de os meios de comunicação de massa, em princípio, fazerem uso do dialeto de prestígio (Rodrigues, 2009: 152).

This view, clearly adopted from Labovian sociolinguistics, is too short-sighted though, because the features apparently separating the two varieties are fluent. Castilho (2010), for example, states that

[q]uando distinguimos PB [Brazilian Portuguese, sn] popular de PB culto, estamos nos referindo a variações socioculturais não separáveis rigidamente. Ninguém é

*exclusivamente ‘falante popular’ nem ‘falante culto’ [...] o que distingue uma de outra é a frequência de uso (Castilho 2010: 205).*

While the *norma culta* was studied thoroughly, associating phonological as well as morphosyntactic features to e.g. gender, linguistic environment and the speaker’s place of residence, a differentiated perspective on *Português Popular* is rather rare in Brazilian sociolinguistics. However, in Rio de Janeiro, various sociolinguistic projects focusing on *Português Popular* had already started in the 70s. They lead to the formation of the current project *Programa de Estudos sobre o Uso de Língua (PEUL)* that also stays in the Labovian variationist tradition (cf. Paiva and Pereira Scherre 1999). So far the project resulted in various studies such as Ferrari (1997), Lopes (2003), Paiva and Duarte (2003), and São Paulo’s such as Rodrigues (1987) and Crespim Pereira (2007).<sup>50</sup> In addition, there are many studies focusing on its single characteristics such as Pereira Scherre (1991 and 1998), Reich (2002), and Rodrigues (2000). A rare “third-wave” study about São Paulo’s *Português Popular* was conducted by Ferreira Coelho (2006), using the concept of social networks to explain the variation between *nós* and *a gente* and the following verbal form.

Beyond that, Bortoni-Ricardo’s 1985 study about rural migrants having come to Brasília constitutes one of the few ethnographic approaches described above. Employing the concept of networks, she shows how urbanized networks are responsible for the loss of rural dialect markers and the acquisition of urban markers. In her complex approaches to describing the reason for the maintenance of rural linguistic features as opposed to approximation to the urban standard, she develops two indices. These are the network integration index and the network urbanization index, which also takes e.g. schooling level, spatial mobility and media exposure into account (cf. Bortoni-Ricardo 1985: 169). Based on this work, her preoccupation with educational questions and Bell’s audience design approach (1984) previously discussed, she suggests a model of three continua to break off the old separation between *norma culta* and *popular*: rural-urban, oral-written and stylistic monitoring (cf. Bortoni-Ricardo 1998: 101 ff.).<sup>51</sup>

Apart from Carvalho (2008) as cited above, language regard was never investigated in the cities, but in the rural areas of Brazil’s south (cf. Preston 1989 and Auer, Arnhold and Bueno-Aniola 2005).

---

<sup>50</sup> The linguistic characteristics of São Paulo’s varieties are presented in 3.4.1 and 6.1.

<sup>51</sup> My own understanding of the varieties important for this study is thoroughly explained in 6.1. Generally spoken, I understand *Norma Culta* and *Popular* as theoretical concepts uniting various linguistic features that – due to their social indexicality – are rather associated with the first or the second.

Inside São Paulo, there are also surprisingly few studies that focus on smaller units. An exception consists of e.g. Beline Mendes' studies on 'gay speech' in São Paulo (cf. e.g. Beline Mendes 2006) that tested whether certain linguistic features were indexical for being gay in São Paulo. Regarding the immigrants' speech communities, apart from Niehoff (2006), no study has actually employed urban sociolinguistics to study the Portuguese of these people. Her study focuses on the Italian immigrants, studying whether linguistic features can actually be identified in São Paulo's Portuguese that still are associated with the influence of the Italian immigrants.

### 3.1.5 Urban sociolinguistics in Bolivia

---

The urban centers of Bolivia, particularly La Paz/El Alto, Santa Cruz and Cochabamba are of particular interest, as they constitute places of massive migration from Bolivian rural areas that has led to language contact between Castilian, Aymará and Quechua. While similar processes are quite well described and analyzed for Lima, the Peruvian capital, sociolinguistic literature about Bolivian cities is limited to very few studies. Compared to that of its neighbor states, the entire linguistic study of Bolivia's language(s) is rather limited.

Gutiérrez Marrone (1992), Mendoza (1991) and Placencia (2001) investigate the *Paceño* variety of La Paz, but the first two focus rather on the influence of the indigenous languages on Castilian than on actual urban characteristics. As they primarily emphasize the differences between bilingual and Castilian monolingual speakers, I would rather perceive them as dialectological studies. Placencia (2001), however, explores stylistic variation as she researches how indigenous people are addressed in Castilian in public institutions.

The broadest study with regard to an urban variety of Bolivia, though, was conducted by Pfänder (2002 and particularly 2009). Focusing on morphosyntactic features evolving due to language contact between Quechua and Castilian in Cochabamba, he presents grammatical evolution as related to social principles:

*[E]l contacto lingüístico urbano debe entenderse como una forma de hibridación. Esta hibridación puede describirse por un lado como reoralización de la norma ejemplar [...], y por otro como influencia recíproca entre diferentes lenguas y registros de una lengua (Pfänder 2009: 96).*

As typically urban he defines personal networks that consist of weak ties, or various close-knit networks that are connected by weak ties that obey different linguistic norms (Pfänder 2009: 97), following not only Milroy's ethnographic studies but also stylistic variation as a product of social practice. Although these studies – complemented by further articles mostly focusing on the dialectologi-

cal description of the whole of Bolivia<sup>52</sup> – are sufficient for my aims, sociolinguistic research investigating language contact in Bolivia’s urban centers remains a desideratum that urgently has to be addressed.

---

<sup>52</sup> Cf. Coello Vila (1996), Crevels and Muysken (2009), Mendoza (2008) and Rocha (1987).

### 3.2 Language Contact and Second Language Acquisition

---

Traditionally, language contact and Second Language Acquisition (SLA) are investigated by different linguistic disciplines focusing on different aspects of the contact phenomenon. While language contact is mostly associated with sociolinguistics, studying “the end results of the social processes as reflected in language change” (Siegel 2003: 204), SLA is mostly studied from a psycholinguistic perspective, focusing rather on the “individual mechanisms that led to the changes in the first place” (l.c.). However, as I endeavor to explain not only the linguistic outcome of the contact situation of Bolivian immigrants in São Paulo through a sociolinguistic approach, but also the processes leading to this outcome in the first place, I include approaches from both traditions, despite the possible differences between the research traditions.

But what are these differences in the first place? SLA’s focus lies on the individual rather than on the speech community. This means, SLA is mainly concerned with model and theory building and the sampling of class room learners of varying proficiency, while language contact theory focuses rather on the nature of linguistic systems stemming from language contact and community members (l.c.).

This fact leads to the estimation of some scholars that sociolinguistics and SLA must be studied separately, because “sociolinguistics [...] is anchored in a research paradigm that has had great success in the study of majority language speech communities, [...] if not entirely monolingual speech communities” (Sankoff 2008: 639). However, “multilingual speakers do not exist in a sociolinguistic vacuum” (Muysken 2010: 287), only “in an ideal world [should] what happens at the micro level of the bilingual community [...] be the direct consequence of the behavior of individual bilinguals” (l.c.). According to Tarone (2007), “there is a long record of research on social causes of interlanguage variation dating back to the beginning of interlanguage study”. She describes SLA’s sociolinguistic approach as

one that studies the relationship between such social contextual variables as interlocutor, topic, or task and the formal features of learner language or interlanguage production (Tarone 2007: 837).

Akin to her study, Leather and James (1991: 309-311) enumerate results of various studies on social constraints of the acquisition of second language speech like attitude, social acceptance or sex – in classic sociolinguistics called variables.

Therefore, a sociolinguistic approach studying language contact seems rather applicable. In a migration situation like the one of Bolivians in São Paulo, i.e. in which immigrants are coming into a well-established society speaking another language, SLA simply must be considered as the central process, by subsuming thus language contact in a natural environment as well as in the classroom environment (also labeled as language learning situation (Faingold 2003: 5)). The distinction between second language acquisition and second language use (Ellis 1994: 13) primarily seems to reflect the learning in different environments too, because SLA is described, as “opposed to L2 [second language, SN] use, [as] concerned with the gradual attainment of linguistic competence in the L2, or, in other words, with the learning of the L2 grammar” (Siegel 2003: 194), while second language use (SLU) refers to the actual “L2 performance” (l.c.: 195). Although I clearly recognize the conceptual differences, the temporal distinction between first SLA and afterwards SLU (l.c.) seems artificial if a natural SLA environment is considered, where language learners like immigrants simply cannot afford to take their time to silently acquire the grammar, and only afterwards start communicating. The temporal aspect might be helpful for the study of pidgins and creole languages as suggested by Siegel (2003), but for a regular migration situation, acquisition and use must be considered as temporarily parallel processes, where learning of grammar and use are interdependent steps of SLA. I maintain, though, the distinction between natural and classroom environments when discussing the state of the art with regard to Spanish-Portuguese language contact in 3.2.2 and 3.2.3.

A further key concept of SLA that differs from sociolinguistic studies of language contact is the idea of interlanguage as “station[ ] in an incomplete process (Sankoff 2008: 639): “Interlanguage is the linguistic system evidenced when an adult L2 learner attempts to express meaning in a secondary language being learned” (cf. Selinker 1972: 213 f., cited in Tarone 2007: 845).

However, the concept of interlanguage is also interpretable from a sociolinguistic view: It is a language system that is characteristic of the variety of a speech community of immigrants, whose “incomplete” status only distinguishes their language slightly from a language system of native speakers that also constantly changes in daily communication. Of course, there are additional social characteristics of this community because of the migration amongst others the place of language learning or the status of assimilation, that have to be considered as further variables influencing the linguistic outcome. Still, it can be described in the same way as every other variety in comparison to former states of the language (L1) or to the varieties in contact (L2).

In the following segment, I will therefore discuss the state of the art of sociolinguistic as well as psycholinguistic approaches of language contact and SLA theories that explain the linguistic processes being effective in language contact situations. I rather concentrate on approaches applied for natural language learning situations, but approaches on Spanish-Portuguese SLA will also be discussed.

### 3.2.1 Process oriented approaches

---

As stated before, the main differences between the process oriented approaches to language contact are firstly the focus on either the individual process or the resulting situation for the community,<sup>53</sup> and secondly, the definition of the relation between internal and external factors leading to the observed linguistic outcome. Despite the different weighing of the single factors, there seems to be common agreement that both linguistic and social factors have to be taken into account. Even in Thomason and Kaufman's approach (1988), the influence of linguistic factors was never completely neglected, as it was often criticized. They merely stated that social pressure may reverse any linguistic constraint like markedness.

Thus, I present studies by way of example<sup>54</sup> in roughly chronological order from sociolinguistics and social psychology that attribute more importance to social factors (3.2.1.2 and 3.2.1.4) as well as one psycholinguistic approach that rather focuses upon linguistic factors (3.2.1.3). Further approaches are dedicated to the processes functioning when narrowly related languages come into contact (3.2.1.5 and 3.2.1.6). As many of these approaches refer to the concept of markedness as one of the linguistic constraints, I would like to briefly discuss the state of the art on this concept to ensure the understanding of the following approaches:

The concept of markedness was introduced by Jakobson and further elaborated by many scholars, e.g. Wurzel (1987) and Mayerthaler (1981). They argue that linguistic forms that are transparent, iconic and uniform are more natural and hence less marked, occur more often in the languages of the world, are learned earlier by children and also in SLA and constitute the base for language change and reanalysis. Additionally, one has to distinguish between universal markedness and language-internal markedness that refers to structural consistency.

---

<sup>53</sup> In my opinion, the individual must be the starting point to see patterns on the macro-level. Therefore I will include both aspects.

<sup>54</sup> Of course, this account is not complete, but only reflects my familiarity with Anglo-American and German literature.

Applied to SLA, the Markedness Differential Hypothesis (MDH) (Eckman 1977: 321, cited in Eckman 2008: 98) asserts that “those areas of the target language which differ from the native language and are more marked than the native language will be difficult”, but those which are less marked will not. Additionally, the relative degree of difficulty corresponds to the relative degree of markedness.

Reich (2004: 63) even adds the dimension of inner-Romance harmony (*harmonia românica*) claiming that “itens lexicais e estruturas gramaticais são menos marcados se as duas línguas românicas os compartilham e mais marcados se existem em só uma dessas línguas.”

Mufwene (1990: 11) however, employs it for several factors such as “simplicity, generality, frequency, semantic transparency and salience”.

According to Faingold (2003:4), though, the principles mentioned above “can be overruled by the borrowing of prestige structures and other sociocommunicational developments, as well as by higher-level developments; for instance, reversals in marked categories or environments – fusion, violations to the principle of constructional iconicity, markedness reversal and so forth.”

Therefore, it seems that apart from the empirical data, there is little agreement about the concept of markedness, and the term is used rather carelessly. For this study, I will refer to markedness only in the sense of Wurzel (1987) and Mayerthaler (1981), but it seems that for typologically narrow related languages, the concept loses its explanatory power, even if Schmid (1995) confirms its validity for Spanish-Italian language contact. Thomason and Kaufman (1988: 165) themselves admit with regards to pidgins and creoles that markedness offers valid explanations “only when the structures of the substrate languages do not coincide substantially”. This finding is obviously due to frequency effects (Siegel 2008: 162) that should be transferable to SLA even if it does not result in a pidgin or creole language.<sup>55</sup>

### 3.2.1.1 Borrowing and shift-induced interference<sup>56</sup>

The publication of Thomason and Kaufman (1988) was very influential in the field of language contact; according to Muysken (2010: 269), it even “caused a major upheaval [...] since they argue that the idea that there are intrinsic constraints on language contact should be abandoned.” Instead, they

---

<sup>55</sup> For the parallels between SLA and the development of pidgins and creoles see e.g. Siegel (2003).

<sup>56</sup> Sankoff (2008: 641) and Krefeld (2004: 38) use the term “substratum interference” instead of shift-induced interference, though Thomason and Kaufman themselves state that “the label ‘substratum’ is inappropriate because it is too narrow” (Thomason and Kaufmann, 1988: 37). Thus, following Thomason’s reference to the study of 1988 (2003: 691), I call it “shift-induced interference”.

argue for the decisive role of social constraints in the contact situation, i.e. relative social pressure of one language or speech community on the other one. Depending on the degree of pressure on the speech community in the contact situation, two types of contact situation can be distinguished: either the speech community maintains its original language, leading to borrowing, or the speech community shifts to a new language, leading to shift-induced interference. Borrowing is broadly defined as “the incorporation of foreign elements into the speakers’ native language” (Thomason and Kaufman 1988: 21), while shift-induced interference occurs when the target language which the speakers have shifted to is not perfectly learned by the new speakers, and the original native speakers of the target language include the “errors“ into their language. In a borrowing situation the interference of vocabulary will be most prominent; phonology and morphosyntax are only borrowed in situations of most intense language contact. However, Thomason (2001) warns that

*[...] the largest class of exceptions will be found in contact situations involving two typologically similar languages. [...] [L]anguages that are typologically very different are likely to follow the borrowing scale closely, while languages that are typologically very similar are likely not to do so in all respects (l.c.: 71).*

According to the Thomason and Kaufman (1988), in a shift situation, there is the obverse scale: The interference of phonology and morphosyntax will be most prominent while the interference of vocabulary will be rather negligible.

The approach was broadly accepted for emphasizing the role of social factors and “bring[ing] together the macro-level of the language and the micro-level of the individual speaker” (Sankoff 2008: 641). However, other aspects were criticized: Krefeld (2004: 38), for example, points to the important aspect that in synchronic observation of migration situations, the speech community of immigrants fulfills the conditions for both contact situations. In the first generation *borrowing* dominates, in the second generation *shift-induced interference*, if the latter acquires the language of its (parents’ new) home-country as its native language. Sankoff (2008), on the other hand, rather criticizes methodical shortcomings that led to disregarding internal linguistic factors:

*[L]acking a quantitative perspective, [they] are forced to deny the importance of internal linguistic factors. [...] [I]n rejecting the contribution of internal linguistic structure, [Thomason and Kaufman] have thrown out the baby with the bathwater” (Sankoff 2008: 640 and 641).*

This equating of method and theory seems to fall rather short, as there are many qualitative studies proving the existence and importance of linguistic factors (cf. e.g. Muysken 2000, Siegel 2008).

However, Thomason herself softens the old thesis, emphasizing that linguistic factors like universal markedness or typological distance between source and recipient languages

*are important, but they are less important than the social factors [...]. By 'less important' I mean that the linguistic factors can be overridden by social factors pushing in an opposite direction" (Thomason 2001: 77).*

This softened thesis, i.e. external and internal factors both have to be considered for the explanation of language contact, is widely agreed upon in sociolinguistics as well as in psycholinguistics.

However, as e.g. Siegel (2008) points out,

*'structural borrowing' or 'substratum interference' [...] are normally descriptive terms that refer to an end result in language change, not to a psycholinguistic process (l.c.: 105).*

Thomason (2003) further modifies the approach, emphasizing that instead of language maintenance and shift, *imperfect learning* constitutes the decisive social factor for either borrowing or shift-induced interference; i.e. in contact situations where imperfect learning is involved, there is the interference of phonology and morphosyntax, in contact situations where imperfect learning is absent, borrowing of lexical items (Thomason 2003: 692). In my opinion, the factor *imperfect learning* points to a rather psycholinguistic perspective of SLA, akin to the constraints established by Coetsem (1988) that will be discussed in 3.2.1.2. As a sociolinguistic factor, the notion of *perfect language learning* seems rather normative, vague and difficult to apply in a communication situation. Considering the complexity of different styles and varieties particularly in an urban environment, who decides whether or not the acquisition is already perfect? Which *language* has to be learned perfectly, e.g. in languages without highly standardized linguistic norms? In Thomason (2001), *imperfect language learning* is classified as one of various social factors like *intensity of contact* and *speaker's attitude*, an approach that seems to be more convincing than the concentration on acquisition alone.

### 3.2.1.2 Borrowing and imposition

---

An account regularly named together with Thomason and Kaufman's approach (1988) is Coetsem (1988). It uses almost the same terms for the characterization of the two transfer types at work in language contact situations, but takes a psycholinguistic perspective. The approach – originally laid out for phonology<sup>57</sup> – defines the two transfer types as *borrowing* and *imposition*, depending on the psycholinguistic criterion of linguistic dominance concerning the individual bilingual speaker; i.e. the language the speaker shows major fluency and proficiency in is linguistically dominant. In his mod-

---

<sup>57</sup> The author states, however, that "the basic principles [...] have a wider range of application than phonology alone" (Coetsem 1988: xvi).

el, Coetsem distinguishes recipient language agentivity leading to borrowing on the one hand and source language agentivity leading to imposition on the other hand. This means that, in cases where the recipient language is linguistically dominant, borrowing of mainly source language vocabulary into the recipient language is found, while borrowing of phonology and morphosyntax will be less prominent. The obverse case of source language agentivity describes imposition as the transfer from source language phonology and morphosyntax to the linguistically nondominant recipient language, while imposition of vocabulary will be more marginal (cf. Coetsem 1988: 7 ff.). SLA as in the case of immigrants coming to an urban environment where another language is dominant, is explicitly discussed as the prototypical case of source language agentivity (cf. l.c.: 47 ff.). As secondary mechanism of source language activity that might occur, apart from imposition that is certain to occur, the author describes imitation leading to second language acquisition:

Imitation as discussed for phonology constitutes an approximation of e.g. phonological categories of the *sl* that are not integrated in the *rl*'s native phonology of the speaker (cf. l.c.: 7). And further: “[...] Imitation is actually *acquisition*. [...] Imitation in *sl* (source language, SN) agentivity is *true* acquisition of the *rl* (recipient language, SN) by the *sl* speaker” (emphases of the author, Coetsem 1988: 17). Therefore, he concludes, “imitation in *sl* agentivity is [...] quite naturally realized in *second language acquisition*” (emphases of the author, l.c.: 17 f.).

Apart from this “externally induced change” (l.c.: 46), Coetsem discusses two further variable factors taking effect in SLA – according to Coetsem the natural case of source language activity: internally induced change and reduction. Internally induced change “represents the activation of latent developments in the *rl* itself” (Coetsem: 41), “obeying universal and language-specific principles” such as e.g. naturalness / markedness in the sense of Mayerthaler (1981) and Wurzel (1987) as discussed under 3.2.1. The fourth process Coetsem proposes as influential in the case of *sl-activity*, is reduction. He defines this process as proficiency-related as “not utilizing parts of the target language, because the speaker lacks the necessary knowledge” (Coetsem 1988: 50).

*In natural second language acquisition the adult learner [...] will normally not utilize or avoid utilizing those parts of the target language [...] that are most specific to the rl and most dissimilar from his own language [...] (Coetsem 1988: 51).*

With regard to the parts of the language that are imposed or reduced he states that the more stable “domains or subdomains” of the language, e.g. phonology, are imposed in the case of source language activity (cf. l.c.: 46). Only in the case of narrowly related languages in contact, inflectional morphology is also imposed (cf. l.c.: 48).

Nonetheless, similarly to the MDH discussed before (cf. 3.2.1), Coetsem points out that

*“[t]here are other cases where imposition (stability) is counteracted. The [...] learner language may have in one of its more stable language domains or subdomains a more marked, less optimal structure in comparison with a less marked, more optimal matching structure in the [...] target language. The sl speaker learning the rl may then readily acquire the less marked, more optimal rl structure, although [...] he would normally impose his own more marked, less optimal structure upon the rl. (Coetsem 1988: 48).*

He fails, though, to point to the fact that markedness can also be overruled by social constraints as argued by Thomason and Kaufman (1988).

During the process of acquisition, until acquisition of the second language is complete, imposition, internally induced change and reduction take place (cf. Coetsem 1988: 55). According to Coetsem the progress in this completing process depends first and foremost on time (cf. l.c.: 49 ff.): The longer people acquire a language, the less prominent imposition, internally induced change and reduction will be.

This rather simplifying concentration on the factor time stands in contrast to Coetsem’s own differentiation regarding the influence of social dominance on a speaker: He e.g. emphasizes that the motivation for borrowing or imposition are social factors like prestige and/or need (cf. l.c.: 14 ff.) and even claims that the social status of the languages in contact might depend on the “area of social activity” (l.c.: 17), hence allowing stylistic variation.

All in all, Coetsem’s approach (1988, 2000) is a more differentiated view of the linguistic processes taking effect in a language contact situation. However, the relation between social prestige and the processes must be elaborated, particularly if imposition is understood not as a one-to-one transfer of single grammatical structures, but like Siegel (2008: 106) as “transfer via [source language] activity” that implies the possibility that

*a feature of the [source language] does not have to be transferred to the [recipient language] as a whole; rather, ‘individual elements or aspects of it’ may be transferred (Coetsem 2000: 76, cited in Siegel 2008: 106).*

Siegel (2008) further elaborates on transfer in second language use (explicitly following Coetsem’s approach) as a strategy to compensate for insufficient proficiency in the second language:

*Transfer is thought to occur as learners (or former learners) fall back on their L1 knowledge when their knowledge of the L2<sup>58</sup> is inadequate to express what they want to say or to interpret what is being said to them (Siegel 2008: 106).*

---

<sup>58</sup> Knowledge in this context must be understood as synonym of *proficiency*.

Reflecting on the approach of Coetsem (2000), Siegel (2008) further states: “L1 knowledge is [...] a resource in communication, used unconsciously to compensate for insufficient L2 knowledge” (l.c.: 121).

### 3.2.1.3 Communication accommodation

---

Although communication accommodation was not primarily designed for explaining language contact situations, it has proved its applicability for describing the psycholinguistic processes in various studies about intercultural and interethnic communication (cf. e.g. Gudykunst 2005) as well as second language acquisition (cf. e.g. Beebe and Giles 1984).

The social psychological approach was first suggested in the 1970s and published in 1982 as speech accommodation theory (Gallois, Ogay and Giles 2005: 125). According to Beebe and Giles (1984), social psychological research differs from sociolinguistics as the latter sees “language [...] as an intrapersonal phenomenon, while for the first it constitutes an interpersonal phenomenon” (l.c.: 6). Thus, research focuses upon factors like subjective attitudes and perceptions of situations (l.c.: 5). With regard to SLA, social psychologists “seem to have concentrated more on the ‘why’ of SLA while SLA researchers tend to concentrate on the ‘what’” (Beebe and Giles 1984: 16).

Based on speech accommodation theory, Giles et al. (1987) presented an expanded interdisciplinary model named communication accommodation theory (CAT) that has been applied particularly to intergroup settings, for example inter-dialect, inter-generational or inter-cultural settings. As basic assumptions for this theoretical approach Gallois, Ogay and Giles (2005: 136) name that communicative interactions are embedded in the sociohistorical context and that communication is not only the exchange of referential meaning but also the negotiation of personal and social identities. The primary motivation for accommodation to their interlocutor’s perceived individual and group characteristics is to evoke social approval (cf. Bell 1991: 74), particularly of socially influential others (Gallois et al. 2005: 130), and/or to increase communication effectiveness (Coupland and Jaworski 1997: 242).

Accommodation is therefore defined as

*the process through which the interactants regulate their communication [...] in order to appear more like (accommodation) or distinct from each other (nonaccommodation) [...]*  
(l.c.: 137).

Perception tests showed that those speakers who adjusted their speech – in terms of e.g. pronunciation, speech rate and message content (Giles and Smith 1979, cited in Giles, Coupland and Coupland

1991: 25) – were, roughly said, evaluated more positively by their interlocutors than non-accommodating speakers (Coupland and Jaworski 1997: 242). Nonetheless, further studies of speech accommodation theory and CAT have revealed that the relation between positive evaluation and accommodation is more complex than that, i.e. it depends on various factors like what the interlocutors intend by the conversation, if their group orientation is convergent or divergent and whether further internal and external conditions of the conversation are estimated as important by them (cf. Gallois, Ogay and Giles 2005: 129). Trying hence to predict the consequences of stylistic choices (cf. Coupland and Jaworski 1997: 243) CAT hypothesized that accommodation

*[...] can function to index and achieve solidarity with or dissociation from a conversational partner reciprocally and dynamically (Giles et al. 1991: 2).*

Thus, the main motivation for accommodation is “the speakers’ need to gain another’s social approval” (l.c.: 19), a factor proportionally rising with the degree of convergence<sup>59</sup>. The costs for accommodation, though, might be the reduction or even loss of personal and social identity (cf. Giles et al. 1991: 21), a fact that has to be considered particularly for the valuation of nonaccommodational behavior.

Thakerar (1982) claimed that our subjective “perceptions and beliefs about [...] sociolinguistic behaviours” (cited in Coupland and Jaworski, 1997: 242) influence the accommodation more than objective conversational data. Hence,

*[t]he fact that speakers often accommodate not to an interlocutor’s actual speech but to what they believe to be the interlocutor’s speech has long been recognized in SAT (Thakerar, Giles and Cheshire 1982, cited in Bell 1991: 73).*

This means also that despite the strong reference of CAT to personal interaction, accommodation occurs even in settings where no direct interaction takes place, as Bell (1984 and 1991), for example, convincingly showed with his study of moderators working at radio stations and other mass media who accommodated to the stereotypical audience they believed themselves to have. If the accommodation occurs regularly, then,

*[...] the mechanics of everyday interpersonal convergences in important social networks are the breeding ground for longer-term shifts in individual as well as group-level usage (Giles et al. 1991: 20).*

---

<sup>59</sup> In speech accommodation theory, the terms *convergence* and *divergence* were used referring to the strategies called accommodation and nonaccommodation in CAT (cf. e.g. Gallois et al. 2005: 131).

Accordingly, even language shift of a whole community can be seen as the strongest manifestation of accommodation. Thus, accommodating strategies are not only important in concrete interpersonal interaction, but

*[a]t another level, accommodation strategies can characterize wholesale realignments of patterns of code or language selection, although again related to constellations of underlying beliefs, attitudes, and sociostructural conditions (Giles et al. 1991: 2).*

Considering this study's situation of language contact, CAT is also applicable here, explaining for example immigrants' maintenance of their native language in another country as nonaccommodation at the group level:

*[...] The same factors leading individuals to diverge in an intergroup encounter also contribute to many immigrant minorities' resistance to acquiring a host language with anything resembling native proficiency (Giles et al. 1991: 300).*

Therefore, CAT offers a broad approach for the inclusion of social context and stylistic variation as discussed earlier.

The question of which linguistic features are accommodated, though, remains unclear, as the authors themselves confirm (Beebe and Giles 1984: 10). Although Giles et al. (1991: 51) suggest exploring “[...] the specific contribution of indexical sociolinguistic variables, such as segmental phonology, rate, or information density”, they also claim that “[...] people use whatever resources are available to them in terms of accommodating to another” (l.c.: 17). Therefore, Bell (1984) rightly claims that

*[t]o the linguist, accommodation theory's chief deficiency is its linguistic naivety. [...] A close linguistic analysis of accommodated speech [...] should prove a rich testing ground for these theories (l.c.: 163).*

However, the studies on the stylistic dimension of linguistic variables that must be seen in the tradition of CAT have already reduced this “linguistic naivety” as discussed in 3.1.2.

#### 3.2.1.4 Koine, koineization and dialect leveling

---

Having focused on approaches aiming for explanations on language contact in general, the following segments concentrate on contact of closely related languages such as Spanish and Portuguese.

Koine, koineization and dialect leveling are terms referring to theoretical approaches that already combine CAT with language contact for the case of typologically related languages. For the elaboration here, I will refer mainly to the work of Siegel (1985), Williams and Kerswill (1999), Kerswill and Williams (2005) and Kerswill (2008 and 2010).

Referring to the original koine, “a variety of Ancient Greek which had come to supplant other, local Greek dialects” (Noonan 2010: 58), *koine* is defined as the stabilized result of mixing of linguistic

subsystems such as regional or literary dialects (Siegel 1985: 363). The linguistic processes that can lead to a koine are described by the further terms.

*Dialect leveling* and *koineization* both concern the contact of languages that are mutually intelligible or share “a superposed, genetically related linguistic system, such as a national standard or literary language” (Siegel 1985: 365). Thus, although the mutual intelligibility of Portuguese and Spanish is ambiguously discussed in literature (cf. Jensen 1989), processes like these must be considered for two languages as closely related as Spanish and Portuguese. Faingold (1996: 113 ff.), for example, discusses Judeo-Ibero-Romance and *Fronterizo*<sup>60</sup>, the Spanish-Portuguese variety spoken at the Brazilian-Uruguayan border, as koine.

According to Siegel (1985: 364 ff.), the distinction between dialect leveling and koineization is often unclear in literature, some even equate the terms. The author defines koineization as

*a dynamic process, usually of dialect leveling and mixing, of which the formation of a stabilized koine may be one stage” (Siegel 1985: 364).*

He emphasizes that a unique feature of koineization is constituted by the fact that “it involves the mixing of features of the different dialects, and leads to a new, compromise dialect” (Siegel 1985: 365), while dialect leveling leads to contact change effects, but not to the development of a compromise dialect. For Kerswill (2010),

*[k]oineization is the leveling of variant forms of the same linguistic items (especially phonemes and morphemes), and simplification<sup>61</sup> – the reduction of phonological and morphophonemic complexity” (l.c.: 231).*

As examples of this simplification, Kerswill (2008) names

*a decrease in irregularity in morphology and an increase in invariable word forms (quoting Trudgill 1986: 103), but also the loss of categories such as gender, of morphologically marked cases, simplified morphophonemics, and a decrease in the number of phonemes (l.c.: 671).*

*Dialect leveling* is defined as a

*[...] process whereby differences between regional varieties are reduced, features which make varieties distinctive disappear, and new features develop and are adopted by speakers over a wider geographical area (Williams and Kerswill 1999: 149)*

or as a process that

*involves the selective simplification and homogenization of patterns as they spread from one community or area to another [...]*” (Muysken 2010: 274).

---

<sup>60</sup> For the characterization and definition of *Fronterizo*, see 2.2.3.

<sup>61</sup> Trudgill (2010) defines simplification as consisting of three linked components: regularization of irregularities, increase in lexical and morphological transparency and the loss of redundancy (307 ff.). For further discussion of markedness and naturalness theory, please see 3.2.1.1

Linguistically, it is more specified as “the reduction of variation in exponents of phonological and morphological categories” (Kerswill and Williams 2005: 1024).

In other words, features that are marked, whether socially or geographically (cf. l.c.: 1041), or features that the minority of the speakers use (cf. Kerswill 2008: 675) are leveled out because the single varieties converge, “allowing one’s speech to conform to another’s” (Gambhir, cited in Siegel 1985: 367). In this way, “[l]eveling leads to a reduction in differences between dialects and hence a gradual homogenization of the vernacular speech of a region” (Kerswill 2008: 671).

Resorting to accommodation theory, it is claimed that

*[w]hen people speak different varieties, as in a new settlement, the dialect differences are likely to be exploited – consciously or passively – as part of accommodation (l.c.: 680).*

This application of CAT means that the accommodation does not take place only by the adoption of certain indexical linguistic markers as described for example by Coupland and Jaworski (1997), but also by the reduction of distinguishing features. For the description of the special features that are reduced first (and taken up first), scholars studying koineization and dialect leveling use the term *salience*<sup>62</sup> (cf. Trudgill 1986: 37, cited in Kerswill 2008), a concept not only used to describe the loss of dialectal features, but also their acquisition. Auer, Barden and Grosskopf (1998) show that this distinction has an important impact on the definition of salience, particularly on the factors influencing the salience of single features. They argue thus for “distinguishing more strictly between dialect loss and dialect acquisition as two components of long-term dialect accommodation” (l.c.: 168). Defining *salience*, both, the subjective perceptual side, including the use of certain features for stereotyping and mimicking, as well as the objective side, including factors like articulatory distance or areal distribution, have to be considered (cf. Auer, Barden and Grosskopf 1998: 167). In a study about dialect accommodation in Germany in the 1990s, Auer, Barden and Grosskopf (1998) find that subjective perceptual salience has more explanatory power than the objective factors enumerated (cf. l.c.: 184).<sup>63</sup> Considering the findings from above, that people accommodate not necessarily to the linguistic features of their interactional partner but to their stereotypical idea of her speech, this result does not come as a surprise, but only underlines the similarity between the concepts. Thus,

---

<sup>62</sup> According to Auer, Barden and Grosskopf (1998: 164), “the notion of salience has a tradition which goes back to the 1920s”, when it was used by the Russian dialectologist Schirmunski. The more common reference to Trudgill (1986) differs only slightly from Schirmunski’s approach, as Auer, Barden and Grosskopf (1998) point out.

<sup>63</sup> The influence of frequency is not considered in Auer, Barden and Grosskopf (1998)’s study, but e.g. Elmentaler, Gessinger and Wirrer (2010: 119) rightfully point out that this factor should be included in the definition of salience.

Lenz's (2010: 95) argument in favor of an unambiguous definition of *salience* at the subjective-attitudinal level, referring alone to cognitive salience (originally *Auffälligkeit*, SN), seems to be quite consistent with the findings of CAT.

Irrespective of the exact definition of *salience*, the occurrence of dialect leveling and koineization can only be explained in certain social situations, namely the continued and long interaction along with high integration of the speakers of different varieties, and the contact of varieties of about equal social prestige in the speakers' opinion (Faingold 1996: 114). In any other cases, social dominance would instead lead to the acquisition of the socially dominant variety by the socially subdued varieties and particularly the maintenance of the socially dominant variety as discussed by Thomason and Kaufman (1988). Thus, on a macro level, it seems rather improbable that dialect leveling or the developing of a koine are processes to describe the outcome of language contact in today's entire city of São Paulo, where Portuguese clearly has more social prestige than Spanish. However, focusing on minor units such as single Communities of Practice inside the city, situations where both languages have the same prestige are at least theoretically possible. In the history of São Paulo, in the 19<sup>th</sup> and the beginning of the 20<sup>th</sup> century, when massive migration of speakers of Romance languages formed the city, the concept of "immigrant koine", referring to a situation where the contact "takes place not in the region where the dialects originate, but in another location where large numbers of speakers of different regional dialects have migrated" (Siegel 1985: 364), might even have been the decisive process in the development of the city's language.

### *3.2.1.5 Code mixing*

---

Studies about code-mixing and code-switching<sup>64</sup> constitute an entire branch of the studies on language contact. According to Myers-Scotton (2002: 10), studies focusing on the "the grammatical structure of code-switching" and those studying the "socio-pragmatic side of code-switching" (l.c.) – the "social motivation for choosing to use one linguistic variety rather than another" (l.c.) – must be distinguished. For this study, the latter are more relevant. Of particular importance for this study is the approach by Muysken (2000), because it explicitly includes linguistic processes leading to *mixed codes* (cf. Muysken 2007) in contact situations of narrowly related languages.

---

<sup>64</sup> As the first term is used by Muysken (2000), I also employ it as cover term for the phenomena discussed below. Other approaches rather use the term code-switching (e.g. Myers-Scotton 2002, 2005; Milroy and Gordon 2006), as the terms are not unambiguously used in literature.

Muysken (cf. 2000: 3) distinguishes three different basic processes that he subsumes under the term code-mixing:

- *insertion* of an alien lexical or phrasal category from language B into a given structure from language A.
- *alternation* between structures from languages, and
- *congruent lexicalization* of material from different lexical inventories into a shared grammatical structure.

*Insertion* is restricted to occurrences where the inserted element of language B and the slot of language A (the matrix language<sup>65</sup>) are categorically or semantically congruent or equivalent (cf. Muysken 2007: 320). Mostly, single bare nouns, bare noun phrases (Muysken 2000: 95) or fixed phrases (Muysken 2007: 321) are inserted, similar to the process described above by borrowing (cf. 3.2.1.1 and 3.2.1.2). However, the author (l.c.) emphasizes that insertion can also concern multi-word constituents.

In the case of alternation, there is “a true switch from one language to the other, involving both grammar and lexicon” (l.c.: 5). With growing proficiency in the “new” language, it can even come to an “overall shift from one matrix language to another one” (Muysken 2000: 247.)

Insertion and alternation underlie syntactic constraints, e.g. they occur mostly at the peripheral position of the clause and at major clause boundaries.

Congruent lexicalization, as defined above, “violate[s] well-documented syntactic constraints” (Lip-ski 2010: 571). Therefore, it can only occur if the languages in contact share grammatical structure and display linear and structural equivalence, i.e. they are typologically closely related or even varieties of one language (cf. Muysken 2000: 122ff.).

According to Muysken (2000),

*congruent lexicalization is akin to language variation and style shifting: switching is grammatically unconstrained and characterizable in terms of alternative lexical insertions (l.c.: 221).*

---

<sup>65</sup> The concept of matrix language is broadly discussed, particularly in light of structural and theoretical approaches. As this aspect is of little interest for this study, I will not present these approaches here. Muysken himself criticizes approaches which only use structural criteria to define the matrix language, but states that “the notion of matrix language is essentially an empirical one [...] not a theoretical prime.” (2000: 68). He further points to the importance of dominance for the definition of the matrix language (cf. l.c.), however, without explaining whether he refers to dominance in the social or the individual sense (cf. 3.2.1.1 and 3.2.1.2). For this study, follow Müller (2003: 9), – I use matrix language as the language that constitutes the frame for insertion from another language.

It can contain a number of different switches of non-constituent elements, even function words. Pre-conditions for its applicability are linear and structural equivalence of both languages/varieties, extralinguistic factors like high bilingual proficiency and “no strong attitudinal barriers against mixing” (Muysken 2000: 231). In the case of congruent lexicalization, the two languages are simultaneously active (l.c.: 252), meaning it can lead to an actual mixture of the two languages. As Spanish and Portuguese are two languages that are as narrowly related as postulated by Muysken (2000), congruent lexicalization could play an important role for the understanding of the Portuguese of the Bolivian immigrants.

Furthermore, the author explicitly counts style-shifting, in sociolinguistics regularly employed for variation inside one historic language, to one “subtype of code-mixing” (l.c.: 123).

Although I agree with his thesis that the distinction between historical languages and varieties is not linguistically but rather politically motivated and thus language contact situations should also include the contact of varieties, it seems that stylistic variation should be the term that also subsumes code-mixing and not vice versa, seeing all varieties and languages as part of the speaker’s options to integrate them into her personal style.

Having presented various approaches dealing with the theoretical aspects of language contact, I next focus on the state of the art concerning language contact between the two languages in question, Spanish<sup>66</sup> and Portuguese.

### 3.2.2 Classroom contact situations between Spanish and Portuguese

Literature regarding classroom situations in which Spanish students learn Portuguese is quite ample. This is probably due to two political and economic developments, on the one side Brazil’s enormous economic progress amplifying its attractiveness for Spanish-speaking students from all over the world and on the other side the creation of the *Mercosur*, facilitating the exchange between people coming from its Hispano-American member countries and Brazil. Interestingly, there is also plentiful literature from the US dealing with Spanish-speaking students that either belong to the strong Hispano minority in the US or have learned Spanish as a second language and want to learn Portuguese next. Some studies in this regard still focus on a contrastive analysis of the two languages, but, according to Carvalho (cf. 2002: 601 ff.) the strong thesis of contrastive analysis to predict and pur-

---

<sup>66</sup> Referring to Spanish all over the world, I use this umbrella term. For the Spanish spoken in Bolivia, I employ the term *Castilian*, because it is used by the Bolivians themselves.

posefully handle errors of students was diminished to the “weak thesis of contrastive analysis” working with error analysis and not with its previewing, because

*“vários estudos subseqüentes mostraram que muitos dos erros previstos pela análise contrastiva não ocorrem em L2, ao mesmo tempo que muitos dos erros que ocorrem não podem ser atribuídos à interferência da L1 [...].”*

Additionally, many studies investigate how to best teach the other language, focusing primarily on the narrow relatedness between the two languages that leads to the “notável rapidez no processo de aprendizagem do português por falantes de espanhol” (Grannier 2000), but also to *fossilization*, the maintenance of normatively false structures leading to a relatively strong accent in Portuguese, even after years of learning<sup>67</sup> (cf. Paes Almeida Filho 2004: 185 ff.).

Therefore, articles like Carita (1998), Grannier (2000 and 2004), Jensen (1989 and 2004), Leiria (1998), Simões and Kelm (1991), Simões, Carvalho and Wiedemann (2004) and Wiedemann (2008) present well informed empirical data about classroom experiences, focusing on typically observed errors, particularly at the phonological and morpho-syntactical level, their possible origination from the transfer of Spanish structures and the consequences teachers have to draw from these results. For my analysis, these studies are very important, since they point once more to the features that are of interest in contact situations between the two languages and also serve as a linguistic database to which my results can be compared, just because the social settings are so different.

The broadest study of Spanish students learning Portuguese is presented by Akerberg (2002): Focusing on L1-influence (*transfêrência*), she also distinguishes between structural, factors including markedness and typological distance (l.c.: 20 ff.), but also takes social and psychological factors into account.. Additionally, she points to the special role of perception for the development of the L2’s phonological system and even its orthographic representations. Arguing with the Speech Learning Model (SLM) developed by Flege (cf. e.g. Flege 1995 and Bohn 1998), she claims that similar sounds are more difficult to perceive and acquire than dissimilar sounds, because “category formation for an L2 sound may be blocked by the mechanism of equivalence classification” (Flege 1995: 239). Applied to Spanish and Portuguese language contact,

*a proximidade das línguas faz com que os alunos de língua espanhola reconheçam muitas palavras da sua primeira língua e não façam um processamento detalhado dos sons (Akerberg 2002: 23).*

---

<sup>67</sup>Akerberg (2002) defines fossilization as “manter itens lingüísticos, regras, e subsistemas da interlíngua apesar de receber informação e explicações de como deveriam ser na língua alvo” (l.c.: 5).

This means also that in many cases SLM may contradict markedness theories, because the features that are less marked but still too similar will not be acquired as predicted by markedness theories. Thus, Leather and James (1991: 326) claim that “the differential effects of L1 influence over time on L2 speech acquisition must be evaluated in the light of the competing forces of the typological (markedness) value of L2 phones as well as learner-perceived similarity between the phones of L2 and L1 phones.” In a single case study, Pestana Allegro and Madureira (2008) confirmed Flege’s SLM for a Brazilian student of Spanish, focusing on the class of vibrants. Even if this study concerned the obverse language acquisition situation, it seems that SLM constitutes a valuable alternative to markedness theory with regard to the phonological analysis of the Bolivian speakers.

### 3.2.3 Natural contact situations between Spanish and Portuguese

---

In contrast to the rather rich literature about classroom contact situations between Spanish and Portuguese, literature focusing on natural contact situations between Spanish and Portuguese seems fairly limited. Despite the continuous geographical neighbourhood of Spanish and Portuguese speaking regions in Europe and Latin America and the mutual migration between the corresponding countries, only the situation at the Brazilian-Uruguayan border has been studied thoroughly. Contact situations caused by migration like that of Bolivian immigrants to São Paulo have never been investigated linguistically.<sup>68</sup>

In Europe, despite the common border of Spain and Portugal, the contact situations are rare as “historical tensions between Spain and Portugal have precluded the formation of stable contact varieties” (Lipski 2010: 566). According to Lipski (l.c.) and Clements (2009: 191), there are some “residual isolates” of Portuguese found in the Extremadura, for example Valencia-de-Alcántara, that are linguistically unstudied. Only *Barranquenho*, “a variety of Spanish-influenced Portuguese spoken in Barrancos” (Clements 2009: 190) has achieved some attention by linguists, lastly by Clements (2009), whose study focuses mainly on phonological and morphological features that define this variety and are all “traceable to Spanish” (l.c.: 208), while taking into account social factors like group identity and social prestige.

---

<sup>68</sup> The so-called *Ladino*, the language spoken by Sephardic Jews that includes Portuguese as well as Spanish linguistic features, because of the history of subsequent expulsion first from Portugal and then from Spain, constitutes a certain counter example (cf. e.g. Busse 1999, Faingold 1996, Rehrmann and Koechert 1999), but as the processes and linguistic origins involved here are much more complex, I will not consider this case.

In contrast to *Barranquenho*, *Fronterizo*<sup>69</sup>, the variety spoken in the border region of Brazil and Uruguay, is very well studied as the only known stable variety stemming from a language contact situation between Spanish and Portuguese (Lipski 2010: 566). Thun (1986) describes even the existence of “Fronterizo linguistics” that has developed since the nineteenth century, including a thorough linguistic description from a sociolinguistic-dialectological perspective that lead to the thorough description and analysis of *dialectos portugueses del Uruguay* (DPU) in the *Atlas lingüístico Diatópico y Diastrático del Uruguay* (Thun, Elizaincín and Boller 2000). Here, dialectology and sociolinguistics were combined, including social factors like residence, age, sex, etc. (Thun 2000: 289). The various studies on DPU, particularly by Elizaincín (1984, 1992, 1996, 2002), Thun (1986, 2000), but also by Frank Kersch (2006) and Lipski (2010), see *Fronterizo* as hybrid dialects inside the Portuguese-Spanish continuum. Faingold (1996: 132 ff.) explicitly calls *Fronterizo* a koiné characterized by the “loss of marked and minority forms”.

Carvalho (2003a, 2003b and 2003c) and Garrido Meirelles (2009), however, point out that the variety is actually located on a continuum “ranging from rural Uruguayan Portuguese to standard Portuguese” (Carvalho, 2003c: 643) that currently – influenced by urbanization and greater exposure to standard Brazilian Portuguese – is undergoing linguistic change “from the highly focused rural dialect to a less stigmatized urban version” (l.c.), even if complemented by the characteristics of border Spanish and its variation. According to Garrido Meirelles (2009: 274), for example, *Português Uruguayo* (PU)’s phonological system clearly derives from standard Brazilian Portuguese; hence PU is too close to Brazilian to affirm the thesis of a mixed language. In her study of Brazilian border regions, Blaser (1995: 161 ff.) comes to the similar conclusion that there is no observable mixed dialect. She emphasizes though, that in Uruguay, the situation is different.

Although the social and linguistic setting reveal great differences between the language contact situation in current São Paulo and *Fronterizo*, the many empirical sociolinguistic studies focusing mainly on oral data, investigating phonological as well as morpho-syntactical phenomena, help with the

---

<sup>69</sup> The name *Fronterizo*, derived from *frontera* (border), stems from Rona (1965), the first author systematically describing “un dialecto mixto, con base aparentemente portuguesa” (Rona 1965, cited in Thun, 1986: 47), but its speakers themselves refer to their language as *português*, *castelhano* or *carimbado* (cf. Thun, 1986: 59 ff.). Contradicting Lipski (2010: 566), Thun (1986: 58) emphasizes that *portuñol* is not used by the speakers themselves, but rather pejoratively by other inhabitants of Uruguay. As more recent terms referring to this variety, first DPU (*dialectos portugueses del Uruguay*), introduced by Elizaincín (1992), next PU (*Português Uruguayo*), used by Carvalho (2003a, 2003b, 2003c), already imply a certain linguistic perspective on the variety, I use the older and apparently more neutral term *Fronterizo* without accepting Rona’s further elaborations. If I explicitly refer to definitions in the sense of Carvalho (2003a, 2003b, 2003c) or Elizaincín (1992), I use their terms.

analysis, because the studies investigating the language contact at the further borders of Brazil with its Hispano-American neighbour countries are very rare.

For the Bolivian border, unfortunately, I am only familiar with the short presentation of Lipski (2010: 568) that sketches some linguistic aspects of the language contact situation in the Bolivian border town *Cobija* and its Brazilian sister city *Brasiléia*<sup>70</sup>. Interestingly, Lipski further compares these data and *Fronterizo* data in regard to code-switching patterns. He explains his findings of “apparent randomness of the language mixture in Bolivia” (Lipski 2010: 573) with the classification of Muysken (2000) as discussed in 3.2.1.5. According to Lipski (2010), the mixing of Portuguese and Castilian elements are attributed

*not only to the high degree of shared structures between the two languages but to the [speakers'] limited proficiency in the second language [...] (l.c.: 573).*

This way he combines the findings of e.g. Thomason's (2006) and Coetsem's (1988, 2000) as discussed above with the approach of Muysken (2000). Literature about the case of Spanish-Portuguese contact is restricted to a few mostly anthropological studies that constitute the basis for Chapter 3 in which the historical, socioeconomic, geographic and linguistic aspects of Bolivian immigration will be discussed. Particularly important are the studies of Silva (2005, 2006 and 2008) as he refers to his own observations at the *Centro de Estudos Migratórios*, but also the sociogeographic approaches of Freire da Silva (2008) and Cymbalista and Xavier (2007), as well as the economical view presented by Gomes Azevedo (2005) and Cacciamali and Gomes Azevedo (2006).

### 3.3 Holistic approaches

---

To conclude this insight into the state of the art of sociolinguistics and language contact, I present two approaches that aim for a holistic approach towards the phenomenon, regarding linguistic processes and social factors to explain language contact caused first and foremost by migration. As I also aim for an approach that enlightens processes and outcome at an individual and societal level, these approaches – *Migrationslinguistik* (migration linguistics) and *Urbane Prozesse* (urban processes) – are of particular relevance for this study.

#### 3.3.1 *Migrationslinguistik*<sup>71</sup>

---

<sup>70</sup> The linguistic investigation of the border areas constitutes a lamentable gap in investigation and major challenge for both Bolivian and Brazilian linguists, particularly as it could shed further light not only on *Fronterizo* but also on language contact in Latin America and language change per se.

<sup>71</sup> *Migration linguistics* (own translation, sn); as I refer only to Krefeld (2004), I use the German term.

The approach *Migrationslinguistik* deals with the particularities of language contact caused by migrational movements. The author claims that mass migration is an important source for linguistic dynamics that has to be studied by the discrete discipline of *Migrationslinguistik* (l.c.: 147), in which sociolinguistic and contact-linguistic aspects are systematically subordinated to the concept of communicational space (Krefeld 2004: 18). Indeed, the studies discussed above do not mention the concept of space and spatiality at all, even though the change of place clearly constitutes the core aspect of migration. According to Krefeld (2004), communicative space consists of three dimensions (l.c.: 22)<sup>72</sup>: the *language's* spatiality, stemming from the association of certain linguistic varieties with regions or national states (areality and territoriality), the *speaker's* spatiality reflecting provenience and mobility and the spatiality of *speaking* itself, referring to the relative proximity/distance<sup>73</sup> between the interacting people. Considering these three dimensions he achieves the classification of various types of communicative spaces, including contact and isolation phenomena in the *language* dimension, social networks and imperfect learning in the *speaker* dimension and also code-switching and code-mixing in the *speaking* dimension. In this way, the author combines various aspects of the language contact theories I already discussed with sociolinguistic approaches, and convincingly applies them to the data of Italian immigrants in Germany.

It seems, however, rather difficult to understand his concept of *communicative space* as explicitly non-metaphoric: Krefeld (2004) repeatedly states that the social aspect is subordinated to the spatial aspect. This concept might be applicable to countries like Germany or Italy – the countries studied by Krefeld (2004) –, where the linguistic division might follow more along the line of local dialects than of social stereotypes. For Brazil and other Latin American countries which are highly socially stratified, social status overrides spatiality. However, the author himself refers to Luckmann and Simmel, German sociologists, who clearly focus on space as constructed by social conditions: “Not geographic space but the allocation and centralization of its parts achieved by the soul have societal

---

<sup>72</sup> Cf. Krefeld (2004: 22): “Der kommunikative Raum konstituiert sich so gewissermaßen als dreidimensionaler Verbund, der die folgenden Einzeldimensionen umfasst:

1. Die Räumlichkeit der SPRACHE (im Sinn der AREALITÄT und der TERRITORIALITÄT),
2. die Räumlichkeit des SPRECHERS (im Sinn der PROVENIENZ und MOBILITÄT),
3. die situative Räumlichkeit des SPRECHENS (im Sinn der POSITIONALITÄT der Kommunikanten und ihrer INTERAKTION).“

<sup>73</sup> The concept of relative proximity and distance, introduced by Koch and Oesterreicher (1990), complements the three variety dimensions (diatopic, diastratic and diaphasic) established by Coseriu (1988, cited in Krefeld 2004: 147).

relevance (Simmel 1908: 615, cited in Krefeld 2004: 20, own translation, sn).<sup>74</sup> Moreover, Krefeld affirms that every speaker constructs her own communicative space by the application of linguistic features (Krefeld 2004: 20). Only this latter interpretation makes sense in regard to the complex context of migration to urban environments, because people from many regions and countries gather there. The meaning of spatiality for social relations and social status is here transferred; it matters only with regard to the social status associated with speakers from a certain area. On the other hand, in order to describe and explain the various social relations between the communicating people, the concepts of social networks or CoP (see 2.1.3) seem to be rather sufficient and more telling than *communicative space*.

Thus, the various dimensions discussed above should be reflected in this study, but the focus upon the spatial aspect instead of the social aspect of space seems rather problematic for the urban environment.

### 3.3.2 Urbane Prozesse<sup>75</sup>

---

The last theoretical approach presented here is the most recent one. Ploog and Reich (2005 and 2006) combine urban sociolinguistics with language contact by explicitly integrating the multitude of languages and varieties stemming from migration. They criticize the model of Krefeld (2004) just discussed, as only theoretically applicable to urban conglomerates, because it does not consider the specific social and therefore linguistic heterogeneity of the city (Ploog and Reich 2006: 225). Instead, they suggest four dimensions that constitute the dynamics of the city as communicative space: *structural resources*, namely the various languages and varieties in contact; *communicative areas*, i.e. “las diferentes situaciones cotidianas en las que hablamos y escribimos” (Ploog and Reich, 2005: 48); *pragmatic requirements*, and *socio-indexical features*<sup>76</sup>. In a natural situation of language contact, the authors confirm a certain influence of transference/interference as discussed in the approaches of Thomason and Kaufman or Coetsem (cf. 3.2.1.1 and 3.2.1.2), although they claim that these approaches do not go far enough. However, they do not differentiate which understanding of

---

<sup>74</sup> Cf. “Nicht der Raum, sondern die von der Seele her erfolgende Gliederung und Zusammenfassung seiner Teile hat gesellschaftliche Bedeutung.“

<sup>75</sup> *Urban Processes* (own translation, sn); I again maintain the German term to avoid misunderstandings.

<sup>76</sup> Although Ploog and Reich (2006) denote this dimension (25) *sozioindexikalische Einstellungen* (socio-indexical attitudes, own translation, sn), I will maintain socio-indexical features, because the authors themselves state that “existen rasgos socio-indexicales en el lenguaje que revelan las actitudes del locutor frente a la comunidad en la que vive y que codifican la subcategorización social” (Ploog and Reich 2005: 48),

transference/interference – e.g. borrowing, imposition, internally induced change or acquisition – underlies their approach:

*Los conceptos propuestos en Thomason and Kaufman 1988, Thomason 2001, Coetsem 1988, Coetsem 2000 nos parecen ciertos, pero insuficientes” (Ploog and Reich 2005: 48).*

But they claim that this transference process is complemented by the “selection of structural options that are already part of the languages’ internal range of variation”<sup>77</sup> (Ploog and Reich 2006: 228, own translation, sn). The selection follows the two principles of transparency and structural harmony (*strukturelle Harmonie*, own translation, sn), i.e. features that are more transparent are (mostly) maintained in a language contact situation and features that are present in both languages are maintained.<sup>78</sup> The authors offer a detailed discussion of transparency, distinguishing the different application to syntax, morphology and phonology and referring to examples in São Paulo, Lima and Abidjan.

As the cities are further characterized by size of population, density and heterogeneity, communicative areas and pragmatic requirements are more plentiful than in rural areas; by the use of socio-indexical features, the recently arrived have to adjust to these requirements and hence achieve “the goal of every immigrant to find his place in society as soon as possible” (Ploog and Reich 2006: 242, own translation, sn). In this way, stylistic variation, as conditioned by different communities of practice (cf. e.g. Eckert 2000), is apparently integrated into the approach.

Thus, *Urbane Prozesse* offers an interesting design combining urban sociolinguistics with language contact. In contrast to the other language contact approaches, they explicitly integrate not only the new communicative space as suggested by Krefeld (2004), but the social composition of the immigrants’ new homes into their design – in Latin America, cities and metropolitan areas of several million inhabitants. Their resort to stylistic variation seems to be the right way to combine language contact with the urban environment for the Bolivian immigrants in São Paulo. However, the linguistic processes taking effect in the contact situation and the interplay between language contact, style and social indexicality, in particular, must be elaborated more thoroughly.

### 3.4 Summing up

---

<sup>77</sup> “[D]ie Auswahl struktureller Optionen, die im internen Variationsfächer der Sprachen schon angelegt sind.“

<sup>78</sup> The authors themselves state that this process is called koineization in the literature (Ploog and Reich 2006: 228); indeed, referring to the definition of koineization as presented in 3.2.1.4 (simplification and the leveling of variant forms) there are parallel processes, in particular in regard of morphology and phonology.

In this chapter I discussed the state of the art of the research lines in whose tradition I locate the study of language contact between Spanish and Portuguese in the special case of Bolivian migration to an urban conglomerate like São Paulo: Urban sociolinguistics, language contact theories and “holistic” approaches that combine social factors caused by migration and language contact theory.

I described how urban sociolinguistics developed three branches of research that differ with regard to their theoretical approach, their perspective and their methodology. Its focus changed from the connection between linguistic features and broad socioeconomic classes studied by quantitative methodology to more detailed local categories like social networks investigated by ethnographic studies, and eventually to the perspective that concentrates on stylistic variation as a manifestation of social practice. In Brazil, however, this advance is hardly reflected in the literature. The quantifying variationist approach is still clearly dominant there. , studies investigating stylistic variation are rare. In addition, the role of language contact in the cities is little investigated, although immigration is highly influential for the development of the Brazilian megacities. Instead, the studies on contact between Spanish and Portuguese focus either on natural contact situations in the Brazilian-Uruguayan border region or on Spanish students learning Portuguese in classroom environments.

In addition, studies on language contact scenarios were discussed above. Different sociolinguistic and psycholinguistic approaches on the linguistic processes caused by language contact and its outcomes were presented and examined for their explanatory power.

The two “holistic” approaches combine many of the elements presented before and present a whole model on the linguistic outcome of language contact caused by migration. Since the latter are not entirely convincing for my purpose, I will present a refined combination of the concepts introduced so far.

#### 4. A theoretical approach to language contact caused by migration

---

In this chapter, I postulate my hypotheses on the linguistic processes effective during language contact situations. Based on some of the theoretical approaches presented in Chapter 3 and a sociological approach that combines transnationalism and assimilation theories, I develop a model of accommodation as central process of a language contact situation due to migration. Depending on the social networks of the migrants they will tend to accommodate or to nonaccommodate by the employment of socio-indexical features. In addition, L1-transfer and simplification are discussed as processes that are also responsible for the linguistic outcome of migrants learning a new language. In the special case of the Bolivian immigrants in São Paulo, the transnational networks and the closeness of the two languages allows for translanguaging networks, i.e. the limits between the languages are fluent. Accordingly, in this case, the L1-transfer can be equaled to congruent lexicalization. The analysis of the Bolivians' speech in their social networks requires a qualitative approach focusing on the attribution of linguistic differences to accommodation, transfer or simplification. As social indexicality is a product of social negotiation inside the receiving society, the perceptual salience of features must be measured by quantitative studies. Thus, I suggest mixed-method research as methodology for the study of language contact in urban environments.

---

My own theoretical approach is based on the approaches discussed in Chapter 3. Many of them have explanatory power regarding one process that must be taken into account, but they do not present a holistic approach towards the speech of immigrants, i.e. a theory that could convincingly describe the outcome of language contact caused by migration as well as the processes underlying it. The two approaches that go in this direction – *Migrationslinguistik* and *Urbane Prozesse* – are not entirely persuasive. While the *Migrationslinguistik's* focus on space seems to make it more applicable for the countryside than for a megacity like São Paulo, *Urbane Prozesse* fails to clearly explain the connection between social indexicality (first and second order), social networks and the learning of another language in a foreign city. Additionally, it does not explain clearly why immigrants speak differently; if actually all urban networks are rather open, uniplex, with weak ties, why should this be? Is it only because of individual choice or because of structural influence of the society?

In order to answer these questions I suggest a theoretical approach that confirms the validity of universal processes and factors such as simplification and the Speech Learning Model as well as of L1-transfer. Additionally, though, I claim that accommodation is one of the central processes in (natural) SLA, conditioning and limiting the influence of the other processes. But who does accommodate, how, and why?

## 4.1 The sociological basis

---

As an extensive account of the many insightful findings of sociology in the field of migration would go beyond the limits of this study, I shortly present an approach that allows for the understanding of the linguistic theory. I refer mostly to Bommers (2005), which combines assimilation theories with transnational / multicultural approaches.

Assimilation theories “claim that a theoretically reflected concept of assimilation still provides the best frame for the analyses of even the most recent immigration processes” (Bommers 2005: 16). Jenny and Obaid (2004: 57) define assimilation as the “‘one-way integration’ whereby newcomers renounce their cultural habits and values in favour of the culture of the receiving society”. However, transnational approaches criticize that today’s society, seen as a “nationally closed container”, cannot be the reference of assimilation any longer, because the importance of transnational networks and the multicultural society has strongly increased in the last years: “[M]ore and more migrants are becoming so-called transmigrants (Bommers 2005: 15). This means that their life is dominated by their participation in transnational social relations, i.e., they frequently migrate between different countries and therefore develop stable transnational networks (cf. Pries 1997, cited in Petendra 2007: 12 ff.). Accordingly, theories that still see immigration as a one-way-movement from one national state to another would have no explanatory power for this *transmigration* (cf. l.c.). Answering this critique, Bommers (2005: 18 f.) points out that in today’s differentiated societies “the society” can indeed not be the point of reference for the migrants. Instead, the migrants “take roles inside organisations and fulfil the bundles of social expectations linked with these roles” (l.c.), e.g. in their profession, in the educational system or the health system of the state they migrated to. In short “even transnational migrants do have to assimilate - to the expectations of those social systems in which they want to participate” (l.c.: 23). Going even further, Bommers (2005: 19) claims that “all individuals in modern society must assimilate”, because everyone in modern society has to orientate to the conditions of participation in the differentiated social systems – and to develop corresponding competence and willingness to participate (cf. l.c.).

Explaining this process more thoroughly, assimilation theories claim that the migrant, mostly driven by economic reasons, undergoes four dimensions of assimilation – cognitive, social, structural and identificational assimilation – to be accepted by national society (Petendra 2007: 12) and therefore gain access to the most important social resources, particularly to education (Esser 2001).

According to Esser (1980), these dimensions contain the following:

- **Cognitive assimilation:** *Individual learning of languages, skills, behavioral and situational patterns, normative knowledge, orientations towards mobility etc.*
- **Structural assimilation:** *formal education, income, legal position, profession, taking membership roles in organizations.*
- **Social assimilation:** *friendships, marriage, clubs and other associations.*
- **Identificational assimilation:** *the claims of belonging and identity made by the migrants themselves.*

Naturally, as one of the foci of this study is the dependency of cognitive assimilation – language learning – from further social factors, the interaction between these dimensions is of interest. According to Bommes (2005: 21), traditional assimilation theory describes a “strict coupling between the different forms of assimilation”, while it is the “main thesis of transnation[a]lism [...] that globalisation leads to a loose coupling of the forms of assimilation” (l.c.: 22).

#### 4.2 The linguistic processes in language contact situations and their cause

---

In accordance with this sociological basis, one central linguistic process of language contact caused by migration must reflect the assimilation towards the receiving society, the receiving organizations or groups in the new country, respectively. As CAT<sup>79</sup> deals with the mutual adjustment between speaker and hearer, this approach offers the linguistic equivalent to assimilation: Accommodation.

Apart from that, L1-transfer does also play an important role in language contact situations because every person will resort to all linguistic sources she has as long as it is not socially restricted. Naturally, in the process of natural language learning, L1-transfer is one of the most important techniques.

But not all linguistic processes underlying the outcome of language contact should be associated with social factors. Universal processes of language development like morphological simplification will certainly play a role and appear even more radical than in “normal”<sup>80</sup> linguistic situations without language contact, due e.g. to the requirements of orality and the almost complete absence of normative language lessons.

All of these processes have been described before; thus, it is essential for my theoretical approach to hypothesize about the interdependencies between these processes.

---

<sup>79</sup> In the following elaboration I refer to secondary literature as presented in 2.2.1.3. I therefore abstain from renewed citation, although the references to the authors and publications cited above are obvious.

<sup>80</sup> From a global point of view, language history has proven that language contact at least between different dialects but also between different (national) languages is rather more far spread than unilingual situations.

#### 4.2.1 Accommodation toward social networks – by social indexicality

---

The logical translation of cognitive assimilation in linguistic terms is accommodation:

To recall its definition, accommodation is

*the process through which the interactants regulate their communication [...] in order to appear more like (accommodation) or distinct from each other (nonaccommodation) [...]*  
(Coupland and Jaworski 1997: 137).

This means that speakers accommodate to their audience, because they seek the approval of their audience. But who is the audience in the case of immigrants? Which language or variety will the immigrants accommodate to? In order to explain this, the social categories relevant for the immigrants have to be observed. As Bommers (2005) has rightly pointed out, today's society cannot be described by a static model of one national society. In accordance, urban sociolinguistics has found that smaller units than e.g. social classes are needed to explain the stylistic variation of the urban speaker. Therefore, I claim that the linguistic outcome of migration must also be investigated by finding smaller local categories, i.e. the actual social networks and CoP of the migrants.

The social network of the migrants is not a priori determined but constructed by various factors: For once, the attitude of herself towards the new country, towards the people living there and towards the language. These factors influence her social network because a positive attitude will cause her to go out and try to get in contact with persons of her new home. Accordingly, on one side, migrants seeking the approval of members of the receiving society will accommodate to their speech and thus learn the language or variety of the people in their newly-built network who are speakers of the "new" language or variety. This means that, different from Thomason and Kaufman (1988), it is not the social dominance in the whole speech communities that is decisive here, but the actual social contacts and linguistic input each immigrant gets. Naturally, factors like limited spatial mobility, illegality, poverty, a negative attitude of the receiving society with regard to immigration and xenophobia constitute obstacles to building a social network with many contacts outside the immigrants' own community.

On the other side, people who seek no approval of persons speaking the new language or have no contact with them at all because of the obstacles mentioned above, will nonaccommodate and consequently maintain their original language. Nonaccommodation, however, is only possible as long as the pressure from the new society on the individual to accommodate is not too high or as long as it does not outweigh the approval she gets by her community for the nonaccommodation.

Thus, the more members of the immigrants' original community her social network contains and the closer-knit the social network is, the higher is the probability that the speaker nonaccommodates. In the case of a looser-knit network which also contains contacts outside the original community, the chances of her accommodation, of learning the new language, are higher. Strong-tie contacts are particularly important, because – by definition – they are the contacts who frequently communicate with the individuals. Hence, they have a serious impact on the immigrants' language learning.

Which language or variety the speakers accommodate to, strongly depends on the contacts in the migrants' social network: The variety of the speakers they perceive as socially most successful will be the variety the immigrant accommodates to. If there are no other authorities like educational institutions that have even more influence on the individual, the variety that the immigrants evaluate as the one earning them most social approval, will become their target variety.

And how does accommodation actually work? In the studies of CAT, different scholars showed that speakers accommodated e.g. by their speech rate, their pronunciation or their message content, but they fail to explain why exactly these factors are used for accommodation.<sup>81</sup> As the motivation for accommodation is defined by the search for social approval, it is only logical that those features are employed that are connected with the social group the immigrant wants to impress or to identify with, i.e. socially indexical features. The concept of social indexicality according to Silverstein (1976, 1992 and 2003) describes how linguistic features are associated with certain social groups and how these features can hence be used for stylistic variation, displaying the speaker's identification with this group. As these indices are not based on the single estimation of individuals but on the "ideological work of society" (Eckert 2008: 465), by definition, only features that are perceived as highly salient can get such an indexical value. This means that language learners will accommodate by employing those features that they perceive as highly salient and socio-indexical for the receiving society. Features which are perceived as salient and socio-indexical for the original society will be reduced. Since vocabulary and phonology are most salient in the perception of the hearer, I suggest that in these linguistic areas the accommodation will be strongest.

---

<sup>81</sup> For more details, see 3.2.1.3.

Nevertheless, it has to be considered that the perception of language learners might differ from the perception of native speakers. As Bell's audience design (1984, 1991 and 2009) suggests, it is not necessarily the actual speech of the interacting person the speaker accommodates to, but the subjective idea the speaker has of it – this is where for example the Speech Learning Model might become relevant: Language learners might fail to realize the indexicality of phonological features because they are too similar to L1 of the language learner.

#### 4.2.2 Further processes: L1-transfer and simplification

---

At this point, it becomes obvious that the influence of the speaker's L1 also must be considered. Repeating the words of Siegel (2008):

*Transfer is thought to occur as learners (or former learners) fall back on their L1 knowledge when their knowledge of the L2<sup>82</sup> is inadequate to express what they want to say or to interpret what is being said to them (l.c.: 106).*

In the terms of Coetsem (1988, 2000), L1-transfer must not be understood as one-to-one transfer of entire structures, but of single features or tendencies. This means also that language-internal tendencies that are common for both languages might be strengthened; structures common for only one area of the linguistic system might be employed in other areas as well.

In some cases, language contact might even lead to simplification and reduction as described by Coetsem (1988) or by e.g. Kerswill (2008):<sup>83</sup>

*A decrease in irregularity in morphology and an increase in invariable word forms (quoting Trudgill 1986: 103), but also the loss of categories such as gender, of morphologically marked cases, simplified morphophonemics, and a decrease in the number of phonemes (Kerswill 2008: 671).*

Whenever L1-transfer and simplification point in the same direction, language change towards less irregular and less marked forms will probably be radicalized. In particular in natural SLA-situations, due to the absence of educational institutions and the pure orality of language learning situation, language intrinsic tendencies like morphological simplification that are associated with orality, might become more important for the linguistic output.

However, social restraints in terms of second order social indexicality will overrule L1-transfer and simplification, i.e. features that are perceived as highly salient for the original variety of the immi-

---

<sup>82</sup> For further information, see 3.2.1.2.

<sup>83</sup> For more details, see 3.2.1.4.

grant, will only be used if the immigrant wants to nonaccommodate to the “new” variety and emphasize that she is different from e.g. the listening Brazilians.

On the other side, L1-transfer or simplification will not take place in cases where immigrants want to accommodate to their Brazilian-speaking audience and perceive the Portuguese feature as highly salient. This means that accommodation influences, conditions and restricts L1-transfer and simplification processes.

### 4.3 Bolivian immigrants in São Paulo

---

So far I have presented an approach that should be valid for any language contact situation due to migration. With regard to the case investigated in this study, the Bolivian immigrants in São Paulo, some specific characteristics must be observed: First, the environment the immigrants move to has to be taken into account. Apart from the size of São Paulo and of the Bolivian community there – which certainly allow for huge variation with regard to the networks of the Bolivians – it was shown in Chapter 2 that São Paulo’s society is socially stratified and highly segregated. This also entails that the quality of the networks of the Bolivians is probably influenced by their social status. Additionally, the barriers against assimilation with regard to spatial mobility are rather strong as was shown before, in particular for those who have no car of their own. Thus, I hypothesize that the Bolivians will employ rather different accommodation strategies – from nonaccommodation to accommodation towards the social “lower-class” they have contact with or even towards the social “upper-class”.

Furthermore, migration theory shows that transnational social networks play a decisive role for migration decisions, e.g. social networks across state frontiers reduce the costs and risks of migration and therefore increase migration, making it at least partly independent from original structural or individual causes (cf. Haug and Sauer: 24). For the Bolivian immigration to Brazil, the importance of social networks reaching from São Paulo to Bolivia for the recruitment of new workers for e.g. sweatshops has already been mentioned. In fact, this means that these networks are transnational, containing persons from Bolivia as well as from Brazil – a fact little surprising if we think of the proximity of both countries and today’s facility of communication via cheap telephone calls or internet.

Linguistically, I hypothesize that the immigrants' transnational social networks also underline the translingual character of these networks, i.e. translingual<sup>84</sup> networks where more than one language is present and spoken: in this case the Castilian varieties and the Portuguese varieties. In contrast to multilingual networks that only signify the presence of various languages, because the speakers are multilingual, e.g. in this case of Aymara, Quechua or English, actual translingualism signifies that the limits between the languages and the speakers are fluent.

Because of the closeness of the two languages Castilian and Portuguese – doubtlessly the most decisive particularity of this language contact situation – the chances for positive L1-transfer, i.e. the (at least linguistically, if not socially) successful employment of L1 even in L2-contexts are very good. While it is an actual decision of the speaker in other language contact situations whether to accommodate or not, i.e. to speak the language of the dialogue partner or not, for the Bolivians there might be no either/or decision in many cases, because there is no clear cut distinction between L1 and L2, neither for the speaker nor for the hearer. Correspondingly, *congruent lexicalization* as described by Muysken (2008), meaning the use of L1 structures complemented by (socio-indexical) vocabulary and phonology stemming out of L2, might be one of the processes relevant for the speech of the Bolivian immigrants.

#### 4.4 Methodology

---

The theoretical approach presented so far has several implications for the choice of the right methodology, too. The fact that migration influences not only the migrant herself but also the receiving society must be reflected in the selected methodology: Therefore, thorough data about the Bolivian speakers must be gathered. The local categories like social networks must be investigated qualitatively because of the various factors like attitude, spatial mobility etc. that have to be considered for the immigrants. In the tradition of Milroy's (1980) and Bortoni-Ricardo's (1985) network studies, this points to an ethnographic approach where the investigator takes part in the studied community. In this way, the role of social and linguistic factors can be explored. It can be discovered whether linguistic differences between the speakers can be attributed only to L1-transfer and maybe universal tendencies like morphological simplification, or if accommodation and socio-indexical features also influence the speech of the Bolivians as suggested by my theoretical approach.

---

<sup>84</sup> With this term I do not refer to its meaning as used in translation sciences but rather to its meaning in literature studies. Kellman (2003: ix), for example, defines translingual authors as "those who write in more than one language or in a language other than their primary one".

The receiving society is particularly important to verify my concept of accommodation, and to gain data about which features are actually perceived as socially indexical. As was stated before, social indexicality of linguistic features is a concept based on social convention. Thus, it is scientifically more correct to ask the persons that establish and live by these conventions than name them as researcher coming from the outside.

As these conventions might differ due to the size and heterogeneity of São Paulo, I have to give a broad overview of the conventions dominant in São Paulo and include not only some individuals but a possibly representative sample of *Paulistanos* who live in various areas of the city, who belong to different socioeconomic classes and have different experiences with Bolivians. This aim clearly implies the employment of a quantitative analysis that will add the necessary breadth to this study.

Therefore, I employ *mixed methods research*<sup>85</sup> constituted of a qualitative study about the speech of Bolivians and a quantitative study about the perception and evaluation of their speech by *Paulistanos*. Compared with traditional quantitative and qualitative methodology mixed methods research still “[...] is an emerging field of study [...]” (Ivankova and Creswell 2009: 137), even in non-linguistic sciences. Only with the beginning of the twenty-first century, the publishing of studies using MMR increased significantly (cf. Ivankova and Kawamura 2010: 252 ff.). The development of MMR in the last years must be understood not only as an answer to the unsatisfying methodological paralysis caused by the so-called paradigm war (cf. Dörnyei 2009: 9 ff.), but first and foremost as a response to the complexity of social phenomena researchers have to face:

*[...] multiple diverse perspectives are important because they are required to explain the complexity of an increasingly pluralistic society (Greene and Caracelli 2003 cited by Teddlie and Tashakkori 2003: 22).*

In literature, many advantages of MMR are identified. Some underline the “unique possibilities for synergy and knowledge growth” (Padgett 2009: 104) and the opportunity to combine the depth of qualitative studies and the breadth of quantitative studies to get over the shortcomings of mono-method studies thus minimizing their weaknesses (cf. Onwuegbuzie and Johnson, 2006). Others, like Teddlie and Tashakkori (2003), emphasize that one of the areas where mixed methods are superior to single quantitative or qualitative approaches, is the “[...] opportunity for presenting a greater diversity of divergent views” (l.c.: 15).

---

<sup>85</sup> For a detailed description of the methodology, see Chapter 5.

Therefore, it seems only logical that MMR should be employed for sociolinguistics where at least the different perspectives of speaker and hearer have to be considered. Labov (1966) actually combined qualitative and quantitative methodology in his pioneer study about language variation in New York City, complementing his quantitative study by doing *exploratory interviews* and asking New Yorkers about their subjective reactions to and evaluations of the phonological variables. The same can be said for some of the Brazilian sociolinguistic studies mentioned in 3.1.4. However, they failed to call their methodology MMR; even worse, the line between the quantitative and qualitative parts is often blurred. In order to establish clarity about the methodological approach, I therefore propose naming the studies that employ both methodologies according to the terms suggested in Chapter 5.

Dörnyei (2009), one of the few authors who explicitly combines MMR and (applied) linguistics, emphasizes that in this domain researchers are often interested in both the exact nature and the distribution of a phenomenon:

*Mixed methods research is particularly appropriate for such multi-level analyses because it allows investigators to obtain data about both the individual and the broader social context (l.c.: 45).*

In this case, the central phenomenon is accommodation and the social indexicality of linguistic markers: A concept based, as was described earlier, on the interplay between various social groups and individuals, receiving society and immigrant. This means that in order to empirically show accommodation by socially indexical features, not only the action at the level of the individual speaker has to be considered, but also the subsequent (re)action at the level of society. Regarding these different elements, their timing and their nature, the necessity of using a sequential exploratory design becomes obvious.

#### 4.5 Summing up

---

Based on a sociological excursus that claims that assimilation processes towards social organization inside a differentiated society are valid to explain the behavior of migrants, my theoretical approach identifies accommodation as the decisive process with regard to language contact due to migration: It explains not only the choice of language maintenance resp. language learning as consequence of individual attitude and the social environment, but also the use of certain socio-indexical features. They are employed in order to accommodate to the variety that is perceived as socially successful in the social network the migrants live in. With regard to social factors influencing these networks, the speakers' attitude and the quantity and quality of Portuguese-speaking contacts have to be consid-

ered. Further linguistic processes that have to be taken into account in this context are L1-transfer and simplification. I claim that these processes are conditioned and influenced by accommodation as well, i.e. they underlie social constraints.

The social constraints of Bolivian immigrants in São Paulo are defined by factors like the size and heterogeneity as well as the strong social stratification of the population of the city and of the Bolivian community. Additionally, the close proximity of the countries allows for transnational networks. In accordance, the concept of a translingual network is developed where two languages are not only present, but actually merge, due to the fluent boundaries between the languages. In this case, the process of L1-transfer could also be described as congruent lexicalization.

Methodologically, the speech of the Bolivians themselves as well as data about their attitudes and their social networks, require qualitative ethnographic research. Since social indexicality is based on social convention, only a quantitative study that includes the social stratification of São Paulo's population can confirm whether the linguistic strategies employed by the Bolivians are actually successful. Generally, migration should be understood as a process where both, the migrants and the receiving society have to be included. Therefore, I suggest that mixed methods research should be employed for studying language contact scenarios caused by migration.

## 5. Methodology: a mixed methods approach

---

In this chapter I describe and explain the methodological approach employed in this study, following the former arguments in favor of mixed methods research (MMR) for the study of hearer-oriented sociolinguistics. I combine qualitative and quantitative data to obtain a complete picture of the dependencies between immigrant and receiving society. The methodological design is characterized as a sequential exploratory one, composed by a (qualitative) interview study with a (quantitative) follow-up questionnaire survey (QUAL → quan). The first part consists of the composition and analysis of a corpus of 28 sociolinguistic interviews with Bolivian immigrants living in São Paulo that were conducted in 2008, giving a detailed presentation of their social characteristics such as age, duration of stay and profession. Secondly, an online questionnaire based on the analysis of the qualitative data was developed and distributed between *Paulistanos* in 2010, resulting in the statistical evaluation of 100 responses. The design of the questionnaire as well as the challenges faced during the fieldwork are explained.

---

*Mixed methods research is the type of research in which a researcher [...] combines elements of qualitative and quantitative research approaches [...] for the broad purposes of breadth and depth of understanding and corroboration” (Johnson, Onwuegbuzie and Turner 2007: 123).*

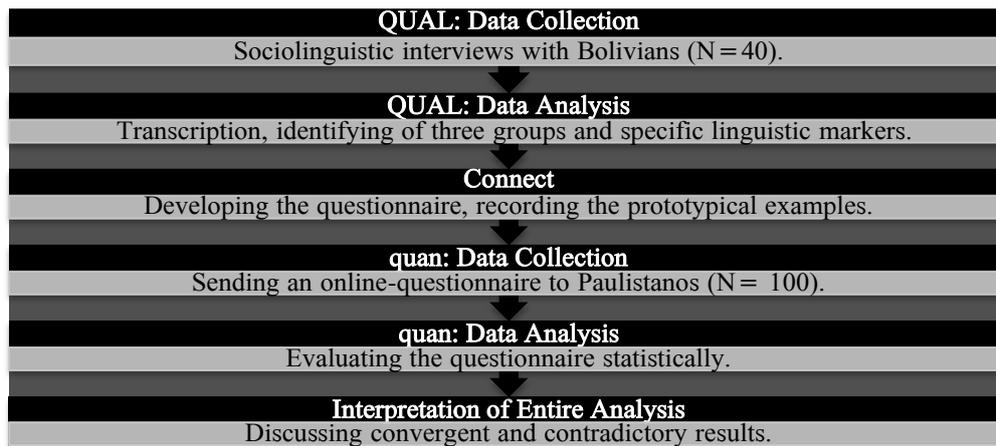
For the characterization of the MMR design employed in this study I refer to a “Basic Design Typology” (Nastasi, Hitchcock and Brown 2010: 313) based on the criteria timing, weighting and mixing (cf. for example Ivankova and Creswell 2009: 138, Creswell et al. 2003: 214 ff.). This means, the order of the qualitative and the quantitative part (indicated by “+” in the case of concurrent and “→” in the case of sequential designs) is considered, the priority given to them (indicated by large and small letters) and the modality of how to combine the two parts.

### 5.1 The sequential exploratory design (QUAL → quan)

---

A sequential exploratory design is a predominantly qualitative research design (indicated by QUAL) where the findings of the qualitative part conducted first, are followed by a quantitative part complementing the prior findings (indicated by quan). As the second part only arises from the first one, they are executed sequentially, one after the other. The mixing of the methods occurs at two points during the study: There has to be a connection between the qualitative findings and the quantitative research design, mostly consisting in the use of the qualitative themes and categories to develop the quantitative measurement instrument (cf. Creswell, 2008). Additionally, the results of both analyses are integrated for the concluding interpretation (cf. Figure 5.1).

Figure 5.1: Sequential Exploratory Design applied to this study (cf. Creswell et al. 2003: 140 ff.)



This design was chosen because the exploratory design is

*particularly appropriate when studying a topic which has been little explored [...]. [It] allows a researcher first to explore a topic by collecting qualitative data to help identify principal themes and possibly generate a theory (Ivankova and Creswell 2009: 140).*

As was pointed out before, the Portuguese of Bolivian immigrants has never been the subject of any study so far. Thus, it was necessary to explore the community and their speech by the collection of qualitative data. Only when the first part of the corpus was analyzed and three groups of Bolivian immigrants were specified according to their linguistic features and social networks, the theory of network-oriented accommodation could be developed. In São Paulo, the success of accommodation is obviously evaluated by the *Paulistanos*. Hence, to obtain a complete picture of the learning situation and to possibly identify the social indexicality of certain linguistic forms, a complementary study testing the perception and evaluation of the Bolivians' Portuguese was developed.

The mixing of the two parts – the qualitative data constituted by a corpus of 28 sociolinguistic interviews and the quantitative data constituted by a corpus of 100 completed questionnaires – occurs at two points: Firstly, the quantitative survey items have to be based on the qualitative data analysis; secondly, the quantitative results have to be compared with the initial qualitative findings.

In this case the first connection poses a particular challenge, because the perception of speech must be investigated by a survey including the evaluation of different speech examples. These examples, however, should differ only in the occurrences of prototypical linguistic markers, because every other difference could influence the participants of the survey and thus bias the results.

While comparing the results of both parts it is particularly important to present both convergent and contradictory results, because both add to the comprehension of the complex social phenomena in

question. Additionally, the numerical data of the quantitative analysis and the verbal data of the qualitative part must be combined in a comprehensible way to achieve one of MMR's main advantages:

*Words can be used to add meaning to numbers and numbers can be used to add precision to words (Dörnyei 2009: 45).*

## 5.2 The qualitative part of the study

---

Working with qualitative empirical data one realizes that, in spite of all the methodological and strategic planning before the actual data collection, in reality, the gathering of the data is conditioned or even impeded by factors that can rarely be anticipated. In exploratory studies, the risk of being confronted with unforeseen obstacles is even higher, because by definition, they are employed for researching topics that are little investigated.

In the case of Bolivian immigrants, the topic had only been studied by a few researchers from an anthropological point of view and some journalists reporting about the *trabalho escravo*, most of whom had either observed the community while working in an NGO dealing with Bolivian immigrants (cf. Silva 2005 and 2008) or interviewed just one or very few Bolivians (Barros 2000, Cymbalista and Xavier 2007, Soraia Vilela 2010, Sydow, 2003 and Varella 2004). Consequently, to obtain the sociolinguistic data I was looking for, I had to get more than “a snapshot view of the phenomenon” (Rallis and Rossmann, 2009: 269) and show prolonged engagement with this group. This deduction was undermined by the fact, that even during the first conversations with the gatekeepers (cf. Heigham and Sakui 2009: 97) all of them agreed that Bolivians are a closed people who do not like talking to strangers. Whether due to their illegality and fear of being discovered and deported or due to cultural or linguistic misunderstandings, later on, this valuation was confirmed by most of the interview partners and by my own experience. Therefore, I chose an ethnographic approach, studying “people’s behavior in naturally occurring, ongoing settings, with a focus on the cultural interpretation of behavior” (Heigham and Sakui 2009: 93). As methods to collect the data I used observation<sup>86</sup> to be able to understand the structure and the dynamics of the Bolivian community and the daily routine of the individual Bolivian immigrant I talked to.

For the corpus, the part of the study being transcribed and linguistically analyzed, I used sociolinguistic semi-structured interviews. My role as researcher could be defined as partially participating,

---

<sup>86</sup> Observation as a qualitative method is defined as “the conscious noticing and detailed examination of participants’ behavior in a naturalistic setting” (Cowie 2009:166).

because I worked as a voluntary Portuguese teacher for a Bolivian class, became friends with my interview partners and met the Bolivian salespeople as a regular client at the Bolivian market, before and while I was conducting the data collection.

### 5.2.1 Ethics

---

As a partially participating observer and researcher during the data collection, hence an “outsider [...] wearing insider’s clothes” (Tedlock 2003: 98), the ethical principles underlying the research must be considered, particularly when interacting with the participants. The most sensitive aspects in my case can be categorized as “relationships” (Dörnyei 2009: 65) on the one hand and “[t]he amount of shared information” (l.c.) on the other hand.

As it is part of the qualitative study to attempt to gain insight into the participants’ life circumstances and their social compoment, I inevitably intrude into their private spheres, which can create a false impression of intimacy and closeness and eventually lead to a strongly unequal and dependent relationship. In this case, the obvious economic distance between a German researcher and an illegal Bolivian seamstress even intensifies this inequality, a fact that could easily lead to the abuse of the informant on my part. Thus, special attention has to be paid to clarify the role of the researcher and also emphasize the transitional character of my involvement. According to Dörnyei (2009: 65), this last point is also important, because otherwise the termination of the project could result in the “concrete dilemma about how to *end* a research project without leaving the participants feeling that they were merely used” (emph. in original).

Furthermore, the ethical principle of sharing all possible information about the project with the participant can be opposed to the finality of the data collection. For example, telling participants everything about the research aim could bias their responses or even make them refrain from participating altogether. Even if I do not believe that this would lead to strong monitoring and hence biasing of the linguistic data, the language learners could easily refuse to talk in Portuguese at all, if they knew their language was of interest, because they already consider themselves as speaking poorly or not well enough. Thus, I told my respondents only after the interview that I was not only interested in *what* they said about their lives, their mobility in the city and their social networks, but also *how* they said it. At this point, the participants had the opportunity to remove their formerly given consent to the study once more, but none of the respondents made use of it. According to Rallis and Rossmann (2009: 276) this consent was built on four basic ethical principles: Participants are as fully

informed as possible about the study's purpose and audience; they understand what their agreement to participate entails; they give their consent willingly and understand that they may withdraw from the study at any time without prejudice. Despite this provision, the incomplete information before the interview and also my work as a voluntary teacher could still be interpreted as deceiving and masking my full identity (cf. Rallis and Rossmann 2009: 277), but they were necessary to conduct the study.

On the whole, I followed the principle as formulated by Rallis and Rossmann (l.c.):

*The key is that the researcher must take every possible precaution to ensure that no harm will come to the participants as a result of participation.*

This means also that the existence of illegal sweatshops was not reported to the police, that I tried to do the interviews outside the sweatshops in order to avoid placing participants at the risk of losing their jobs<sup>87</sup>, that questions about the illegal status of many participants were usually not recorded as part of the interview and that none of the informants is identifiable by name.

## 5.2.2 Participants

---

When investigating a very little explored community in a foreign country without personal contacts within this community, the selection of the informants is particularly difficult to determine before starting with the data collection. Thus, two steps were necessary to compose the corpus: Firstly, I tried to establish contact and talk to as many informants as possible to obtain a clearer picture of the whole situation of Bolivian immigrants in São Paulo, before I selected the sample for the analysis in a second step.

### 5.2.2.1 Making contact

---

According to Heigham and Sakui (2009) the first steps of an ethnographic researcher are crucial for the type and place of your data collection. "Once you have decided these, you select your research site(s) and arrange entry" (l.c.: 96). In this situation, though, even the selection of the research sites posed a challenge because of the poor data basis I had to rely on. Thus, I looked for churches or other cultural institutions with Bolivian representation, because priests often come into contact with marginalized people and often function as community gatekeepers due to the authority of their pro-

---

<sup>87</sup> As researcher I once entered one of the illegal sweat shops with one of the workers to talk to his colleagues. I explained the aims of the study to the  *dono*  confirming that I was not interested in any legal status. Although he allowed me to do the research, my contact person was threatened with dismissal after I had left.

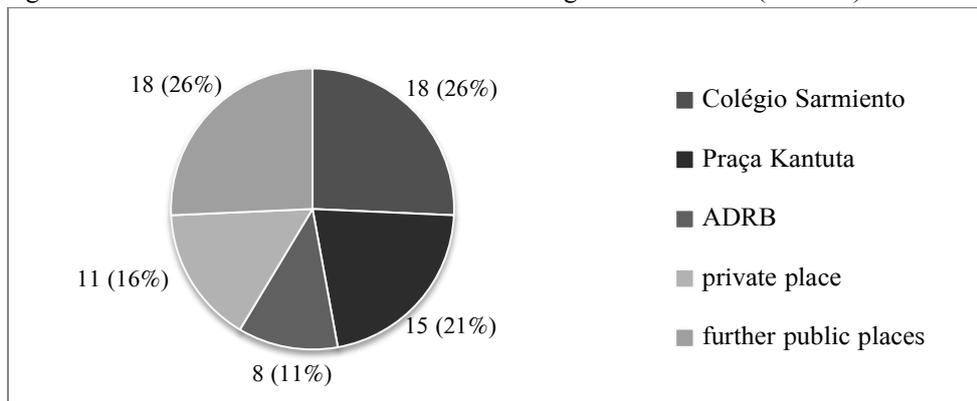
fession. Eventually, three gathering places of Bolivians were identified where it was expected to be possible to make contact with gatekeepers:

- The *Centro de Pastoral dos Migrantes Nossa Senhora da Paz*
- The organization *ADRB*
- Sunday's Bolivian market at *Praça Kantuta*.

Having identified these places I asked the leaders of each organization for help, so they could either introduce me directly or to allow me to refer to them as recommendation when addressing other members of their organization. At the *Centro de Pastoral dos Migrantes Nossa Senhora da Paz* I learned about the *Colégio Sarmiento*, an NGO where Brazilian students offer Portuguese lessons to Bolivians every Sunday. As there were few teachers available because of university vacations at the time, I offered to help and worked several weeks as a voluntary Portuguese teacher.

This means that the main research sites for observation and interviews were the *Colégio Sarmiento*, the market at *Praça Kantuta* and the *ADRB*. Apart from these places, where the majority of the interviews was conducted (18 at *Colégio Sarmiento*, 15 at *Praça Kantuta* and eight at the *ADRB*), 18 further interviews were held in public places like restaurants, a bus stop or near a football field. It was almost impossible to meet the informants in their homes, because most of them live in very poor conditions, often at their place of work sharing with many other colleagues. There, they feel too observed and uncomfortable to speak freely. Additionally, as the *donas* fear that their illegal business will be discovered and raided, they often prohibit strangers at the sewing shop. Thus, I only conducted 16% of all interviews at a private place, as Figure 5.2 shows:

Figure 5.2: All interviews before selection according to research site (N = 70)



#### 5.2.2.2 The sampling strategy

It has already been pointed out that both observation and sociolinguistic interviews were methods to collect the data. Only by the close observation of the Bolivians at the research sites could I compre-

hend their social comportment and their life circumstances – a highly important prerequisite to accomplish the interviews. The sociolinguistic analysis, however, is based on the interviews of the selected sample alone.

Qualitative studies are not representative of the whole population. Hence, the selection of the sample that will be analyzed sociolinguistically is not necessarily dependent on characteristics like age, gender or social class. The aim of my research, however, was clearly stated as describing the speech of Bolivian immigrants (without restriction to single subgroups) and analyzing the influence of different social factors. Thus, the intention was to include as many important subgroups of different socioeconomic classes and language proficiency as possible.

Therefore, I first tried to get as many respondents as possible by chain sampling (cf. Dörnyei 2009: 129). This means, every participant was asked to recruit further participants from his network. According to Dörnyei (l.c.) this strategy is “ideal in situations where the experience in question is rare.” After having iteratively collected and analyzed some data, I decided to employ the strategy of maximum variation sampling<sup>88</sup> (cf. l.c.: 128) to be able to explore the sociolinguistic variation between the respondents, particularly in reference to their social networks. As seen before, these strategies resulted in a first sample of 70 interviews (N = 70) until saturation regarding maximum variation was reached (cf. Figure 5.2). Naturally, this sample was not only too broad for a qualitative analysis; there were also many cases too similar to offer new information or the sound quality of the interviews was not good enough for the phonologic analysis.<sup>89</sup>

### 5.2.2.3 Selecting the informants

---

For the eventual corpus of the linguistic analysis, the sample was cut down to a size of 28 participants, considering mainly sociolinguistic features (cf. Chapter 6), but also criteria like profession (cf. Figure 5.3), age (cf. Figure 5.8), gender (cf. Figure 5.10), and how long the immigrants have lived in São Paulo (cf. Figure 5.5) to check for the influence of these classic sociolinguistic variables. The still rather large size of the sample is explained by the linguistic phenomena I want to investigate. To

---

<sup>88</sup>“The researcher selects cases with markedly different forms of experience (for example, L2 learners from all developmental levels” (Dörnyei 2009: 128).

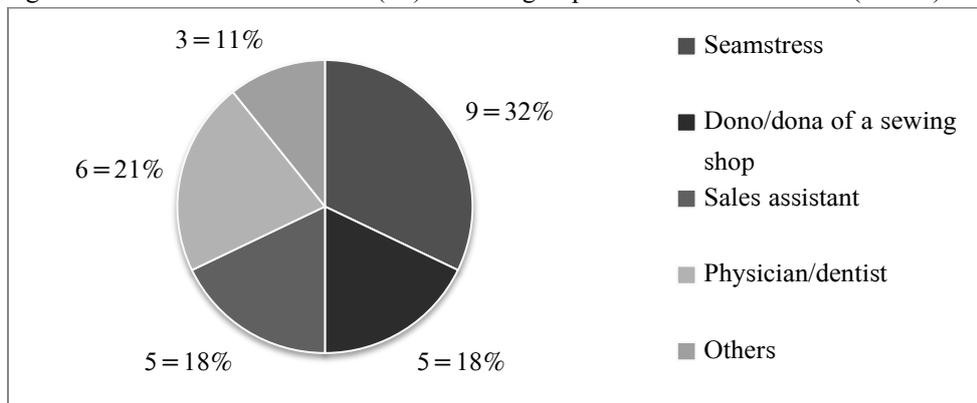
<sup>89</sup> As Figure 5.2 shows, only 27% of all interviews (at the ADRB and at private places) were held somewhere where the background noise was at least partly controllable. Thus, many interviews conducted at the other places had to be excluded from analysis, because the background noises would have impeded the phonologic analysis.

include various examples of rarer morphosyntactic features the corpus has to contain more speakers (cf. e.g. Machline Oliveira e Silva 1992: 108).

Though the numbers of the participants are not representative of the population size, the main dynamics and characteristics of the Bolivian community are reflected in the corpus, for example the strong immigration into the textile industry described in Chapter 3. Figure 5.3 clearly shows that 14 participants, i.e. 50% of the corpus, are directly involved with this business, either as seamstress or as *dona* of their own sewing shop. Additionally, two of the five sales assistants started working at a sewing shop or even continue to work there during the week; selling goods at Sunday’s Bolivian market at *Praça Kantuta* is just an auxiliary income. This distribution underscores the dominant role of the textile industry for the community of Bolivian immigrants, particularly in light of the mass migration of the last decades.

The second biggest professional group is constituted by the doctors and dentists. These six plus also an engineer and a pharmacist (subsumed in the category “others”) represent the academic immigration<sup>90</sup> from Bolivia to Brazil that already existed in the 1970s.

Figure 5.3: Number of interviews (NI) according to profession after selection (N=28)



The dominance of these professions is strongly conditioned by the selection of the research sites, as Figure 5.4 shows. Obviously, at *Praça Kantuta* I encounter principally sales assistants and at the ADRB Bolivian doctors and dentists. As the *Colégio Sarmiento* offers additional professional qualification for Bolivians, which is particularly interesting for people who suffer from bad working conditions, here seamstresses can be met who are trying to improve their lives.

<sup>90</sup> This does not mean that none of the other informants working in the textile industry are academics. On the contrary, there are various *donas* as well as seamstresses who went to university before coming to Brazil, but they did not finish their studies and do not work in their studied profession as the doctors, dentists and engineers do. Apart from that, all informants attended school for at least five years, most of them eight or more.

Figure 5.4: NI according to profession and research site (N = 28)

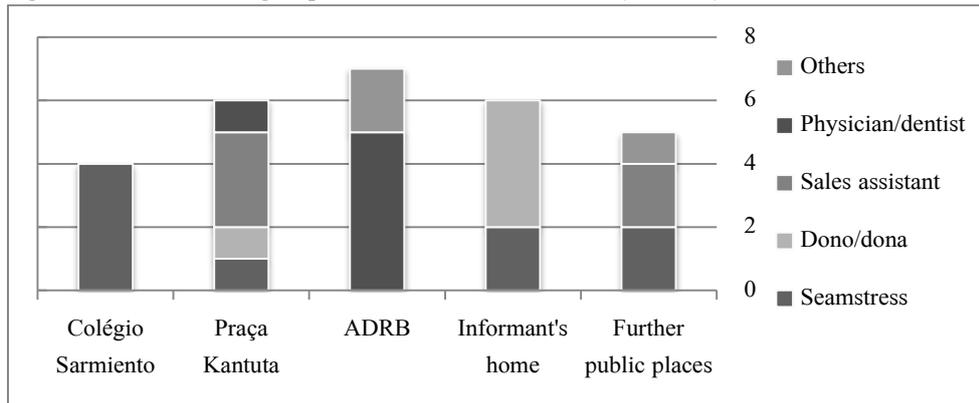
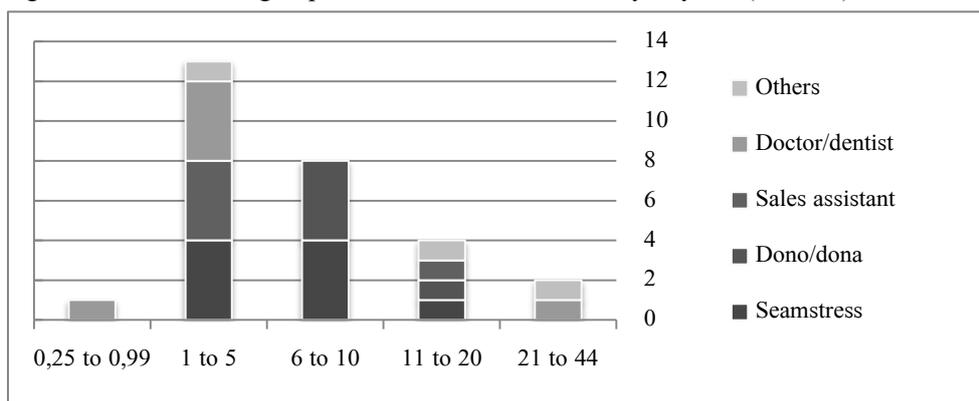


Figure 5.3 shows that only at *Praça Kantuta* and in other public places the distribution is more heterogeneous. The market and for example Bolivian restaurants at *Rua Coimbra* are indeed places where all members of the Bolivian community, whether seamstresses or academics, meet to enjoy the food of their home country. The other places display a strong separation according to profession and therefore of socioeconomic class. For example, at the *Colégio Sarmiento* I only met seamstresses, while at the ADRB, I only encountered doctors, dentists, the engineer administrating the association and the receptionist.

Interestingly, only the *donas* and two seamstresses offered their homes as the interview place, even if I had to stay in the yard in some cases. This is obviously due to the fact that many seamstresses are ashamed or even forbidden to bring strangers to their place of residence, because they live at their working places, the illegal sewing shops. The two seamstresses who agreed to do an interview in private are the only ones who already had spent more than nine years in São Paulo; they live in their own houses independent from their *donas*. The other Bolivians working in the textile industry who have been in Brazil for such a long time have all become *donas* of their own sweatshops, as columns three and four of Figure 5.5 indicate. My data also underscores the finding that only after some years of working as a seamstress one can become a *dona*, because the first two columns include every other professional group but them.

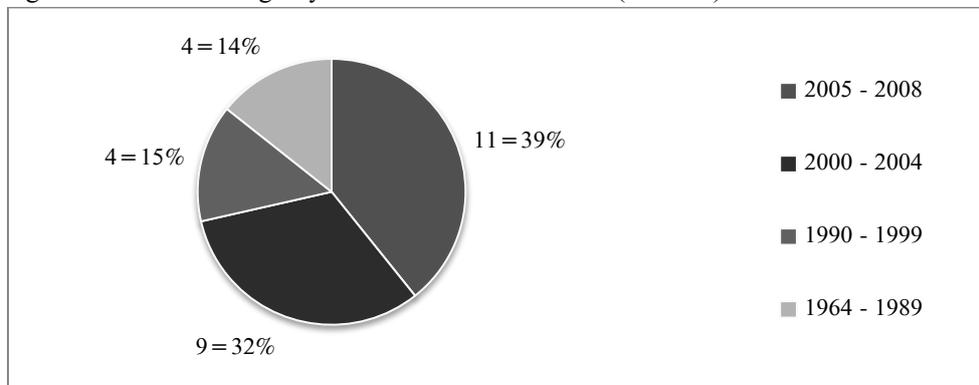
Figure 5.5: NI according to profession and duration of stay in years (N = 28)



The fact that the number of informants diminishes strongly after ten years and the only informants having spent more than 20 years in São Paulo are academics, reflects the increasing immigration of Bolivians, starting in the 1960s and 70s with the immigration of academics and increasing to the economically driven mass migration in the twenty-first century.<sup>91</sup> The same aspect is illuminated by Figure 5.6: Only 29% of the informants had arrived in São Paulo before 2000, i.e. the majority of the Bolivians had stayed in the city less than ten years at the time of the interview in 2008.

Following the sampling strategy of maximum variation, extreme cases were included such as that of a physician who had come to São Paulo only three months before the interview and an engineer who had lived there since 1964.

Figure 5.6: NI according to year of arrival in São Paulo (N = 28)



This pattern of maximum variation also becomes obvious when comparing the age of the informants and their age at time of the arrival in São Paulo (cf. Figure 5.7 and 5.8):

Figure 5.7: NI according to age at time of the arrival in São Paulo (N = 28).

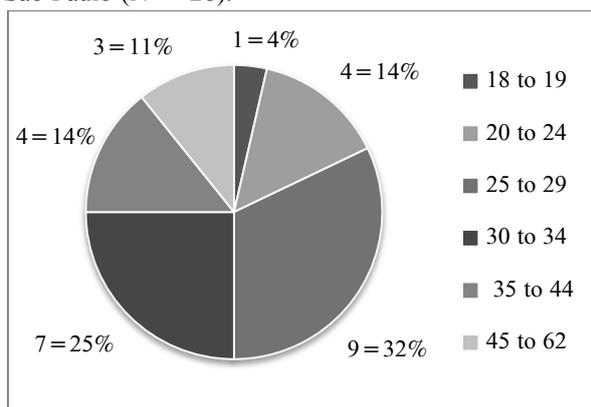
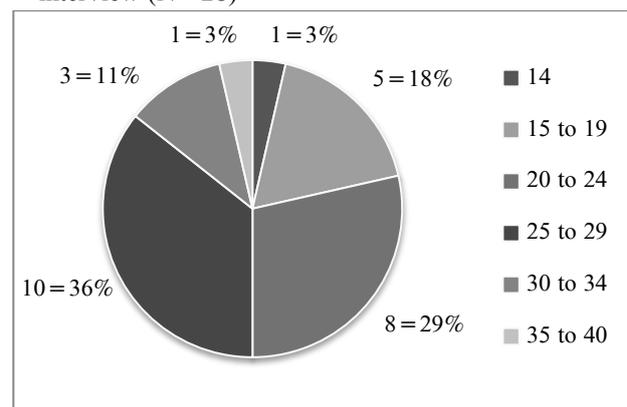


Figure 5.8: NI according to age at time of the interview (N = 28)



<sup>91</sup> The presence of many academics in the first two columns originates from my aim to compare the Portuguese of diverse social and professional groups, independent of their duration of stay.

There is one informant, for instance, who had come to São Paulo at the age of only 14 years; another one was already 40 years old at the time of her arrival to Brazil. As said previously, most of the Bolivians get to São Paulo between the ages of approximately 20 and 35; this was the case for more than 70% of the whole sample. Correspondingly, at the time of interview, 61% of the informants were between 20 and 30 years old, but I also included informants older than 60 years of age as well as one of only eighteen years of age. Naturally, these atypical cases may not be overestimated, but they can underscore linguistic commonalities despite the obvious social and professional variety within the Bolivian community.

To complete this overview about the informants' characteristics, Figure 5.9 and Figure 5.10 show the number of informants according to place of origin in Bolivia and gender. Regarding the place, there are different aspects to be considered: Almost all informants come from an urban area, only two informants stem from the countryside;<sup>92</sup> following the linguistic-geographic division by Coello Vila (1996), the corpus is dominated by informants from one zone,

*región andina centro y sudoccidental. Comprende los departamentos de La Paz, Oruro, Cochabamba, Potosí y Chuquisaca” (l.c.: 27),*

whose variety is characterized as

*mercado bilingüismo castellano-aymara o castellano-quechua” (l.c.: 28).*

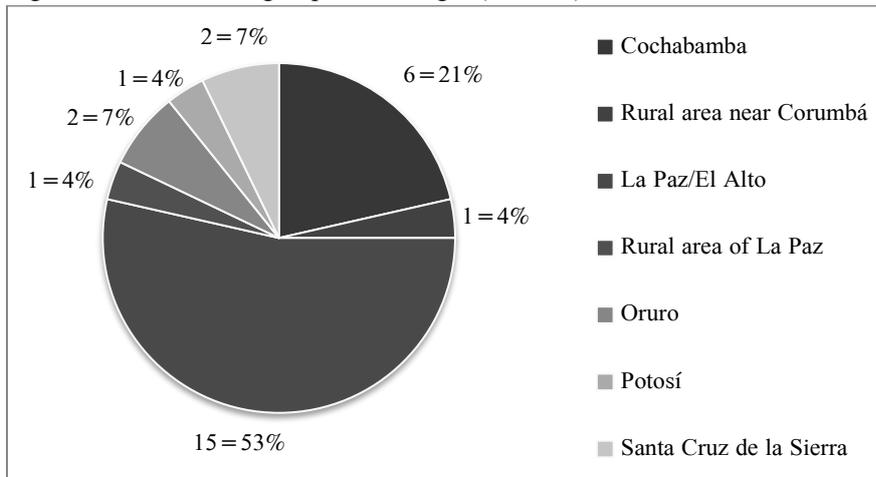
Only three speakers come from the “región de los Llanos del norte y del oriente” (l.c.) where not only the influence of *tupí-guaraní*-languages can be observed but also of Brazilian Portuguese. This means that for the analysis, linguistic features must be included that are discussed for the varieties of the Andean region where the indigenous languages like Aymara (La Paz, Oruro) and Quechua (Cochabamba, Potosí) influence the Castilian spoken there.<sup>93</sup> The cases from Santa Cruz and Corumbá as well as from rural La Paz, however, serve as further extreme cases to compare the data in regard of the influence of the place of origin.

---

<sup>92</sup> This finding underscores the observation of Silva: “ A população boliviana é majoritariamente jovem e urbana [...]. Os grandes centros aglutinadores de população são as cidades de La Paz e El Alto, no altiplano, Santa Cruz de la Sierra, no oriente boliviano, e Cochabamba, na região central. São desses centros urbanos que saem grande parte dos que emigram para o Brasil [...]”(Silva, 2005: 7).

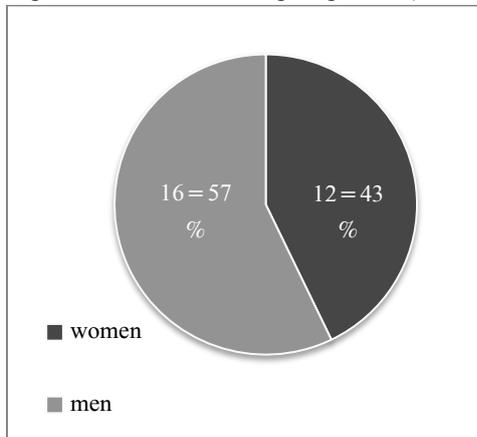
<sup>93</sup> As Quechua and Aymara display the same grammar structures, the varieties stemming from their influence on Castilian are mostly subsumed as *Español Andino* without distinguishing the differences of the indigenous languages. As we have no means to falsify or verify this approach, I follow the distinction cited above relying on the corresponding literature. A clearer distinction of the varieties spoken in Bolivia, however, clearly constitutes a further desideratum for research.

Figure 5.9: NI according to place of origin (N = 28)



As mentioned before, these numbers are not representative, but the proportions reflect the observation regarding the dominance of people coming from the *Altiplano*. The distribution according to gender shows 57% men and only 43% women, confirming my own observation<sup>94</sup> and the finding of Silva (2005: 16) that most of the immigrants are young men even if the presence of women has increased in the last years.

Figure 5.10: NI according to gender (N = 28)



In summary, the composition of the sample is not representative regarding gender, profession, research site or age. However, by combining chain and maximum variation sampling strategies, I achieved a broad and deep sample that shows not only the main dynamics and characteristics of the

<sup>94</sup> Obviously, this could be due to the fact that many women do not leave their houses out of fear or too much work, hence they are not seen on the streets or at the institutions where the observations were made. However, as it is a typically female obligation to buy Bolivian ingredients to prepare food for the family at least on Sundays, an equal proportion of both sexes at least at *Praça Kantuta's* Bolivian market could be expected, if there were as many women as men in São Paulo.

Bolivian community, but also the variety between the participants by including cases of maximum variation. Almost paradoxically, these extreme cases will be useful to detect common features and verify my findings:

*If a pattern holds across the sampled diversity, we can assume that it is reasonably stable (Dörnyei 2009: 128).*

### 5.2.3 Sources of data

---

The data sources for this study are my observations and communications with the Bolivian immigrants at the places described earlier. As the Portuguese spoken by them also constitutes the research object of this study, I have to make the difference between sources usable for sociolinguistic analysis of 28 speakers and further data that are necessary for obtaining the knowledge about the social background of the informants. Only in this way can the linguistic data be integrated into the big picture. Methodologically, this means that the data source for the actual sample is a homogeneous composition of semi-structured sociolinguistic interviews, while the additional data include a heterogeneous mixture of different kinds of data that will be described further on.

For the recordings, I used a Marantz PMD 620 recorder, because the small size of the device offers major advantages. The device can be carried around and put everywhere, i.e. the interview can happen outside, for example at a bus stop or at a food stand at the market place as described earlier. The interviewed persons also easily forget that they are recorded. Furthermore, as many Bolivians live in areas of São Paulo with high criminality rates, it is important that the recorder can easily be hidden to avoid robbery and theft.<sup>95</sup>

The different data sources require different procedures of data collection that will be presented further on.

#### 5.2.3.1 Observation

---

As a partially participant observer, I made an effort to become part of the community of practice at every research site: At the *Colégio Sarmiento* I became one of the voluntary teachers, at the *Praça Kantuta* a regular customer of Bolivian goods and friend of the market's leader and at the ADRB an interested supporter of their case, regularly attending their events and spending time at the office. This way, I gained the trust of the Bolivians and obtained the information I needed to conduct the

---

<sup>95</sup> Because of these circumstances, instruments like videotaping or bigger microphones that were successfully employed in sociolinguistic studies to obtain better sound quality, could not be employed in this case.

interviews. Furthermore, also being foreign in Brazil, knowing about difficulties with the *Polícia Federal* and having learned Portuguese in Brazil helped to express solidarity with the Bolivians. It would have been impossible to participate more, because the life of a German researcher and a Bolivian seamstress differ too much.

Nevertheless, I had the opportunity to collect data from various sources: I conducted structured, semi-structured and unstructured interviews in Portuguese and also in Spanish<sup>96</sup>, talking to immigrants but also to their Brazilian wives or husbands, to their children or even to the Bolivian consul. These talks were either recorded or accompanied by taking notes. Additionally, I recorded the Portuguese lessons I gave at the *Colégio Sarmiento*, collected Bolivian newspapers in São Paulo, took pictures at the different research sites and maintained personal email communication with Bolivian immigrants in São Paulo. The big picture of Bolivian immigrants in São Paulo was completed with my observations in Bolivia itself, where I met family members of São Paulo's immigrants and took notes of the conversations I had there about emigration.

This diversity of sources enabled me to triangulate my experiences and the perceived literature with the linguistic data I obtained from the interviews.

#### 5.2.3.2 The sociolinguistic interview

---

A particularity of sociolinguistic interviews is certainly the fact that the point of interest lies not only in what content is communicated, but also how it is done. Therefore, a highly structured interview that allows little participation of the respondent and “[...] little room for variation and spontaneity” (Dörnyei 2009: 135) cannot be used for a broad research design like mine based on the purpose of giving an overview of a variety of linguistic features. On the other hand, in the pilot interviews I saw that many informants did not feel comfortable with an unstructured interview design, because – probably influenced by the mass media – they are used to a design where questions are asked and answered. Thus, I conducted the interviews in a way, “which offers a compromise between the two extremes” (Dörnyei 2009: 136): the semi-structured interview. Here, the researcher prepares to a large extent what topics need to be covered and how the questions are asked, allowing hence a certain degree of comparability (cf. Richards 2009: 185), but at the same time it gives the respondent the possibility to steer the talk into new directions.

---

<sup>96</sup> For the actual corpus, only interviews conducted in Portuguese were chosen.

In my case the topics included some biographic data like their age and the duration of their stay, life in Bolivia, particularly their language competences and their family status, the story of their emigration, for example their reasons for leaving Bolivia and the route they took, their living conditions, social network and spatial mobility in São Paulo (sometimes assisted by a map of the city) and the attitude toward São Paulo, the inhabitants of the city and the Portuguese language.

The questions were formulated to stimulate personal views and opinions and mainly narrative passages where the interviewed person could speak for a longer time without the interference of the interviewer.<sup>97</sup> The speaker should always feel free to talk about additional topics if she wanted to, as is required by the semi-structured method (cf. Richards 2009: 186). This freedom resulted in the fact that the interviews differ considerably regarding their length: the shortest six minutes, the longest 40 minutes. Like the differences within the data shown before, these are cases of maximum variation that have to be considered for the analysis.

The interviews were held as face-to-face interviews, mostly with only two participants – the researcher and the respondent – but some of the interviews included more respondents, because the participants only agreed to the interview if done together with their partners or friends. In this situation I tried to do the interviews separately, talking to only one participant at the time of the interview. However, I did not try to stop the interaction between the respondents to maintain the natural flow of the communication.

Regarding the interviewer, traditional literature holds it as a “key principle that the interviewer should try to be neutral, without imposing any personal bias” (Dörnyei 2009: 141), but Fontana and Frey (2009) convincingly show that the researcher is no neutral tool, but a person “carrying unavoidable conscious and unconscious motives, desires, feelings, and biases [...]” (l.c.: 696). Additionally, the power differential between the interlocutors places the interviewer in a hierarchically stronger position than the respondent. Even in a semi-structured interview, where the respondent – in comparison to the structured interview – has more opportunities to choose the topics and speak without restriction, the power differences are also present, since the reduction of explicit control by the researcher does not automatically signify the empowerment of the respondent (cf. Briggs 2005: 1056). For these reasons, the interviewer cannot be neutral:

---

<sup>97</sup> Traditional sociolinguistics claims that narrative passages are a *gold mine*, because the narrator completely forgets about the linguistic forms when counting personal experiences. Thus, narratives are more natural than other discourse types (cf. e.g. Tarallo 2001). As I will expand on later, I do not believe in the *naturalness of interviews*, but narratives are still useful for getting a longer and undisturbed passage for analysis.

*Interviewers are [...] active participants in an interaction with respondents, and interviews are seen as negotiated accomplishments of both interviewers and respondents that are shaped by the contexts and situations in which they take place (Fontana and Frey 2009: 696)*

Thus, in the interviews I followed an empathetic approach (cf. l.c.), signaling agreement with the respondent and revealing personal feelings and private situations, even if I did not share the opinion of my informant outside the interview.

The influence of the researcher as one of the interlocutors has to be considered at the linguistic level, too. I chose Portuguese as the language of the interview, although many informants would have preferred to speak Spanish.<sup>98</sup> Of course, sociolinguistic literature described the effect of the context and the interviewer as early as the 1960s and 1970s when the concept of the vernacular and the observer's paradox was developed<sup>99</sup>, but only Bell's audience design (1984) emphasized that this effect cannot be reduced but will lead to the stylistic variation of the interlocutors. In accordance, Pappas (2008) points out: "It is perhaps overly optimistic to assume that the data obtained during a sociolinguistic interview could ever be truly representative of the vernacular, or, to be more specific, of that variety that participants use when they are talking with their closest childhood friend without a microphone and a recorder present" (Pappas 2008: 498).

In this study the stylistic shift becomes obvious when comparing the utterances of the interviewer in the different interviews. Depending on my social role in the different CoP and the characteristics of the respondent, the same questions are formulated differently and the discourse signals differ slightly, i.e. for example informants older than the interviewer are treated differently and interviews with men differ from interviews with women.<sup>100</sup>

This shows clearly, that particularly in cases like mine, with strong power asymmetry between the interlocutors, a German researcher and a Bolivian economic refugee, the linguistic influence of the interviewer on the speech of the informants has to be considered when interpreting the data.

---

<sup>98</sup> Mostly, the language was not chosen explicitly, but I started asking questions in Portuguese and the respondents also answered in Portuguese. If asked explicitly, I told them it would be better if they spoke Portuguese. Only one of the informants explicitly decided to do the interview only in Spanish, because as he perceived the situation as an official discourse, he wanted to speak his first language. After the recorder was turned off, he immediately switched back to fluent Portuguese.

<sup>99</sup> For more details, see 2.1.

<sup>100</sup> Unfortunately, these interesting linguistic facts cannot be investigated more narrowly because of the focus of this study; it seems though, as if this point has not been studied enough. Particularly in the case of huge variationist studies where data from various fieldworkers are gathered together and studied without differentiation according to the researchers' age, gender or ethnicity, the effects could be strongly biasing for "[t]he effects of the race of the interviewer on sociolinguistic fieldwork" (cf. Cukor-Avila and Bailey 2001).

### 5.2.3 Transcription and analysis

---

As was shown before, every step of data collection, beginning with the selection of the research sites and continuing until the structure of the interview, is influenced and even biased by the researcher, because it is impossible to neutralize her personality or her research interest. This influence has also to be considered during the transcription and analyzing process. As Rodrigues and Ferreira Netto (2000) point out: “Aliás, transcrição nenhuma é neutra!” (l.c.: 173), because, by transferring the oral communication into another medium, the transcribing person is subordinated not only by her own language ideology determining what she hears and understands, but also by the structural limitations and orthographic conventions of written language. In other words:

*[A]ll transcription is representation, and there is no natural or objective way in which talk can be written (Roberts 1997: 168).*

This means that the transcription of oral communication has to be seen as an interpretive retelling of the original conversation (cf. Dörnyei 2009: 247) and not as a pure reproduction in another medium. Following the distinction between conceptual and medial speech and writing as formulated by Koch and Oesterreicher (1990: 5 ff.), the transcription is an oral concept transferred into the written medium. This transference, however, changes also the content because of the interpretive work of the transcriber.

The analysis of qualitative research is also fundamentally interpretive, because the researcher interprets the data according to her opinions and research interest, even if this happens subconsciously, because she cannot neutralize her own background: “[T]he researcher filters the data through a personal lens that is situated in a specific sociopolitical and historical moment” (Creswell 2003: 182).

Despite this unquestionable influence of the researcher, the transcription and analysis procedures should be as transparent as possible. The first step to achieve this goal is to use a transcription program, in this case *Exmaralda*, which offers the possibility to get not only the written word, but also the audio trail (cf. annex I). This is particularly important for the comprehension of the phonological analysis, but also the analysis of morphosyntax and vocabulary become more plausible if supported by the original sound. As the transcription followed orthographic norms to ensure its readability, features of the interlanguage could be lost.

Although transcription protocols of Brazilian linguists recommend and successfully used a different system that implements linguistic features of Brazilian Portuguese directly into the transcription (cf. e.g. Bortoni-Ricardo 1985 and Rodrigues and Ferreira Netto 2000), I decided for the orthographic

option following mostly the transcription norms of the project NURC (cf. Preti 2003: 13 ff.), because I already had to deal with the transcription of code-mixing between Spanish and Portuguese. The close relatedness of the two languages makes it often impossible to decide whether the word is Portuguese or Spanish, either Portuguese with Spanish pronunciation or Spanish influenced by Portuguese, etc. This difficulty will be further considered during analysis and discussion. Thus, I only voted for Spanish orthography if the word or the morpheme was without doubt Spanish or if the word was ambiguous but situated in a purely Spanish context, i.e. words before and after were undoubtedly Spanish. When the transcription into Spanish orthography led to misunderstandings, because the same word existed in Portuguese with a different meaning, I labeled it to be recognizable. Thus, for example, Spanish *no* (no) was transcribed as *nó*, so it could not be confused with Portuguese *no* (*em* (in) + *o* (art.)). The remains were normatively transcribed into Portuguese, paying little attention to phonological discrepancies, because these can either be heard directly when working with the corpus or the phonetic transcription will be given and discussed during the phonological chapter. Additionally, I used punctuation marks to facilitate the reading, but tried to reflect the discourse structure by also indicating pauses longer than 0.1 seconds.<sup>101</sup>

It is obvious that these transcription norms lead to a biased representation reflecting the research interest of the Portuguese spoken by Bolivian immigrants. For the analysis every linguistic example was hence double-checked with the original oral data and not only with the written representation.

During and after the transcription I analyzed the single interviews according to their linguistic and non-linguistic contents. The important elements were coded and the single utterances added to the characterization of the single informant.<sup>102</sup> In an iterative process of noting frequent and extraordinary linguistic features that either had already been described for similar situations of language contact or had caught my attention during the transcription, 15 subcategories and the additional category of code-mixing were selected.

### 5.3 Connecting the two parts

---

In a sequential exploratory design the integration of the qualitative and quantitative types of data occurs at two stages in the research process: initially in connecting the qualitative and quantitative research parts, and then in interpreting the results by comparing and discussing both study parts. As

---

<sup>101</sup> For further transcription norms of this study see the transcription protocol (cf. annex II).

<sup>102</sup> The computer program used was *citavi*. Although this program was originally developed to organize bibliographies, I found it extremely convenient for the purpose of corpus organization.

the purpose of this strategy is “to use quantitative data and results **to assist** in the interpretation of the qualitative findings” (Creswell 2003: 215, my emphasis, sn), a survey was created, where the questions of the questionnaire were deduced from the first part. As my focus was on the language regard and evaluation of the speech of representatives of the different network types, the questionnaire is based on three audio examples of Bolivian immigrants.

As the speakers should be evaluated alone on their linguistic characteristics and not on the content of their speech, it was impossible to take examples from the interviews already conducted, because these would offer too much information about the speaker only by listening to **what** she was saying. This study, however, should measure the perception of **how** she was talking. Thus, I had to record new examples with Bolivian immigrants that would give no explicit information about the speaker. Additionally, the speaker should employ more or less the same linguistic features and social characteristics that had been analyzed as being typical for the different groups of the qualitative study.

### 5.3.1 Participants

---

Almost two years after the recordings of the qualitative research I went back to the same research sites. I had maintained contact with some of the Bolivian informants during the whole time, so I had no problems talking to gatekeepers and making the first contact as two years before.

As sampling strategy typical sampling was employed, i.e. the selection of the participants occurred according to typical characteristics of the informants in terms of the research focus (cf. Dörnyei 2009: 128). For this study these characteristics were primarily related to their linguistic features, representing the different groups of Bolivians established in the qualitative study. This way, I recorded twelve Bolivians. From these recordings three examples were selected for the questionnaires, because they represented the social and linguistic qualities of the groups established in the qualitative analysis

### 5.3.2 Sources of data

---

As a recording device again the Marantz PMD 620 was used because of its small size and easy handling. To stimulate a speech passage offering personal information about the speaker only by linguistic forms I wanted the speaker to tell us a story according to a cartoon strip. Therefore, I used an easily comprehensible picture story concerning a lion and a mouse (cf. Figure III.1). At the time of the recording I sat down, showed the informants the pictures arranged in proper order and asked

them to tell the story as if they were telling it to their Brazilian friends. This way, I tried to prompt a colloquial style in Portuguese that they would employ when talking to their friends. I additionally tried to produce the same situation as two years previously, holding the interviews at the same place in pretty much the same situation, evoking hence the same style by the speaker as in the original interviews to ensure a certain degree of comparability.

However, regarding the place where the recording took place, I had the same problem in 2010 as I had in 2008 recording the interviews of the qualitative part of the study: The recordings had to be done in a public place, because only there would the Bolivians talk freely. This meant, of course, that possible background noises during the recording could influence the results of the questionnaire; a fact that had to be considered for the selection of the examples used in the questionnaire and for the final analysis. After transferring the wav-data to the computer, I used the program Adobe Audition 2.0 to cut coherent pieces of 29 to 60 seconds, because I wanted every example to contain approximately the same amount of words and still present a comprehensible part of the story with a beginning and an ending. Eventually, I manipulated some speech examples with the computer program *Praat* in order to reduce the influence of voice quality. However, when piloting the samples, the pilot group estimated the “computer quality” of the examples as too irritating, leading to the elimination of the voice-manipulated data.<sup>103</sup>

### 5.3.3 Selecting the examples

---

In a further step, the different examples were analyzed according to their linguistic qualities. After having found the most typical ones according to the former analysis of the corpus, six examples were piloted with a group of 10 *Paulistanos* who were asked to write down the characteristics and particularities of the audio examples. As stated above, this process led to the elimination of the voice-manipulated data. . However, I still wanted to control at least some variables influencing the voice quality such as sex and age. Thus, three young men were selected whose age and place of origin coincide and whose voices sounded similar, so that the judgment was not influenced by these factors. Additionally, the examples showing massive influence from Castilian appeared to be too foreign; i.e. people dropped out of the questionnaire once they heard these examples, because they considered

---

<sup>103</sup> My first intention to manipulate the speech examples in a way that the examples would only differ in one single linguistic feature (cf. Campbell-Kibler 2008) appeared to be too difficult for this study. To achieve more validity and reliability, though, the exclusion of all other differences in order to control the different perception of one single linguistic feature seems to be the right approach for future studies. In this study, I tried to control the different factors influencing the answers by adding open-item questions (cf. 5.4.2.2).

them as “not understandable” and “too difficult”. Thus, I decided to select two examples representing the orientation toward PC on the one hand and PP on the other hand, both with high levels of proficiency, and one representing a lower level of Portuguese competence.

Thus, after the piloting, the three examples S I, S II and S III (cf. section 7.1 and Annex III) were selected as usable to investigate the perception and evaluation of the Bolivian speakers by the *Paulistanos* in a quantitative survey. Generally spoken, they differ regarding L1-transfer of Bolivian Castilian and their orientation toward PC or PP respectively (cf. section 7.1 for a more detailed analysis of the linguistic markers of these speakers).

#### 5.4 The quantitative part of the study

---

As was pointed out before, the employment of perception tests to measure the attitude of native speakers toward a group of language learners is not common in sociolinguistics or second language acquisition. Thus, to develop the questionnaire I referred to the survey that Carvalho (2008) conducted in Rio de Janeiro to test the attitude of native *Cariocas*<sup>104</sup> toward the Portuguese spoken by native Spanish speakers learning Portuguese in a classroom environment. Although her point of interest differs slightly from the focus of my study, I adopted various questions into the questionnaire, but complemented them with further items, particularly in regard of the classification of the variety.

The test was directed by the hypotheses that, because of the strong presence of Bolivians and further Latin-American immigrants in São Paulo, the interlanguage of the three speakers would be recognized and – in contrast to Carvalho’s test – reveal social stereotypes connected with the Bolivian immigrants, at least for participants coming from areas of strong Bolivian presence. I further expected that the perception and evaluation of the three speech examples depended not only on the fluency of Portuguese demonstrated by the Bolivians, but on the employment of linguistic markers with social indexicality.

These directional hypotheses made a certain stratification of the sample necessary, because I wanted to compare the possible influence of residence, profession and knowledge of foreign languages as well as gender and age to get a thorough overview, generalizable for the population of São Paulo.

##### 5.4.1 Participants

---

<sup>104</sup> *Carioca* is the Brazilian expression for the inhabitants of Rio de Janeiro, equivalent to *Paulistanos* in relation to São Paulo.

As the population of the study contains all people living in the metropolitan area of São Paulo who are native speakers of Portuguese and consider themselves *Paulistanos*, theoretically, there are almost 20 million potential participants. Given these numbers and the directional hypotheses, it is obvious that no random sample could be composed, but a stratified sample that represents various subgroups according to gender, age, place of residence and profession.

To obtain such a sample it would be ideal to personally select the participants and organize one meeting where all participants can listen to the examples and respond to the questionnaire. Unfortunately, because of limited time and resources in São Paulo, I had no possibility to conduct the survey in this way. Thus, I opted for an Internet survey, relying on the enthusiasm of Brazilians for social networks like orkut, facebook, etc. Naturally, by using an Internet survey, the sampling relies on the self-selection of the participants and already excludes persons who have no access to this medium, i.e. on the one hand more elderly people, on the other hand socially disadvantaged people.<sup>105</sup> I tried to meet this challenge by initiating a chain reaction, actively addressing NGOs working in the poor areas of São Paulo and talking to gatekeepers of the communities to get them to help. Despite the helpfulness of the people, the technical necessity of the survey to listen to the speech examples via the internet heavily complicated the sampling.

All in all, the combination of stratified random sampling and snowball sampling lead to a sample of one hundred completed questionnaires (N = 100), another 23 reached me incomplete<sup>106</sup> and were hence erased. The distribution according to the variables listed above could not be reached for every factor, therefore limiting the generalizability of these results.

---

<sup>105</sup> My original idea of getting one inhabitant of the poorer areas to accompany me and walk around the neighborhood with my notebook, asking people to participate, was frustrated by the refusal of my acquaintances to go along with such a possibly dangerous task.

<sup>106</sup> If single questions like the semantic differential were left out, the questionnaire was included into the survey; the absence of the informants' social data or the evaluation of one speaker leads to exclusion.

Figure 5.11: NI according to gender (N = 100)

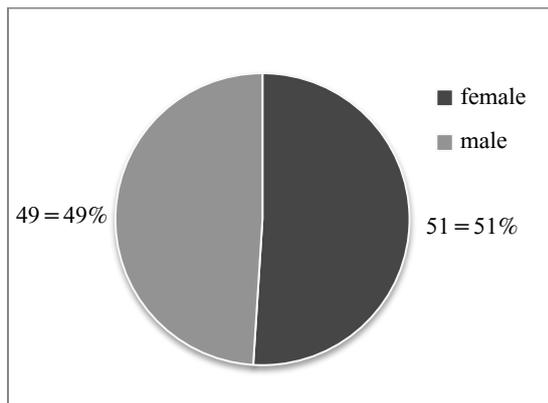
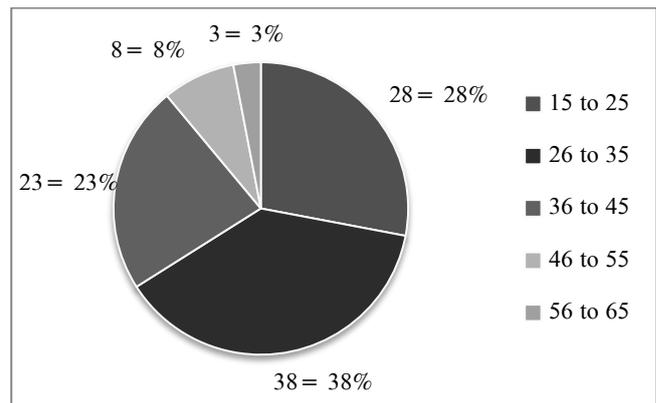


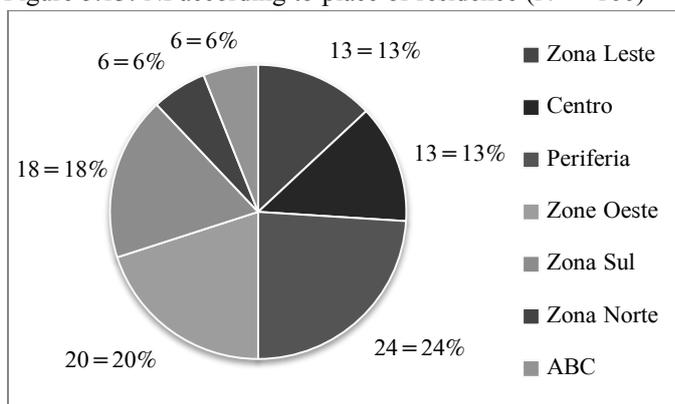
Figure 5.12: NI according to age in years (N = 100)



While the distribution according to gender almost shows equality between men and women, Figure 5.12 shows that the age groups of 15 to 35 years dominate the sample; only 34% of the whole sample are older than 35 and only 11% older than 55 years. This does not represent the age structure of São Paulo’s population, but it indicates how a survey on the internet can exclude certain strata from participation.

Considering the informants’ place of residence, the distribution is given insofar that all major regions of São Paulo – *Zona Leste (ZL)*, *Zona Oeste (ZO)*, *Zona Sul (ZS)*, *Zona Norte (ZN)*, *Centro*, *Periferia*<sup>107</sup> and *ABC*<sup>108</sup> – are represented in the corpus even if the distribution is not equal and not necessarily proportional to the population in the various areas. Referring to the zones defined in Figure 3.5, Norte 2, Leste 2 and Sul 2 were counted as periphery because of their similar socioeconomic characteristics. ZN corresponds hence to Norte 1, ZS to Sul 1 and ZL to Leste 1:

Figure 5.13: NI according to place of residence (N = 100)



<sup>107</sup>

<sup>108</sup> For further explanations regarding the different neighborhoods and my coding procedure see 5.4.4.

Most of the participants live in the periphery regions of São Paulo<sup>109</sup>, only six come from the ZN. Keeping in mind that the neighborhoods of strongest Bolivian presence are located in the ZL, the *Centro* and the *Periferia*, we see that exactly 50% of the participants come from these areas, while ZO, ZS, ZN and ABC constitute quarters with fewer Bolivian immigrants. The social structure of these big areas is not unambiguous, considering that each hosts more than one million persons. But according to the socio-geographic classification given above, I consider ZO, ZS and ABC as socially rather privileged, ZL, ZN, *Centro* and *Periferia* as rather unprivileged.

Thus, social status derives already from the place of residence, complemented by data about the informant's level of education, deduced from the profession and the knowledge of foreign languages<sup>110</sup>. Figure 5.14 shows that only 40% of the sample have non-academic professions while 60% work as academics or are still students, a variable obviously related to the young age of the sample. Figure 5.15 shows that 74% of the informants claim to speak a foreign language, 41% of all even have Spanish competence. Only 26% admit to have no foreign language competence at all.

Figure 5.14: NI according to profession (N = 100)

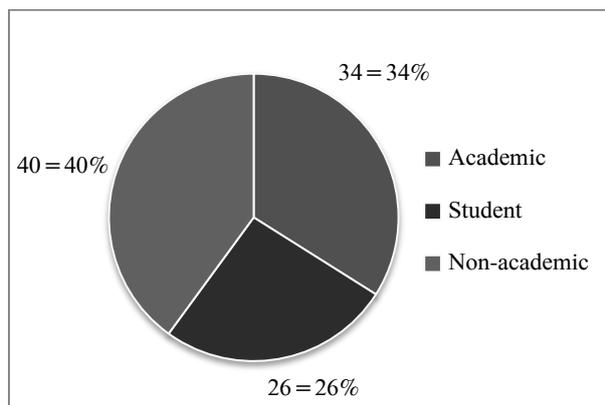
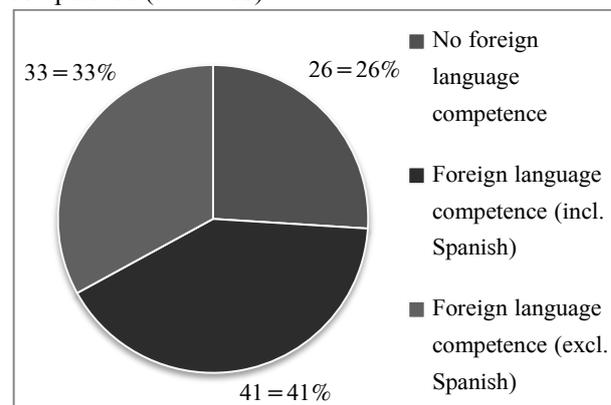


Figure 5.15: NI according to foreign language competence (N = 100)



Relating profession and foreign language competence, it is confirmed that all academics have competence in foreign languages, while nineteen students and fifteen non-academics speak no further language after Portuguese. To illuminate how these factors – obviously indicators for a lower level of education – coincide with the residential neighborhood, Figure 5.16 and Figure 5.17 put profession and foreign language knowledge in relation to neighborhood.

<sup>109</sup> The term *periphéria*, of course, includes all regions far from the center of the city, resulting in an area much larger and more populated than the other areas listed above.

<sup>110</sup> As Carvalho (2008) described this variable as significant for her study in Rio de Janeiro, I also included it in my survey.

Figure 5.16: NI according to residential neighborhood and profession (N = 100)

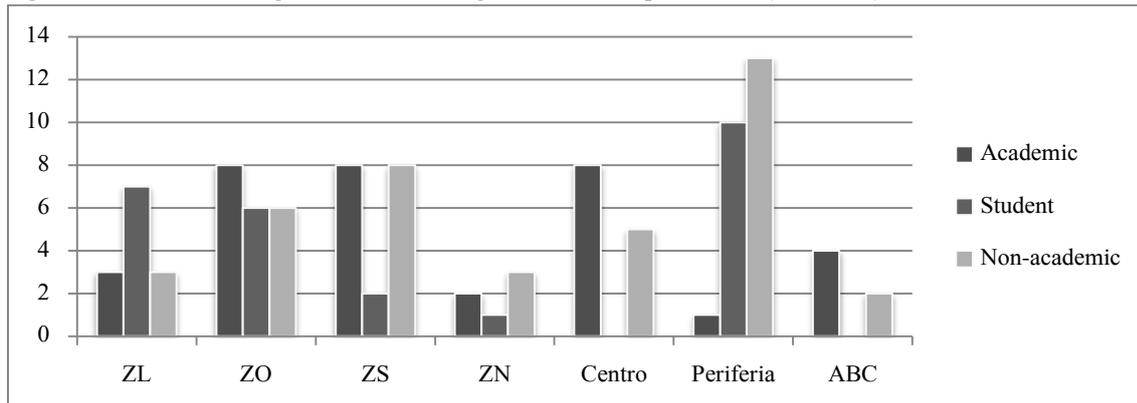
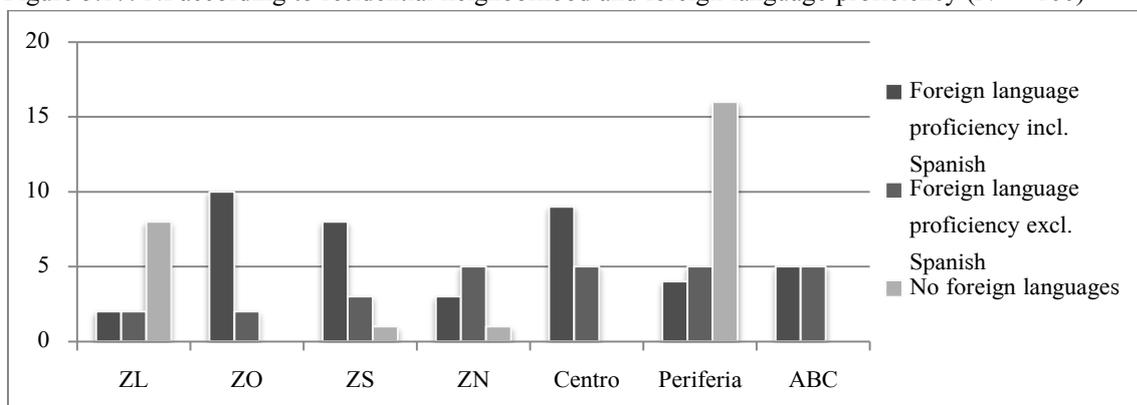


Figure 5.17: NI according to residential neighborhood and foreign language proficiency (N = 100)



Both figures clearly underscore that the *Periferia* and the *ZL* display a different profile from the others – while here, the participants without further foreign language competence outnumber the other participants, only in the *ZS* and *ZN* two further speakers without further language knowledge are found. The distribution in terms of profession is a little more equilibrated, but there is the unequal distribution between non-academics and academics in the *Periferia* as well.<sup>111</sup> While the concentration of the students in the survey could partly be due to the snowball sampling strategy<sup>112</sup> – as I had asked students from these areas to help me with the survey – the quantity of non-academics in the *Periferia* underscores that the sample represents more or less the social distribution of São Paulo’s population.

<sup>111</sup> A 7x3 Chi-square analysis revealed that there is a significant relationship between neighborhood and profession of the participants ( $\chi^2(7, 100) = 28.20, p < .01$ ) as well as between neighborhood and foreign language competence ( $\chi^2(7, 100) = 52.24, p < .001$ ), a fact confirming my discussion of socioeconomic distribution in SP.

<sup>112</sup> The concentration of students in the *ZO* is probably due to the fact that the *Universidade de São Paulo* is located there.

#### 5.4.2 The questionnaire

---

The most important parts of the questionnaire constitute the speech examples whose collection and selection were already described above. After having chosen the examples the questionnaire was piloted with the same Brazilians who had already helped me with the speech example.

In order to keep the questionnaire as short as possible, I developed only five questions that were repeated for every speech example, complemented by two open answer questions about the recent immigration to São Paulo and discrimination in São Paulo's society, and the inquiry about statistically relevant data like gender, age, etc. Because of the self-selection of the participants I avoided direct questions that might have discouraged the participants from participating, e.g. about their salary or their level of education. Instead, I decided to derive the data concerning the social status of the participants from questions about their profession, their knowledge about foreign languages and their residential neighborhood.

The questions about the speech examples are strongly based on Carvalho (2008). As she distinguishes among three components of her survey, the linguistic one measured by the test of dialectal perception, the social one measured by the text of subjective reaction and the personal one measured by semantic differentials (l.c.: 163), I also made this difference and adopted her questions into my survey. In some cases, they were slightly modified; items were added concerning the linguistic and the social aspect, because I wanted to obtain more explicit answers on the social indexicality of linguistic markers.

Thus, in the part about the dialectal perception, I directly asked for linguistic markers that had caught the ear of the informant, before I wanted her to describe the language just heard.<sup>113</sup> These items about the dialectal perception were complemented by the request to choose the estimated place of origin and the current place of residence of the speaker from a given selection. Here, different options were offered ranging from the different areas in São Paulo city to "outside Brazil". If the participant chose an area outside São Paulo, she was asked to specify her opinion.

To control the motivation for the selection and obtain more explicit meta-linguistic comments, I encouraged the participants to give their reasons for the choice. The question about the current place of residence aimed not only for the participants' experience with Bolivian immigrants, but even more for the social status they associated with the speakers.

---

<sup>113</sup> Question 1.2: *Como você descreveria a língua que você está ouvindo? (How would you describe the language you are hearing?)* See Annex III for the original questionnaire in Portuguese.

To obtain more explicit data about the social component, I adopted the question about subjective reaction from Carvalho (2008), where people should choose from various options which job they would offer the speaker if they were president of a company. The options ranged from cleaning staff to manager; this way the social values attributed to the speech example were supposed to become explicit. Like before, the participants were asked to give the reason for their choice.

The third component about the personal qualities of the speaker contained three semantic differential scales between two bipolar adjectives, at the extremes aiming once more for the social characteristics attributed to the speakers. The adjective pairs *pobre – rico* (*poor – rich*), *preguiçoso – trabalhador* (*lazy – hard-working*) and *ignorante – bem informado* (*ignorant – knowledgeable*) were taken from Carvalho's test (2008: 169). The further scales concerning the category of solidarity offered scales between *simpático* and *pouco simpático* (*likeable, little likeable*), *sociável* and *pouco sociável* (*sociable, little sociable*) as well as *atraente* and *pouco atraente* (*attractive, little attractive*).

The scales contained four steps, deliberately offering no possibility to choose the middle between the two poles and hence forcing the participants to select one side. To counter superficial responding, the position of the positive poles was changed once.

Altogether, several questions were included about the social and linguistic component to get more explicit answers about the social status of the speakers in São Paulo and particularly about the influence of linguistic markers.<sup>114</sup>

### 5.4.3 Procedure

The html version of the questionnaire was uploaded on the server of *Freie Universität Berlin*, reachable via [http://userpage.fu-berlin.de/sniewhoff/Como\\_fala\\_Sao\\_Paulo.php](http://userpage.fu-berlin.de/sniewhoff/Como_fala_Sao_Paulo.php).<sup>115</sup> This link was active from April 2010 until April 2011. The audio files could be played using Firefox or Internet Explorer, but I still received some complaints that it was impossible to listen to the files. Thus, the option to down-

---

<sup>114</sup> In Carvalho's test (2008: 163), she implies that the speakers were evaluated only because of their level of Portuguese, i.e. the speech example with least Spanish influence is more advanced. This way, however, she neglects factors like the more personal way of telling the story by the *less advanced* speaker using diminutives like *bonitinha* or emotional expressions like *sinto muita pena, é muito triste para mim* (l.c.: 162) that easily could have influenced the evaluation of the example. Naturally, these factors can only be controlled in experimental designs where the examples only differ with regard to one linguistic feature without the interference of any personal narrative style. Despite this fact, it seems that employing classifying questions helps to illuminate the reason for the evaluation.

<sup>115</sup> The title *Como fala São Paulo* was used in order to awaken more interest when posting the link in social network groups like *Eu amo São Paulo* on Facebook or *Zona Leste* on Orkut.

load the mp3 files was added. To make the comparison between the examples easier, they could be stopped and restarted at every moment and played as often as the participant wanted. Additionally, all questions and examples were visible on one page; hence the participants could change their opinion about the first participant after having listened to the third one, etc.

The link was sent to all my personal acquaintances living in São Paulo asking not only for their participation, but also for them to forward the link to any *Paulistano*. Additionally, I posted the link in groups about São Paulo on Facebook and Orkut, actively trying to appeal to people from the *Zona Leste*, and addressed NGOs working in São Paulo.

Unfortunately, despite the piloting process, I had to erase the third question about the subjective reaction after four weeks, because it became obvious that the question was not only answered reluctantly, but even lead to incomplete results that were called off at exactly this question. As there were other questions also aiming for the social classification of the speakers, I estimated that it was better to erase the question than to risk the failure of the survey by too many discontinued questionnaires.

#### 5.4.4 Data analysis

---

Before I started with the statistical analysis, the online responses had to be coded and transferred into tables using the spreadsheet and calculation software 'Microsoft Excel' and the statistical analysis software SPSS. As the questionnaire contained many simple open-ended items, e.g. asking for the place of residence of the informants, I first assigned a new number to every new area that was mentioned. Afterwards, I built the seven categories according to the *Zonas* enumerated above. For the analysis according to Bolivian presence in and social status of the areas, the seven categories were further combined until two or three categories were achieved. In order to maintain the important correlations and differences, before merging the categories, various correlation tests were conducted to confirm the right combination.

Apart from erasing the incomplete questionnaires and the answers to the question that had been removed, no further data cleaning was conducted. After having coded all responses, the data was quantitatively computed by using SPSS. In order to test the data for correlations and statistical significance, I employed not only descriptive but also inferential statistics like chi-square tests, correlation tests, t-tests as well as one-way and two-way analyses of variance (cf. Dörnyei 2009: 197-241, Duller 2006, Fromm and Baur 2008, Janssen and Laatz 2007 and Rowntree 2004). To illuminate the correlations and differences, diagrams and tables were used (cf. Chapter 7).

After the presentation of the quantitative analysis, the qualitative and quantitative parts are interpreted together, indicating not only congruent but also contradictory results.

## 5.5 Summing up

---

Since the aim of my study is to describe and explain the speech of Bolivian immigrants in São Paulo, with special regard to the interaction between immigrants and the receiving society, I employed a sequential exploratory approach, consisting of a dominant qualitative part concerning the speech of Bolivian immigrants and a complementary quantitative survey measuring the evaluation of the immigrants' speech by the *Paulistanos*. The basis of the qualitative analysis is the prolonged engagement with the Bolivian community in São Paulo enabling me to observe their way of living and to conduct sociolinguistic interviews with 70 immigrants. From these interviews, 28 examples were selected for the linguistic analysis according to factors like social status, social network and their access to educational institutions in Brazil. In order to cover the whole range of the Bolivians' living situations in São Paulo, non-representative extreme cases were included in the sample. The observations about their social position were triangulated with the linguistic facts I obtained from my interviews. The analysis led to the hypothesis that three groups could be differentiated according to their linguistic behavior, because they pursued different strategies of accommodation toward the receiving society. To confirm this hypothesis and overcome my biased opinion as the result of my involvement as researcher at every step of the qualitative study, an online survey was conducted asking 100 *Paulistanos* about their linguistic, social and personal evaluation of three prototypical speech examples of Bolivians. The examples consisted of three Bolivian immigrants of the same gender, age and roughly similar voice quality telling a short fable that was stimulated by a comic strip. These speakers were selected to represent the different groups obtained from the qualitative study with regard to the linguistic qualities, e.g. language proficiency, accommodation toward the Portuguese norm and to norm-distant varieties.

For the quantitative part, I achieved an almost equal dispersion of the sample regarding gender and place of residence in regard to Bolivian presence. As a whole, however, the participants of the online survey are not representative of the population of São Paulo, because the form of an online-survey excluded older people as well as those living in poor social conditions without access to a computer or the internet. Additionally, the technique of snowball sampling and the participants' self-selection, inevitable for an online survey, seem to have favored *Paulistanos* who know foreign languages as

well as students and academics over those with non-academic professions. Thus, the results can only be regarded as tendencies.

The responses obtained during various months were coded, statistically analyzed employing SPSS and, upon finalization, compared with the results of the qualitative survey. In this way, both qualitative and quantitative aspects of the study were played to their strengths, combining the depth of the first part with the breadth of the second one and allowing hence a complete picture of the Bolivian immigrants' Portuguese acquisition as an interactive product of them and the receiving society of São Paulo.

The results of the analyses will be presented in the further chapters following the chronological order of qualitative analysis, quantitative analysis and concluding discussion.

## 6. Characteristics of the Bolivian's Portuguese in São Paulo

---

This chapter contains the detailed linguistic description and qualitative analysis of the first part of the study regarding 15 linguistic features from phonology, morphosyntax and vocabulary. In order to show which underlying processes are responsible for the speech of the Bolivian immigrants in São Paulo, I compare the examples with its counterparts of the varieties in contact, i.e. the Portuguese varieties PP and PC and the Spanish varieties CP and CC. Additionally, the distribution of the linguistic particularities among the Bolivians' speech are described and analyzed. Considering the group and the individual level, similarities among the whole group will be enumerated as well as differences between the speakers in order to discover underlying patterns characterizing their Portuguese.

The analysis shows that L1-transfer, accommodation (toward different target varieties) and general processes of language learning like simplification each play their role for the actual speech of the Bolivian immigrants. The differences can be attributed to three groups that are not only distinguishable by their linguistic performance but also their social networks.

---

*Então você tem a dúvida se está falando ou o português mesmo ou já está utilizando o término espanhol (A18 (30) m. 2004 P. Ce: 3.38).*

The aim of my study is to offer a thorough insight into the Portuguese of Bolivian immigrants. Naturally, not all aspects of their (Portuguese) speech can be equally considered, but I set out to analyze and discuss features from phonology, morphosyntax and vocabulary, that show which linguistic processes are responsible for the speech of the Bolivians, and which linguistic and social factors might trigger these processes. Leaning on many examples from the corpus, the analysis will not only focus on the individual but also on the group level.

One main difficulty for the identification of these factors concerning my study lies in the close relatedness of Spanish and Portuguese:

*Nunca, frente a un fenómeno concreto, puede el hablante (ni el lingüista) estar seguro de cuál sea su origen, si lusitano o hispánico (Elizaincín 1996: 417).*

The two languages not only stem from the same Latin origin, but also share a good part of their history – Portugal even belonging to the Spanish empire of Philipp II during the 16<sup>th</sup>/17<sup>th</sup> century (1580-1640) – and, of course, the geographical proximity on the Iberian Peninsula and in South America that resulted in continuous language contact since the diverging development of the two languages in medieval times (cf. Bollée and Neumann-Holzschuh 2008: 34 ff.). Therefore, it is common opinion that the two languages are mutually intelligible, although “Portuguese speakers understand Spanish better than vice-versa” (Jensen 1989: 851). However, Jensen (1989) shows that by measuring the mutual intelligibility of electronically reproduced voices, its level amounts only up to 60% (cf. l.c.).

This fact implies that the intelligibility depends on the one hand on the direct conversation where the pragmatic context of the conversation as well as gestures and the facial expression of the talking persons facilitate the understanding. On the other hand, it also shows that written texts are easier comprehended than oral utterances because of the lack of phonetic interferences that obscure the common vocabulary and the more normative and conservative style typically associated with written language:

*[Q]quanto mais a oralidade se distancia da ‘norma culta’, maior é a distancia entre as duas línguas (Ribas Fialho 2005: 2.2).*

Even though this seems like a rather rough generalization, it directs the attention to an important point for my study: I have to consider all varieties in contact, not only the standard norms as it is often done. Already the choice of the American or the Iberian varieties strongly influences the findings when comparing the two languages, as the vocabulary shows. Various researchers claim that 89% of Spanish and Portuguese vocabulary are cognate (cf. Carvalho 2002: 599), but according to Simões and Kelm (1991), this level of similarity applies only to the varieties of the Iberian Peninsula and not to the American varieties. Because of the influence of lexical borrowing from the Indian and African languages present in Latin America’s (particularly Brazil’s) history, they assume that at most 55-60% of the vocabulary are similar there (Simões and Kelm 1991: 655).

## 6.1 The varieties in contact

---

In the case of São Paulo, because of its spatial, social and cultural separation, the socioeconomic classes differ strongly with regard to the social networks, level of education and literacy rates. Due to these differences I claim that it is reasonable to indeed speak of two varieties dominant in the city: São Paulo’s *Português Popular* (PP) and *Português Culto* (PC). I understand these varieties as theoretical concepts that help with the analysis of the study, although they mainly constitute the poles of three continua with regard to alphabetization (more or less literate speakers), urbanization (social network) and stylistic monitoring e.g. communicative distance versus communicative proximity (cf. Koch and Oesterreicher 2011: 10 ff.) as suggested by Bortoni-Ricardo (1998). However, it has to be considered that in communicative practice, no speaker reaches these poles, but chooses linguistic features that rather indicate one or the other variety – dependent on her abilities, e.g. for stylistic variation, and the requirements of the communicative situation such as the spontaneity of the communication, the emotional involvement and the relationship between speaker and hearer (cf. Koch and Oesterreicher 2011: 7). The continua condition the speech of urban speakers. They unite various

linguistic features that – due to their social indexicality – are rather associated with PP or PC. PP constitutes the theoretical default of communicative proximity, i.e. every parameter conditioning the communication indicates proximity. PC, on the other hand, corresponds to the concept of communicative distance.

While speakers of the higher socioeconomic classes might use features rather associated with PP for stylistic reasons and in informal situations, the possibilities of stylistic variation are limited for speakers from the lower socioeconomic class, because their communication is mostly limited to contacts inside their own social class and few CoP.

This means that speakers of low social class mostly employ features associated with PP, because they are not familiar with PC. However, this does not automatically entail that they would like to speak PC as the variety of higher prestige. On the contrary, it is probably socially more rewarding to refer to influential social groups inside their lives, in the regions and networks they live in, than to refer to an abstract norm spoken at university.

The most salient features of PP are stigmatized in the perception of members of the higher socioeconomic class that have little direct contact with speakers of the lower class, because they hear PP mainly through the media in reports about crime and violence, even if they themselves employ these stigmatized features in informal situations.

The Bolivian Castilian varieties that have to be considered in this contact situation differ most with regard to the influence of the indigenous languages; at the census of 2001, 50% of Bolivia's people claimed to speak one of the indigenous languages, but mostly also Castilian<sup>116</sup> (cf. Cancino 2008: 4). Considering the distribution of the languages in Bolivia, the indigenous languages, mainly *Aymara* and *Quechua*<sup>117</sup>, are spoken first and foremost in rural areas of Bolivia. The *Quechua* speaking population constitutes the biggest contingent with more than four million people speaking it as their native language, concentrated in the departments of *Cochabamba*, *Chuquisaca* and *Potosí* (cf. Cancino 2008: 5); the *Aymara* speaking population primarily lives in the departments of *La Paz* and *Oruro*.

---

<sup>116</sup> According to the census 2001, only 11% of the population is monolingual in one of the indigenous languages; this confirms the ongoing process of bilingualization in Bolivia pictured by Pfänder (2002).

<sup>117</sup> Since the constitution of 2006, there are 26 official languages in Bolivia: *Aymara*, *Baure*, *Besiro*, *Canchama*, *Cavineño*, *Cayubaba*, *Chácobo*, *Chimán*, *Ese Ejja*, *Guaraní*, *Guarasuwe*, *Guarayu*, *Itonoma*, *Leco*, *Majineri*, *Mojeño Ignaciano*, *More Mostén*, *Movida*, *Pacawara*, *Quechua*, *Reyesano*, *Sirionó*, *Tacana*, *Tapiete*, *Toromona*, *Uru Chipaya*, *Weenhayek*, *Yaminawa*, *Auki* and *Yuracaré* (cf. Cancino 2008: 10). As *Quechua* and *Aymara* are the languages with highest speaker numbers, by far, I will focus on these languages.

In the cities, *Aymara* or *Quechua* proficiency is less spread than in the rural areas, but a considerable part of the population living in urban agglomerations claims to be bilingual: In 2001, 38% of La Paz' population declared to speak *Aymara* (cf. Cancino 2008: 4). Considering the first language of rural and urban speakers, though, there are major differences: While in the rural areas 61.19% of the total adolescent population has one of the indigenous languages as L1, in the city, 80.17% learned Spanish as L1 (Sachdev, Arnold and Yapita 2006: 110). Despite of the official status of the indigenous languages since the constitution of 2009, some still consider them inferior to Castilian representing thus another separation inside Bolivia:

*[...] dialeto. Esses só são dialetos. Eu [o falo] não , mas dentro da família do meu pai, sim. Eles falam o quechua que é o dialeto que se utiliza lá em Cochabamba.[...] Eu intendendo alguma coisa mas falar assim direitinho não (A18 (30) m. 2004 P. Ce: 1.12).*

Regarding the immigrants coming to São Paulo, there is no data available about their language competence. Because of the difficulties connected with the immigration it seems rather implausible though, that there are people coming to São Paulo with no competence in Castilian at all. As most of them come from urban centers, however, they represent the above mentioned distribution of bilinguals that grew up with Castilian. Most of them confirm that at least one of the parents speak either *Quechua* or *Aymara*, but they themselves admit to have less competence in the indigenous languages:

*A gente fala escuta pessoas que sempre escutam al vez falam, né, você fica ya tanto escutando muito ya fica sabendo algumas coisas; mas falar falar eu não sei não (A5 (28) f. 2000 D. ZL: 1.38).*

This is attributed to the low divulgence of *Aymara* in urban society:

*Eu deveria falar né Aymara mais, mas não falo né [...] a sociedade só fala espanhol (A14 (28) m. 2006 S. ZN: 0.34).*

The immigrants coming from rural areas, however, affirm that the indigenous languages are the regularly employed languages in their family:

*Eu falava Aymara com minha mãe com mis irmãos com minha pai eu falo Aymara. Eles falam Aymara constante, então eu tem que responder em Aymara; não pode responder em el castilhano, es falsa, no (B4 (27) m. 2003 S. Ce: 10.27).*

Additionally, the narrow and prolonged contact with the indigenous languages influenced the Castilian of South America itself, resulting in the formation of a language variety called *Español Andino* that is particularly well studied in Peru (cf. e.g. Escobar 2000, Godenzzi 2007 and Klee and Caravedo 2009). As the name suggests, *Español Andino* is first and foremost diatopically defined as (originally) spoken in the whole Andean region stretching from Colombia to Bolivia. According to Men-

doza (2008), particularly phonetic/phonological features are only found in the speech of bilingual speakers, while morphosyntactic features are already part of the language of monolinguals as well:

*[O]curre con frecuencia que el bilingüe en su aprendizaje del castellano supera generalmente todas o casi todas las transferencias fonético-fonológicas. Entonces estos rasgos no llegan a sentar presencia en el habla de monolingües. Muy diferente es el comportamiento de las transferencias morfosintáticas [...]. Muchos de estos rasgos han llegado a constituirse en parte del habla de los monolingües tanto de la variedad culta como de la variedad popular, especialmente en el registro oral informal (Mendoza 2008: 217).*

Thus, similar to the situation in Brazil, the urbanization and massive migration to the cities has led to the variety's spread in the whole country, becoming diastratically marked as having little social prestige. However, some features are already part of the norm oriented variety as well. Compared with the above mentioned division of *Norma Culta* and *Popular* in Brazil, I consider Bolivia's varieties as theoretical concepts. Similar to the situation in Brazil, Bolivian *Español Andino* and *Bolivian Castellano Culto* constitute the poles of the continua like urbanization, alphabetization and stylistic monitoring, complemented by the continuum of bilingualism. Since the literature suggests that this last continuum divides the country's speech community, it is assumed for this study that two varieties must be considered, even if these extremes are never spoken in communicative practice. In analogy to Portuguese the Bolivian variety of communicative proximity will be called *Castellano Popular* (CP); the Bolivian variety of communicative distance will be called *Castellano Culto* (CC).

Thus, when analyzing the corpus, I have to consider the influence of the varieties important for my study. As was explained before, this concerns São Paulo's Portuguese varieties PC and PP and Bolivia's Castilian varieties CC and CP. The distinction between PC and PP is particularly interesting because linguistic features that differ according to these two varieties offer valuable information if and to what degree social indexicality and social networks are important for the Bolivians

## 6.2 The analysis of the corpus

---

The features analyzed in this study are selected in order to verify or falsify the thesis presented before that various processes like accommodation, L1-transfer and congruent lexicalization can be observed when observing the individual data of an entire group of immigrants. Depending on their social networks, accommodating processes towards PP will be more present for some, while for others rather accommodation towards PC, or no accommodation at all with a strong L1-transfer, is observable. This does not mean, however, that the different occurrences are unambiguously attributable. Because of the similarity of the languages and the strong variation between the speakers I will

indicate multiple possibilities where applicable. To ensure the transparency of my analysis, I will point out the relevant data from the Bolivian (source) varieties and the Brazilian (target) varieties.

As this is a qualitative study, I will not count the occurrences of each feature, but comment on the distribution inside the group. Nevertheless, valid examples show the linguistic particularities of my informants' different interlanguages and do the complexity of language learning in a natural environment justice.

### 6.2.1 Semi-open vowels $\epsilon$ and $\text{ɔ}$

---

*It is especially challenging for me to learn and incorporate the nasal sounds non-existent in Spanish, as well as the open and closed vowel sounds in spontaneous conversation (Cowles and Pires 2008: 245).*

This quote comes from a Hispanic student studying Portuguese in the USA. She has these difficulties although she learned in a classroom-environment, where teachers systematically attend to this particular problem of perception and production (cf. Simões and Kelm 1991).

The underlying process can be explained by L1-influence, because – while both languages share the oral vowels /i/, /e/, /a/, /o/ and /u/, the Brazilian system also contains the phonemes / $\epsilon$ / and / $\text{ɔ}$ / (see Tables 6.1 and 6.2) and a system of nasal vowels and diphthongs. In Castilian, [e], [ɛ] and [o], [ɔ] have allophonic status, as Gordon (1980: 349) confirms.

Table 6.1: Castilian vowel system (cf. e.g. da Hora 2000: 16)

	Front	Central	Back
High	i		u
Middle	e		o
Low		a	

Table 6.2: Brazilian Portuguese vowel system (cf. e.g. da Hora 2000:17)

		Front	Central	Back
High		i		u
Middle	Closed	e		o
	Open	$\epsilon$		$\text{ɔ}$
Low			a	

In accordance with these findings, it does not come as a surprise that my corpus contains many examples for the free allophone variation between [e,  $\epsilon$ , o,  $\text{ɔ}$ ] without realizing the phonemic distinction

of /e/ and /ɛ/, and /o/ and /ɔ/ as in (1). In this example the speaker realizes the same word, *nove*, first with closed [o] and then with open [ɔ].

- (1) *Estoy **nove** [ˈnovi] anos ya em janeiro ya **nove** [ˈnovi] anos que sou cumprido ya (A5 (28) f. 2000 D. ZL: 1.16).*

A similar case can be seen in (3), where it is not clear if the speaker is talking about his grandmother *avó* or grandfather *avô*<sup>118</sup> because the vowel is neither really open nor really closed.

- (2) *Es a [...] a mãe de mi de mi **avó**. A irmã de mi **avó** (A10 (22) m. 2006 S. Pe: 2.31).*

As *avó* [aˈvɔ] and *avô* [aˈvo] represent one of the most salient minimal pairs with /o/ and /ɔ/, Portuguese speakers tend to pronounce these words very carefully to facilitate the differentiation by the hearer.

For /e/ and /ɛ/, my data are similar: /ɛ/ is pronounced as [e] as in (3), while /e/ is pronounced as [ɛ] as in (4) and (5).

- (3) *Algunas veces me siento muy mal también [...] de mis **pés** [pes] (A25 (60) f. 1989 S/SA ZL: 3.51).*  
 (4) *Ele trabalha aqui como médico ((0,3s)) né e **ele** [ɛɛ] ((0,3s)) ele [ɛɛ] é bom (A23 (45) f. 1990 Ph ZO: 2.56).*  
 (5) *Estoy fazendo curso do **português** [portu'ges] (A22 (27) f. 2008 P. ZS: 0.8).*

In accordance, this free variation of [o] and [ɔ] also concerns the minimal pairs *sou* (1<sup>st</sup> pers. sg. of the verb *ser*) that is regularly monophthongized to [so] in oral Portuguese, while *só* (just, only) has to be pronounced as [sɔ]. In the following examples of my corpus, the two words become homophones:

- (6) *Tinha um boulevard ((0,4s)) e aí tinha outro tipo shopping. ((0,2s)) Só. [so] (A18 (30) m. 2004 P. Ce: 16.05).*  
 (7) *Todo que eu f... eu f... tô facturando estou facturando só [so] mando à Bolívia porque tenho família lá (A4 (21) m. 2006 S. ZL: 12.0).*  
 (8) *Com a leitura sou obrigado [sɔ ɔbri'gado] né a faz ... aprender português (A6 (34) m. 1999 S. ZL: 30:48).*

Both languages, Castilian and Portuguese contain the semivowels [j] and [ɥ] occurring in one syllable after a full vowel, forming falling diphthongs like [aj], [ej], [oj], [aɥ], [eɥ] and [ouɥ], as in Portuguese *caixa* ['kajʃa], *coisa* ['kojʃa], *couro* ['koɹu] or in Castilian *veinte* ['bejnte] or *feudal* ['feɯðal]. In oral Portuguese, however, these diphthongs are often monophthongized, resulting in [aj] being pronounced as [a], [ej] as [e] and [ouɥ] as [o], for example *caixa* ['kaʃa], *brasileiro* [brazil'leru] and *vassoura* [va'sora] (cf. Castilho 2010: 206).

In CP, the phenomenon of monophthongization is only described for rising diphthongs as in the following examples: *quiero* ['kero], *aunque* ['anke] (cf. Mendoza 2008: 220; Gordon 1980: 349).

<sup>118</sup> As A10 (22) m. 2006 S. Pe uses the genus-indifferent Castilian possessives, the identification is further complicated.

Interestingly, the use of open [ɛ] and [ɔ] is even found for the diphthongs /ou/ and /ei/ as in (9) and (10). I suggest that the variation between closed and open vowels is applied to all sounds that the Bolivians perceive as /o/ or /e/, even if the original sound was [ou̯] or [ei̯].

(9) *Trabalhava com brasileiros* [ˌbrasiˈlerus] (A2 (18), f. 2005 SA ZL: 0.43).

(10) *Nos prevestibulares que eu fui que participei* [ˌpartisiˈpeɪ] [...] *ya passei* [paˈsɛ] *então* (A6 (34), m. 1999 S. ZL: 33.31).

While the realization of *sou* as [so] and of *dinheiro* as [diˈɲeru] and *brasileiro* as [ˌbrazɪˈleru] is common in Brazilian Portuguese, the diphthongs in *participei* [ˌpartisiˈpeɪ] and *passei* [paˈsɛ], however, would not be monophthongized. Thus, the use of the open [ɛ] in [ˌpartisiˈpeɪ] and [paˈsɛ] can either be attributed to the transference of the corresponding Spanish morpheme (1<sup>st</sup> pers. sg. past tense) -é as in *participé*, the insertion of the whole word, or the analogous application of the monophthongizing rule to every diphthong. However, the same speaker realizes the diphthong in *leitura* correctly as [leɪˈtura]. Thus, this last option seems rather unlikely; the context with *já* pronounced as Castilian [ja] rather confirms the interpretation as L1-insertion / congruent lexicalization as described by Muysken (2000).

Regarding the distribution of this feature, my examples show that almost all speakers use it, be it a young woman having arrived in São Paulo only some months ago, an elder pharmacist having been in São Paulo for eighteen years or seamstresses having worked and lived in the city for already nine or just two years.

### 6.2.2 Diphthongization before final /s/

For Brazilian Portuguese spoken in São Paulo – independent from social class – the reversed phonetic process can also be described, i.e. the diphthongization of monophthongs by adding an epenthetic [j]. This phenomenon concerns the stressed vowel if followed by a final sibilant: thus, *luz* is realized as [lujz] and *atrás* as [aˈtrajs]. Although the non-diphthongized variant, which is mostly equivalent to the Castilian realization, is used as well, the diphthongization is a rather salient feature. Therefore, the underlying process for examples of this feature in my corpus must be interpreted as accommodation. My corpus includes many examples of this feature, particularly of the diphthongization of /os/ to [ois] as in (11), of /es/ to [ejs] and [ɛjs] in (12) to (14) and the most common process from /as/ to [ajs] in (11), (15) and (16).

(11) *Agora as autoridade de aqui a polícia consulado a embaixada da Bolívia faz* [fajz] *nada pra nos* [nos] *por nos* [nojs] (A9 (42) m. 1994 D. ZL: 18.08).

- (12) **Português** [ˌportuˈgejs] direito porque nó queria atrapalhar na escola (A12 (39) m. 1990 D. ZL: 3.34).
- (13) Falam espanhol **português** [ˌportuˈgejs] e **inglês** [iŋˈglejs] (A2 (18) f. 2005 SA ZL: 2.03).
- (14) Eu pretendo ficar aqui até dois mil e **dez** [dejs] (A1 (23) m. 2006 SA Ce: 2.56).
- (15) Ya aconteceu um um **rapaz** [haˈpai] aí de repente aconteceu um problema e apareceu o **rapaz** [haˈpaɪs] que morava lá na minha casa (A5 (28) f. 2000 D. ZL: 3.21).
- (16) **Faz** [fais] tiempo (A25 (60) f. 1989 S/SA ZL: 0.16).

The speakers display a lot of variation regarding diphthongization. In the corpus, there are instances of the same vocabulary spoken by the same speaker as in (11) to (16) but without diphthongization. The discourse marker *mas*, however, is regularly realized as [maɪs] by all speakers as in (17) to (19), even though it becomes homophone with *mais* as in (122) and (123).

- (17) **Mas** [maɪs] dei mis primeiros passos lá (A16 (32) f. 2004 P. ZL: 2.25).
- (18) Eles faziam igual encontro tipo Kantuta **mas** [maɪs] era **mais** [maɪs] pequena (A7 (32) f. 1994 S. Ce: 9.28).
- (19) Não, sou da cidade **mas** [maɪs] eu deveria falar né ((0,4s)) aymarará **mais** [maɪs]. **Mas** [maɪs] não falo né (A14 (28) m. 2006 S. ZN: 0.31).
- (20) Ah é difícil lidar com minhas a ser com meus irmãos bolivianos [...] **Mais/mas** [maɪs] com ... ((0,5s)) falo com eles **mas** [maɪs] pouco **mais** [mas] é brasileiro (A9 (42), m. 1994 D. ZL: 6.30).

In (20), the speaker even monophthongizes *mais* pronouncing it [mas], while he diphthongizes *mas* [maɪs], resulting in an ambiguous transcription; since it is impossible to decide what the speaker wanted to say in this case. This may signify that accommodation due to the social indexicality of this feature is more important to the speaker than the semantic distinction – a clear sign for the importance of accommodation.

In the case of some speakers, diphthongization even occurs with their city of origin, the Bolivian La Paz [laˈpaɪs] as in (21) to (23). As Bolivian names and toponyms are normally maintained in the original language, I interpret these cases as hypercorrection:

- (21) Mas também eu sou de La **Paz** [laˈpaɪs] (A28 (28) m. 2002 S. ZL: 1.31).
- (22) De quatro anos eu voltei pra La **Paz** [laˈpaɪs] (A5 (28) f. 2000 D. ZL: 5.59).
- (23) Eu vim de la cidade de La **Paz** [laˈpaɪs] (A12 (39) m. 1990 D. ZL: 0.35).

These hypercorrections are not found for the informants working i.e. as physicians – outside the textile industry –, but mainly for speakers having spent a lot of time in São Paulo in a poorer social environment. While the common pronunciation of *mas* as [maɪs] could be due to the frequent use of the discourse marker in Brazilian speech, the generalization of this rule concerns most the speaker group of lower social class that has frequent contact with Brazilians. This finding points again to accommodation as the underlying process for this feature – and that for highly salient features ac-

commodation clearly overrules L1-transfer, at least for those immigrants who are rather oriented toward social assimilation and not as norm-oriented as the academics.

### 6.2.3 Nasal vowels

Another phenomenon that Spanish and Hispano-American students of Portuguese regularly describe as most difficult, are the nasal vowels. In contrast to Bolivian Castilian<sup>119</sup>, Portuguese contains nasal vowels and diphthongs that have phonemic character<sup>120</sup>: The nasal vowels /ĩ/, /ẽ/, /ã/, /õ/ and /ũ/, and the nasal diphthongs /ẽĩ/, /ãĩ/, /õĩ/, /ũĩ/ and /ãũ/ that originate from the combination of nasal vowels with the nasalized semivowels [ĩ] and [ũ]. In spoken Brazilian Portuguese and depending on the phonetic context, these nasal diphthongs are often reduced to monophthongs the same way as observed for the oral diphthongs. In rather norm-distant PP, the monophthongization often occurs together with denazalization: *eles comem, os homens* or *eles falaram* are pronounced as [eĩs 'kõmi], [uzũ'õmi], [eis fa'laru] (cf. Castilho 2010: 206).

Accordingly, the “difficulties” with the nasals can rather be explained by markedness – learners tend to learn more marked fields of the target language last. As nasals are considered as universally more marked than oral vowels, speakers of any source language should have these difficulties. Additionally, because of the closeness of the languages, this phenomenon could also be described as morphological integration occurring in the process of congruent lexicalization, i.e. using the Castilian morpheme instead of the Portuguese one (cf. Muysken 2000: 134).

Accordingly, regarding the pronunciation of the nasal diphthong [ãũ] a great amount of variation can be discovered in my corpus, not only concerning the whole group but also individual speakers. In (24) and (25), there are three occurrences in each example. Each of these occurrences is pronounced differently:

(24) *Então* [en'tõ] *não* [nõ] *sofrí discriminação* [diskrimina'sãũ] (A23 (45) f. 1990 Ph. ZO: 5.29).

(25) *Na realidade eu acho que ((0,5s)) esse é o problema do da educação* [eduka'sãũ] *do Brasil. [...] Com várias ramificações né não* [nõ] *tem uma solução* [solu'sõũ] *só* (A19 (42) m. 1988 P. Ce: 6.31).

<sup>119</sup> For Mexico and the Antilles, allophonic nasalization before a nasal consonant is described as regular, but not for Andean Spanish (cf. da Hora 2000: 21).

<sup>120</sup> The discussion about the phonological status of the nasal vowels in Portuguese has been going on for many years now. While some claim that there are no phonemic nasal vowels, but only oral vowels followed by a nasal archiphoneme (cf. e.g. Quilis 1979: 5-7), Marques Abaurre and Pagotto (2002) act on the assumption that it is agreed on that “[...] há, na língua, dois tipos de nasalidade, a nasalidade fonológica, com função distintiva e a nasalidade fonética (não-distintiva)” (l.c.: 492), even if this fact is interpreted differently. For a more detailed overview about this discussion see e.g. da Hora (2000: 18-19). As it is not the aim of this study to resolve that question I claim that there are five phonemic nasals, as proposed by e.g. Castilho (2010: 49).

As the examples already show, the realization of *-ão* oscillates between as distinct sounds as the nasalized Brazilian Portuguese standard [ã̃], the denasalized diphthongs [aũ] and [oũ], vowel-nasal combinations like [ɔ̃n] (and [an]), the nasalized vowel [õ<sup>n</sup>] (and even [ã]), and [ɔ]. At first sight, the different realizations seem to appear in free variation, but at a closer look a certain pattern appears: The use of the variants depends mainly on the Castilian counterpart of the vocabulary, more specifically on the (morpho)-phonological material of the Spanish counterpart of the words that contain the nasal diphthong in Portuguese. Thus, the Portuguese *São* is pronounced [an] according to Castilian *San* [san] in (26); analogously, the verb morpheme indicating 3<sup>rd</sup> pers. pl. of indicative present in Portuguese, *-am* [ã̃], is realized as [an] in (27) and (28), the equivalent morpheme in Spanish. Considering (26) and (28), it becomes obvious that the distinction between the different realizations must depend on the Spanish influence as the homophone *São* (holy) and *são* (are) realized differently as [san] and [sɔ̃n].

- (26) *Já fui no São [san] Camilo na universidade São [san] Camilo eu já fui no São [san] Marcos também (A17 (31) m. 2007 P. Ce: 3.16).*
- (27) *Eles **trabalham** [tra'baʎan] juntos. [...] Chega um e **ajudam** [a'ʒudan] ele [...] Aqui **não** [nãũ] ((0,3s)) somos muito desunidos (A16 (32) f. 2004 P. ZL: 7.21).*
- (28) ***Deveriam** [deve'rian] estar cadastrados porque ((0,4s)) muitas vezes **são** [sɔ̃n] assassinados aqui na Zona Leste ninguém fica sabendo os familiares **não** [nɔ] **ficam** [fikan] sabendo **são** [sɔ̃n] enterrados como (indígenos?) (A21 (62) m. 1964 E. Ce: 12.11).*
- (29) ***Então** [en'tãũ] eles se **privam** [privaŋ] ou se **limitam** [li'mitaŋ] para fazer algumas coisas. [...] De aqui nasce o problema de: **exploração** [esplora'sãũ] ((A6 (34) m. 1999 S. ZL: 9.28).*
- (30) ***Não** [nãũ]. ((0,6s)) Mas bien quando viaje a Argentina sim ((0,2s)) sentí que **discriminaçam**<sup>121</sup> [diskrimina'sã] (A1 (23) m. 2006 SA Ce: 6.57).*

In further cases, the 3<sup>rd</sup> pers. pl. morpheme is realized with the velar nasal [aŋ] as in (29) or nasalized as [ã̃] in (30). The fact that the other nasal diphthongs in these examples are pronounced according to Brazilian Portuguese norm [ã̃] could also indicate that it is not transparent for my informants that the pronunciation of *-am* and *-ão* is the same in Brazilian Portuguese<sup>122</sup>.

Another good example for this phenomenon is the discourse marker *então* that is regularly realized as [en'tɔ̃n] by almost all speakers, obviously stemming from a mixed form of Bolivian Castilian *entonces* and Brazilian Portuguese *então*.

- (31) ***Então** [en'tɔ̃n] eu vou sempre viajando para esses lugares (A7 (32) f. 1994 S. Ce: 14.24).*

<sup>121</sup> Because of its syntactic position, I interpret this word as equivalent to *discriminam*, i.e. 3<sup>rd</sup> pers. pl. of *discriminar*. The verb *discriminaçar* is probably wrongly derived by the speaker from the noun *discriminação*.

<sup>122</sup> This interpretation would imply that the orthographic realization of the vocabulary influences the pronunciation, a thesis defended e.g. by Akerberg (2004).

- (32) *Então* [en'tɔ̃] eu: falei com meu cunhado *então* [en'tɔ̃] vamo vamo pra Brasil (A5 (28) f. 2000 D. ZL: 3.32).
- (33) *Então* [en'tɔ̃] desse lado:: um pouco de problema (A18 (30) m. 2004 P. Ce: 2.24).
- (34) *Então* [en'tɔ̃] a gente tá estudando essa parte para *não* [nɔ̃]... para se um dia cair né ((0,4s)) você se levanta rápido né. [...] *Então* [en'tõ] esse é o:: ideia né o idealismo. (A13 (33) m. 2000 D. ZL: 5.58)

Only in some cases, slight nasalization of [ɔ̃] resulting in the pronunciation [õ] is found, as in (34), [en'tõ].

For *não*, I mostly find the Spanish equivalent [nɔ̃], as for example in (24), (25) and (34). In doubled negations, though, – a construction very common in Brazilian Portuguese – the first *não*, positioned right before the verb form, is realized as [nɔ̃], but the second, more focused than the first one and positioned right after the verb, as [nãũ] (cf. (35) to (36))<sup>123</sup>. In (37) and (38) only the second *não* in post-verbal position is found, also a typical construction in Brazilian Portuguese. As these are also pronounced as [nãũ], syntactical factors obviously influence the pronunciation of the nasal diphthong.

- (35) *Falando sobre a ((0,6s)) especificamente sobre ((0,7s)) a imigração* [imigra'sãũ] do boliviano pra cá pro Brasil [...] existem motivos mas *não* [nɔ̃] é de *educação não* [nãũ] (A19 (42) m. 1988 P. Ce, S. 6.54).
- (36) *Não* [nɔ̃] tem problema *não* [nɔ̃] *não* [nãũ] (A23 (45) f. 1990 Ph ZO, S. 4.08).
- (37) *Eu consigo não* [nãũ] (A5 (28) f. 2000 D. ZL, S. 1.48).
- (38) *São de Cochabamba não* [nãũ] (A17 (31) m. 2007 P. Ce, S. 0.31).

Considering the whole group of my informants, the nasal diphthong clearly constitutes one of the phenomena that sets them apart from other *Paulistanos* because the pronunciation of this phoneme differs so strongly that it is highly salient for the receiving society. Among my informants, only some of the *donos* having spent many years in São Paulo and the younger informants working outside the textile industry pronounce it mostly according to the norm of Brazilian Portuguese. But for almost all of them, the strategy of leaning on the equivalent Spanish vocabulary and morphemes can be observed.

#### 6.2.4 Reduction vs. variation of o/u and e/i

The reduction of o/u and e/i is a further example in the context of variation in pronouncing the vowels by the Bolivian immigrants. In Brazilian Portuguese, the pronunciation of the vowels varies according to the position of the vowel in regard to primary stress. This means, that the unstressed pre-

<sup>123</sup> In Brazilian Portuguese, the two elements are also realized differently: the first as [nɔ̃], the second as [nãũ].

tonic syllable contains only /i/, /e/, /a/, /o/ and /u/ as in *feliz* [fɛ'lis] or *fogão* [fo'gãũ]; in final unstressed position, though, the vowel system is further reduced to /i/, /a/ and /u/, despite their graphic realization as <e> and <o>, for example *gelo* [ʒɛlu] or *sabe* ['sabi] (cf. Moraes, Callou and Leite 2002: 34 ff.).

CP, however, displays free allophonic variation between /i/ and /e/ as well as between /o/ and /u/ as in *mesa* ['misa] or *carro* ['kaRu] (cf. Rocha 1987: 10) and *prisión* [pre'son] and *suspiro* [sos'piro] (cf. Noll 2001: 25). However, this variation is rather socially stigmatized, as (39) underlines:

(39) *E eles do campo [...] eles falam 'qui ti pasa' quando eles podiam falar 'que te pasa' (A14 (28), m. 2006 S. ZN: 9.30).*

Additionally, in CP, the atonal final vowel is elided, particularly when the vowel stands between two voiceless consonants and follows an occlusive, e.g. *meses* [mes], *cartas* [karts] and *policía* [pol'sia] (Gordon 1980: 35).

Hence, the occurrences in my corpus could illuminate the influence of CP on my informants' speech. The analysis of this phenomenon, however, is complicated because many Castilian and Brazilian words only differ with regard to one vowel, for example *mismo* (Castilian) and *mesmo* (Brazilian), a fact also observed by one of my speakers:

(40) *Mas que português es quasí castelhano mal falado parece; porque aí os bolivianos que moram por os cantos falam como tipo português [...] eles falam 'mesmo' ['mezmu] e então aqui falam mismo ['mizmu] en la ciudad né (A5 (28) f. 2000 D. ZL: 6.39).*

Additionally, the reduction of the final atonal vowels in Brazilian Portuguese makes it impossible to decide if there is a case of vowel reduction regular in Brazilian Portuguese or o/u-variation according to CP (ex. (41)), or a case of non-reduction according to CC or o/u-variation (ex. (42)).

(41) *Os os cinco [us us 'sĩnku] dois filhos ['fiʎus] mais (A13 (33) m. 2000 D. ZL, S. 0.17).*

(42) *Eu solicitei para um vereador um político [po'litiko] para ver se [si] nos poderia ajudar [...] para conseguir um.: terreno [te'reno] para nós poder construir e ter um lugar próprio ['prɔprio] né (A21 (62) m. 1964 E. Ce: 2.37).*

Reduction in the pre-tonal syllables, e.g. *menino* [mi'ninu], is also described by Brazilian literature, but rather rarely and associated with spoken Brazilian Portuguese or socially indexicalized PP.

Thus, considering only pre-tonal vowels and vowels in stressed syllables, where the reduction in Brazilian Portuguese would rather not take place, I find the following examples:

- (43) *Vim com três filhas uno de quinze años uno de nueve años e [e]<sup>124</sup> ocho años (A25 (60) f. 1989 S/SA ZL: 0.23).*
- (44) *Tomara que **mude tudo** ['mode 'todo] isso aí (A18 (30) m. 2004 P. Ce: 11.45).*
- (45) *Aqui agarra você ((0,7s)) coloca contra parede **apontado** [apun'tado] dois armas (A14 (28) m. 2006 S. ZN: 3.56).*

While ['mode] and [e] seem to represent the free variation unambiguously, because in both languages the words e/y (“and”) as well as mudar (“to change”) are regularly pronounced as [i] and *mudar*, ['tudu] in (46) could be a case of hypercorrection, where the speaker applies Brazilian Portuguese vowel reduction not only to the second vowel in final position, but also to the first stressed vowel:

- (46) *Cê vai ver nas praca **todo** ['tudu] mundo estudando (A19 (42) m. 1988 P. Ce: 30.21).*

['todo] in (47) and (48), on the other hand, can also be interpreted as cases of insertion<sup>125</sup> where the speaker employs the word of her L1 instead of L2. Of course, the free variation of the vowels in CP could even cause my informants to have difficulties in realizing the phonemic distinction between e.g. *tudo* and *todo*.

- (47) *Tem de **tudo** ['todo] aqui no consultório (A16 (32) f. 2004 P. ZL: 3.28).*
- (48) *Tem brasileiro también tem boliviano también. [...] Assim **tudo** *misturado* ['todo ,mistu'rado] (A7 (32) f. 1994 S. Ce: 11.04).*

In addition to the phonological implications of this phenomenon, there are also morphological processes to be considered. In examples like (49) and (50), the variation of the vowel signifies a change of the verb morphology from 1<sup>st</sup> person to 3<sup>rd</sup> person sg.

- (49) *Eu **fui** [foi] pra San Mateus (A4 (21), m. 2006 S. ZL: 9.05).*
- (50) *Algumas coisas ((0,5s)) dentro da minha profissão **tive** ['tevi] que aprender (A23 (45) f. 1990 Ph ZO: 7.17).*

These examples could also be the manifestation of a morphological generalization of the 3<sup>rd</sup> person at the expense of the 1<sup>st</sup> person, continuing a process that is already present in PP.<sup>126</sup> However, the fact that I also find opposite examples (1<sup>st</sup>/2<sup>nd</sup> person instead of 3<sup>rd</sup> pers. sg.) as in (51) and (52) might indicate that it is really a phonological process and not necessarily a morphological one.

- (51) *E o filho da mãe né **fez** [fis] isso com nos (A6 (34), m. 1999 S. ZL: 6.44).*
- (52) *'Como ha sido Brasil? Ay, que **fez** [fis] você lá?' (A14 (28) m. 2006 S. ZN: 15.56).*

<sup>124</sup> It is possible that A25 is influenced by Portuguese orthography where [i] is written as {e}. Considering the pure Bolivian Castilian context of the quote *nueve años ocho años*, this seems less probable, though.

<sup>125</sup> See e.g. 6.2.16.

<sup>126</sup> For further information see 6.2.7 and 6.2.10.

Because of this variety of realizations and possible influential factors, there is no clear pattern recognizable. Some speakers – independent of age, gender or duration of stay – seem to prefer the reduction of *-o* to [u] and *-e* to [i] as typical for oral Brazilian Portuguese while others maintain the lower vowels in every position as typical for CC. Further examples rather indicate that free variation stemming from CP might also play its role for the interlanguage of the speakers.

### 6.2.5 Palatization of *-ti/di*

---

In Brazilian Portuguese, the realization of the dental occlusives represents an often discussed feature because when they occur before /i/ as in *dia* or *gente* their palatization is one of the most salient features for diatopic variation in Brazil. The palatization of *-ti/di* is closely connected with the reduction of *-e* to [i] in atonal position, because due to the latter the sound combination *-te/de* in atonal position is also palatized. According to a study of Abaurre and Pagotto (2002) the north eastern and the southern dialects of Brazil realize *cidade* as [si'dadi] in 100% of the occurrences, while in Rio de Janeiro 90% of the occurrences are pronounced as [si'daʒi] (cf. Abaurre and Gozze Pagotto 2002: 596). São Paulo is one of the regions where the palatization takes place, even though there is more non-palatized pronunciation than in Rio de Janeiro, probably due to the immigration of *Nordestinos* to São Paulo mentioned before. Further studies show that palatization is spreading even in southern rural areas according to the urbanization of the speakers (Battistiani and Dornelles Filho, Pires Lucas and Bovo 2007: 16). It can even be observed for the bilingual communities at the Uruguayan-Brazil border, where the palatization is estimated as a “linguistic stereotype of the urban monolingual Brazilian accent” (Carvalho 2003: 7). Therefore, speakers might consciously decide if they include the feature into their speech<sup>127</sup>:

*No português da televisão há muito [tʃ]. [...] Yo hablo portuguôol, não falo [tʃi]* (Carvalho 2004: 147).

Thus, the palatization of the dental occlusives is a highly indexical feature regarding not only the diatopic origin of the speaker, but also the social dimension of urbanization. In São Paulo, palatization is socially indexical for being *Paulistano* – in the quarters where many *Nordestinos* live, for not being a poor *Nordestino*. Consequently, it is socially rather positive to use this feature.

In Bolivian Castilian, the palatization of dental occlusive is a linguistic feature almost nonexistent in the phonetic inventory. Considering the occlusives, the reduction of intervocalic /d/ to [ð] is de-

---

<sup>127</sup> This is clearly a case of first-and-second-order social indexicality according to Silverstein 1992, as described before.

scribed as a regular characteristic, e.g. *metido* [me'tiðo]. In some regions though, particularly in the “Región de los Llanos del norte y del oriente”<sup>128</sup> (Coello Vila 1996: 28), the reduction goes further until the elision of the occlusive. This occurs mainly in intervocalic position, e.g. *medico* ['meiko], concerning mostly the past participles ending in -ido or -ado, e.g. *metido* [me'tio] or *salado* [sa'lao]. [ð] in word final position is also affected, e.g. *verdad* [ber'da] (cf. Rocha 1987: 14).

This means that in contrast to the variation just seen, where many factors possibly influence the realization of e/i/o/u in the different contexts, reliable patterns of palatized -*ti* and -*di* are clearly due to accommodation.

Interestingly, there are many examples in my corpus like (53) to (58):

- (53) *Á princípio era **difícil** [ʧi'fisil] é **difícil** [ʧi'fisil] aqui em São Paulo é muito **grande** [granʧi] la **gente** [ʒentʃi] es muy **diferente** [dife'rentʃi]* (A22 (27) f. 2008 P. ZS: 5.26).
- (54) *Mas tem que seguir **adelante** [ade'lantʃi]* (A24 (25) m. 2004 S/SA Ce: 5.35).
- (55) *Não **pode** [pɔʧi] sair daí mas eu fui a caminhar* (A14 (28) m. 2006 S. ZN: 3.42).
- (56) *Só también elas trazem pouquinhos nomás también nó: **monte** [montʃi] ne camionada* (A25 (60) f. 1989 S/SA ZL: 3.20).
- (57) *Minha **idade** [i'daʧi]? Minha **idade** [i'daʧi]é ... tenho **vinte** [vĩntʃi] anos* (A11 (20) m. 2002 S. Pe: 0.13).
- (58) *Um **monte** [mõntʃi] amigos brasileiros* (A10 (22) m. 2006 S. Pe: 9.48).

As the examples show, this phenomenon concerns mostly the last syllable, be it -*de* or -*te*. Regarding the palatization of the first syllable as in (53), I find fewer occurrences, but it can also be observed for almost all speakers. The preposition *de*, though, is mostly realized as [de] or [di] and not as [ʧi] as (59) and (60) illustrate. However, this is as it is common in São Paulo:

- (59) *Ah deve ter la **internet** [inter'neʃi]<sup>129</sup> **de** [de] lá **de** [de] Bolivia né. [...] **De** [de] Santa Cruz tem a **internet** [inter'neʃi]* (A1 (23) m. 2006 SA Ce: 5.03)?
- (60) *Eu vim **de** [de] la **cidade** [sidaʧi] **de** [di] La Paz* (A12 (39) m. 1990 D. ZL: 0.35).

Irrespective of the linguistic context, however, I observe a lot of variation concerning one speaker, as e.g. (61) and (62) show. While pronouncing *de* as [di] in (61), the same informant realizes *de* as [ʧi] a few seconds before, despite the similar linguistic context.

- (61) *La no estação **de** [di] trem mesmo nó do metro* (A11 (20) m. 2002 S. Pe: 11.56).
- (62) *Estação **de** [ʧi] trem São Miguel* (A11 (20) m. 2002 S. Pe: 11.49).

<sup>128</sup> As there are strong migrational movements inside Bolivia too, the spreading of single linguistic features from one region is an issue. Therefore, despite my focus on the Andean region, salient characteristics of other regions will be mentioned as well.

<sup>129</sup> Although *internet* has no regular -*te*, the pronunciation as [inter'neʃi] corresponds to Brazilian Portuguese because of its prosodic structure. It is a common phonological process to insert an epenthetic vowel [i] in order to avoid a consonant in the syllable coda, e.g. *advogado* [aʧivo'gadu].

This variation is also reflected in the following examples where it seems that the ending *-ente* as in *gente* and in adverbs like *imediatamente* is more frequently palatized than *-te* in other linguistic contexts as in *parte* or *neste*, although (65) and (66) show that even for *-ente/ende* the variation between [te] and [ti] as well as [[di] and [tʃi] is observable inside one phrase.

- (63) *Es parte* [parte] *de* [de] *La Paz seria Oruro Potosí e ((0,7s)) toda gente* [ʒentʃi] *daí* (A14 (28) m. 2006 S. ZN: 1.04).
- (64) *Mas neste* [neste] *caso é diferente* [ʧiferente] (A18 (30) m. 2004 P. Ce: 3.29).
- (65) *Então aqui você pode* [pode] *ter percebido a gente* [ʒente] *atende* [a'tendi] *imediatamente* [ime,djata'mêntʃi] ((0,5s)) *né nó nó precisa fazer éh agendamento* (A21 (62) m. 1964 E. Ce: 1.52).
- (66) *Entendo completamente* [kom'pleta,mente]. *Só que a gente* [ʒentʃi] *nó fala [...]* *Agora a minha mãe fala entende* [in'tendi] *e tudo* (A3 (34) f. 2004 SA ZL: 0.47).

As *a gente* [ʒentʃi] has become a personal pronoun in Brazilian Portuguese while the Castilian *la gente* [xente] has preserved its original semantics, it seems possible that the phonetic realizations of *a gente* differ according to the different semantics.

Thus, I looked at one speaker's realization of *a gente* with the Portuguese meaning (*we*) on the one hand (cf. (67), (69) and (71)) and with the Castilian meaning (*speaker excluded*) on the other hand (cf. (68), (70) and (72)):

- (67) *Acaba mi trabalho a gente* [xentʃi] *vai embora su casa para descansar* (A24 (25) m. 2004 S/SA Ce: 8.43).
- (68) *Porque a gente* [xentʃi] *quando tá bebada a gente* [xentʃi] *perde sentido né. Entonc a gente* [xentʃi] *sempre uno quer ser muito mais grande de outro ((0,8s)) pero mas eu nó* (A24 (25) m. 2004 S/SA Ce: 8.32).
- (69) *A gente* [ʃentʃi] *branca. ((1,0s)) Eles mexeram com tudo e agora eles querem nostro campo* (A14 (28) m. 2006 S. ZN: 1.45).
- (70) *Desculpa né mas ((0,4s)) es a diferencia né la gente* [ʒentʃi] *do campo nó entende muitas coisas* (A14 (28) m. 2006 S. ZN: 8.57).
- (71) *Como aqui é bonito a gente* [ʒentʃi] *gosta de aqui. (A16 (32) f. 2004 P. ZL: 6.18).*
- (72) *O que eles me f... éh ah muita gente* [ʒentʃi] *me fala que eu pareço mais assim peruana ou chilena* (A16 (32) f. 2004 P. ZL: 4.32).

As the examples clearly show, there are differences between the speakers concerning the palatization, but also the realization of the voiced sibilant at the beginning of the word. However, there is no pattern reflecting the different semantics; the realization of *gente* is quite uniform for the same speaker. In the case of different pronunciations as in (69) / (70) and (73) / (74), there is no connection detectable between phonetics and semantics: the semantics are the same for all occurrences (*speaker excluded*) but the realization differs strongly:

- (73) *Venho aqui para ajudar a la a la gente* [ʒentʃi] *boliviana* (A22 (27) f. 2008 P. ZS: 0.45).

- (74) *Entonces têm mais confiança com.: com gente ['xente] que fala espanhol [...] entonces éh ajudamos muito essa gente ['ʒentʃi] (A22 (27) f. 2008 P. ZS: 1.34).*

Only in (75), the phonetic representation seems to reflect the semantics, contrasting the first and third instance of *gente* signifying the Brazilian *we* – realized akin to Brazilian Portuguese norm<sup>130</sup> as ['ʒentʃi] with palatization – with the second excluding the speaker according to Castilian semantics – realized the Castilian way as ['xente].

- (75) *Sim, como aqui con colegas en la casa con colegas gente ['ʒentʃi] con gente ['xente] es nuestra: de nuestra terra con essa nomás a gente ['ʒentʃi] fala entón aí que nó: aprende bem. (A25 (60) f. 1989 S/SA ZL: 8.58)*

Additionally, there are examples of hypercorrection, i.e. sounds like *-to*, *-tem* or *-dem*<sup>131</sup> are palatized analogously, although in Brazilian Portuguese, the palatization would not be applied in these cases:

- (76) *Tengo muitas ['mujʃas] amizades (A24 (25) m. 2004 S/SA Ce: 7.04).*  
 (77) *Muito ['mujʃu] ((0,2s)) muita gente nó faz isso né ((0,3s)) pero faço porque tem que ganhar dinheiro né (A14 (28) m. 2006 S. ZN: 6.47).*  
 (78) *E.: eles acham porque eles nó têm a documentação podem ['pocʃi] ser explorados (A20 (27) f. 2006 P. Gu: 7.18).*  
 (79) *Eu ouvi que muitos éh se sentem ['sêntʃi] discriminados (A20 (27) f. 2006 P. Gu: 8.00).*

These hypercorrections indicate that the rule of palatization is rather consciously applied by the Bolivian informants to accommodate their speech toward the variety of Brazilian Portuguese spoken in São Paulo.

However, looking more closely, I observe a surprising division of the group: On one side, there are few occurrences for the informants having spent the least time in São Paulo and having least contact to Brazilians, but also for male academics working outside the textile industry. On the other side, particularly the young men who are working in the textile industry and the younger women working outside the textile industry, having an advanced level of Portuguese proficiency, palatize *-ti/di* almost every time and apply the palatization even to other phonetic contexts (cf. (76) to (79)).

## 6.2.6 Vibrants

In Brazilian Portuguese, the realization of the vibrants represents a highly variable and often discussed feature. Its allophones reveal first and foremost the diatopic origin of the speaker, but are also

<sup>130</sup> The norm would also require the nasalization of the vowel followed by the nasal consonant *n* resulting in ['ʒẽntʃi] or ['ʒẽntʃi].

<sup>131</sup> As I will show in 6.2.11, these occurrences could also be due to the absence of verbal agreement resulting in [...] *eles* [...] *pode* *ser explorados*, a verb form that is regularly palatized.

indexical of the speaker's social status and style. Depending on the position of the vibrant, Callou et al. (2002) enumerate in their quantitative survey seven different phones that are used by *Paulistanos*. The most common feature for the syllable onset in São Paulo is the aspirated [h], e.g. *rato* ['hatu]. Uvular [x] is more typical for Rio de Janeiro, while the single tap [r] or even [R] are sometimes heard by elder people with Italian migratory background, e.g. *guerra* ['gɛRa] (cf. Niehoff 2006: 39). In the syllable coda [r] as in *carne* ['karni] is found most often (87%), but sometimes also [∅] (5%: ['kani]), the retroflex [ɻ] (2%: ['kaɻni]) and rarely the uvular [x] (1%) ['kaxni] (cf. Callou and Moraes Leite 2002: 466 ff.) According to Callou, Moraes and Leite (2002: 467 ff.) the syllable-coda vibrant positioned in the middle of the word and at the end of the word have to be distinguished. In medial context, 87% are realized as [r] and only 2% as [∅]. In word final context, though, 49% of the occurrences are pronounced as [∅] and only 41% [r]. Beline Mendes (2009) has pointed out that

*each variant is popularly associated with particular social identities: [...] the tap is typically 'Paulistano', associated with cosmopolitan and educated speakers (l.c.).*

The retroflex allophone is called *r-caipira* and considered as rural and uneducated. Although there are no profound sociolinguistic studies about this subject, Beline Mendes (2009) hints at the possibility that the *r-caipira* has “some level of covert prestige in the city” because for example women favor this pronunciation. The fact that a famous Brazilian TV presenter adopted the [ɻ] seems to confirm this thesis (cf. Mendes and Matte 2008).

Another stigmatized feature of PP with regard to the vibrants is the *rhotacism* of the laterals. This means that /l/ is substituted by [r] in the syllable coda or in consonantal groups, for example *malvado* [mar'vadu] or *planta* ['prāta] (cf. Castilho 2010: 206).

In CP, multiple vibrants are often assibilated to [z] as in *perro* ['pezo]. According to Gordon (1980: 351) the assibilated realization is also rather prominent in consonant clusters with the vibrant following the occlusive /t/ as in *atrasado* [at<sup>s</sup>a'sadu] or *tres* [t<sup>s</sup>es]. Apart from that, the two dominant realizations of Bolivian Castilian are [r] and the multiple [R], which 74% of the speakers in the western areas of Bolivia realize (cf. l.c.). This is confirmed by Coello Vila's finding that the people around Santa Cruz show

*[p]lena realización de la vibrante múltiple /r/ en posición inicial, intervocálica o después de consonante nasal (Coello Vila 1996: 29).*

#### 6.2.6.1 Vibrants at syllable onset

---

Similar to the phenomena presented so far, there is much idiosyncratic variation to be observed in my corpus. In (80) and (81), for example, the same speaker realizes /r/ three times as [h] while realizing it as [x] some minutes before. A similar case is demonstrated by (82). The further examples show that even inside a single clause speakers vary between [h] and [r] in (84) and between [R] and [h] in (85).

- (80) *Nós descemos ((0,2)) eu só ví assim um caminho não uma **rodovia** [hodo'via] outra **rodovia** [hodo'via] [...] falei 'cadê as **ruas**' ['huas] (A18 (30) m. 2004 P. Ce: 15.44).*
- (81) *Ele éh.: tem muito **remorso** [xe'morsu] (A18 (30) m. 2004 P. Ce: 10.18).*
- (82) *Sábado es en Brás y en Bom **Retiro** [re'tjĩro] (A1 (23) m. 2006 SA: Ce: 6.10).*
- (83) *Tem outra feirinha também que é no Bom **Retiro** [he'tjĩro] (A1 (23) m. 2006 SA Ce: 5.53).*
- (84) *Quando **erro** ['ehu] alguma coisa eles me **corrigem** [kɔ'riʒẽĩ.] (A2 (18) f. 2005 SA ZL: 5.26).*
- (85) *E além disso o atendimento era muito **ruim** [Ruĩ] né, foi uma experiência **horible** [o'hible] né (A6 (34) m. 1999 S. ZL: 5.38).*

(85) is particularly interesting because a Castilian word (*horible*) is inserted – but with Portuguese phonetics, while [R], typically associated with Bolivian Castilian, is found in the Portuguese vocabulary *ruim*. Most studies on code mixing, however, claim that Portuguese phonology normally coincides with Portuguese vocabulary, while Spanish pronunciation is applied to Spanish words, as it is the case in (86). Here, the whole word *registrados* is realized in the Castilian way [r̄rexis'trados] and not [heʒiz'tradus]. (87) seems to confirm this assumption as the speaker starts to say the word *region* with the aspirated [h], but interrupts himself in order to repeat the whole Castilian word with Castilian pronunciation, as if she had realized that phonetics and vocabulary did not fit together.

- (86) *Temos pero nó está **registrado** [r̄rexis'trado]nó temos **registrados** [r̄rexis'trados] (A28 (28) m. 2002 S. ZL: 9.45).*
- (87) *Né então yo estoy de acuerdo então fazer uma coisa bien típica ((0,3s)) de nossa **reg** [heg] éh **región** [rex'jɔn] né que es a Bolívia. ((laughing)) (A23 (45) f. 1990 Ph ZO: 9.52)*

A form employed quite frequently by my informants is the simple alveolar vibrant [r] as (88) and (89) show. Additionally, there are a few cases of the multiple vibrant [R] as in (90) and also the assibilated form [z] in (91):

- (88) *También ven...vendía calçados [...] que llegava de la de **roupas** ['ropas] americanas (A26 (25) f. 2005 S. ZL: 5.40).*
- (89) *É o **regionalismo** [reʒiona'lizmo]. [...] Está **errado** [e'rado] porque Bolívia é um todo (A21 (62) m. 1964 E. Ce: 14.04).*
- (90) *Na **realidade** [r̄reali'daʒi] eu acho ah eu nó acho tão bom a Bolívia (A2 (18) f. 2005 SA ZL: 4.41).*
- (91) *Na Bolívia, quando a polícia te **agarra** [a'gaza] te dice ((0,3s)) 'que estás haciendo?' ((0,7s)) Aquí **agarra** [a'gaza] você ((0,7s)) coloca contra parede apontado dois armas (A14 (28) m. 2006 S. ZN: 3.56).*

The by far most spread and most used realization, though, is the aspirated form [h] typical of São Paulo's Portuguese:

- (92) [...] monte de serviço gente entregava unos quinhentos **reales recibía** [he'ales heci'bia] (A25 (60) f. 1989 S/SA ZL: 8.03).
- (93) Um **rapaz** [ha'pais] aí de **repente** [he'pẽnfi] aconteceu um problema e apareceu o **rapaz** [ha'pais] (A5 (28) f. 2000 D. ZL: 3.22).
- (94) Éh primeiro cheguei aí no **bairro** [bajho] da luz (A16 (32) f. 2004 P. ZL: 2.32).
- (95) Da minha **terra** [teha] (A17 (31) m. 2007 P. Ce: 9.00).
- (96) Se **morre** [mõhe] um boliviano num acidente alguma coisa assim ele **morre morreu** [mõhe mõ'reu] como um **cachorro** [ka'fõhu] aí (A9 (42) m. 1994 D. ZL: 18.19).

I associate this realization of /R/ with accommodation toward São Paulo's Portuguese. At a closer look, it stands out that akin to palatization, some of the speakers, particularly the women working as physicians, show very little variation and always employ the aspirated form, while other speakers – particularly seamstresses with little contact to Brazilians, but also some of the men working outside the textile industry – use [r] most of the time. These differences might point to the fact that there are two processes underlying the employment of [r]: L1-transfer in the case of the seamstresses; in the case of the academics, however, rather a form of non-accommodation of informants with high social status.

#### 6.2.6.2 Vibrants in syllable coda

In the syllable coda, there are very few occurrences of r-*caipira*:

- (97) E a:: ((0,4s)) a dificuldade de:: se **comunicar** [komuni'kaɪ] o idioma (A20 (27) f. 2006 P. Gu: 7.30).
- (98) Eu acho que a discriminação veio principalmente dos das pessoas ((0,5s)) que nó têm uma uma um: grado **escolar** [esko'laɪ] também ((0,7s)) razoável né (A20 (27) f. 2006 P. Gu: 8.30).
- (99) Elas também gostam de **falar** [fã'laɪ] só português elas não falam espanhol (A17 (31) m. 2007 P. Ce: 4.36).
- (100) O que o gosto de **fazer** [fã'zeɪ] é falar com as pessoas (A17 (31) m. 2007 P. Ce: 1.58).
- (101) Não pode sair daí mas eu fui a **caminar** [kami'naɪ] (A14 (28) m. 2006 S. ZN: 3.42).
- (102) Todo mundo vai **colocar** [kolo'kaɪ] (A14 (28) m. 2006 S. ZN: 6.43).
- (103) E también está fazendo curso de português para ver para **revalidar intentar revalidar** [revali'daɪ inten'taɪ ,revali'daɪ] (A22 (27) f. 2008 P. ZS: 4.11).
- (104) Entonces los pacientes que vem a **tratar-se** [tra'taɪse] tenemos esses medicamentos eles nó precisam comprar (A22 (27) f. 2008 P. ZS: 15.15).

Remarkably, although r-*caipira* is generally associated with the rural areas and is said to be socially stigmatized, only one of the four speakers using the r-*caipira* is a seamstress, the other ones are young academics. The speaker of (97) and (98), has already spent three years in Brazil and has

reached an extremely high level of Portuguese; hence, a low level of language proficiency is no valid explanation for these occurrences. On the contrary, - if the thesis of covert prestige of the r-caipira is accepted - the employment of this feature could rather point to accommodation toward a prestigious variety.

### 6.2.6.3 Rhotacism

---

Rhotacism, however, has no covert prestige but is described as a socially stigmatized feature associated with uneducated speakers of rural varieties. The few examples that are found in my corpus, however, show that seamstresses with a high level of proficiency in Portuguese – a *dono* and an academic – employ this feature:

- (105) *Mas a gente novinha que não pode ter ainda não tem a documentação certa que é isso que é o **problema** [pro'brema] que a maioria dos trabalhos ou requisito é o CPF ou RG que a gente não tem. Esse que é o **problema** [pro'brema] (A4 (21) m. 2006 S. ZL: 2.53/15.22).*
- (106) *Jogaram os copos **plásticos** [pras'fikus] e os **pratinhos** [pla'tɨjɨnus] aí jogaram (A7 (32) f. 1994 S. Ce, S. 9.47)<sup>132</sup>*
- (107) *Bom tenho muitos **planos** [pranos] (A13 (33) m. 2000 D. ZL: 4.00).*
- (108) *Árvores eu **plantei** [prantei] muitas (A21 (62) m. 1964 E. Ce: 15.39).*

This means that for the Bolivians who work and live in rather low social networks, the feature might have a different indexicality than for *Paulistanos*: First and foremost, it is important for the immigrants to speak like the others in their network and to be accepted as an “insider” of the network, a Brazilian, no matter if this feature is considered norm-distant or not.

### 6.2.7 Sibilants inside the word

---

The development of the sibilant system in the Iberian languages from Latin is one of the most complex processes described in diachronic studies. Today's consonantal inventories show that the Castilian system is reduced to the voiceless sibilants [s] and [ʃ], while Brazilian Portuguese maintained the opposition between voiced and voiceless sibilants [z] and [s] as well as [ʃ], and [ʒ] (Boller 2002: 197). This means, that e.g. in Brazilian Portuguese, *casa* ['kaza] and *caça* ['kasa] form a minimal pair, but in Castilian they are both pronounced ['kasa].

The voiced [z] only occurs in CP originating from the assibilation of the multiple vibrant as described before, e.g. *rato* ['zatu] or *carro* [kazu] (cf. Mendoza 2008: 221).

---

<sup>132</sup> In (106), I find not only the substitution of [l] by [r] in *plásticos*, but also the contrary process substituting [r] by [l] in *pratinhos*.

The realization of /s/ is rather complex in Bolivian Castilian. Gordon (1980), for example, describes strong diatopical variation: “El fonema fricativo /s/ tiene seis alófonos en Bolivia” (l.c.: 350), whereby [s], [z], [h] and [∅] are the most common ones. The latter are mainly used in the lower regions of Bolivia like Santa Cruz for implosive /s/: “‘eh que soy de Santa Cruh, pueh’” (Coello Vila 1996: 29.) In the higher regions around La Paz, Oruro and Potosí, 90% of Gordon’s informants used the [s] even for the implosive position. He claims that the pattern in Bolivia is highly unsystematic and dominated by individual variation:

*Más bien reinó la anarquía; un mismo informante pudo articular en un momento [s] y a los dos minutos, en la misma palabra, [h] o [∅] (Gordon 1980: 350).*

Additionally, in the Andean region, when the alveolar /s/ is followed by an /i/ it is palatized and realized as [ʃ] while the vowel is elided. This results in occurrences like *siempre* [ʃjẽmpre] or *canción* [kanʃon].

Therefore, the pronunciation of the sibilants inside the word can be revealing for L1-transfer:

In the corpus, the majority of the speakers realize the voiced sibilant /z/ and voiceless /s/ the same way, i.e. as [s]:

(109) *Casou* [ka'so] muy menina (A25 (60) f. 1989 S/SA ZL: 15.57).

(110) *Eu tinha quinze* [kĩnse] anos quando eu vim (A2 (18) f. 2005 SA ZL: 13.52).

(111) *Algumas coisas* [koĩsas] ((0,5s)) dentro de mia profissão teve que aprender (A23 (45) f. 1990 Ph ZO: 7.16/7.51).

(112) *Nós damos assim atenção personalizado* [personalisado] como se pagassem [pa'gasẽĩ] ((0,4s)) realmente uma consulta (A21 (62) m. 1964 E. Ce: 2.04).

The same phenomenon is found for the realization of /ʒ/ as /ʃ/, but only for one speaker:

(113) *As crianças me falaram que não se fala desse jeito* [ʃeĩtu] se fala do outro *jeito* [ʃeĩtu] (A2 (18) f. 2005 SA ZL: 2.20).

(114) *Ajuda* [a'juda] eles muito (A2 (18) f. 2005 SA ZL: 8.51).

This could indicate that – in accordance with the SLM (cf. 3.2.2) – the distinction between /ʒ/ and /ʃ/ is easier to perceive and realize than /z/ and /s/. However, as there are only few occurrences for the palatal sibilants and broad variation regarding its realization, e.g. as [ʧ] in (115) and [j] in (116), it is not possible to verify this thesis by my corpus.

(115) *Desse jeito* [ʧeĩto] (A14 (28) m. 2006 S. ZN: 5.18).

(116) *Eu vejo* [vejo] em outras comunidades assim os coreanos os árabes (A16 (32) f. 2004 P. ZL: 7.15).

While the correct comprehension of the meaning of (109) to (114) is nonetheless given, *casa* ['kaza] (“house”) and *caça* ['kasa] (“hunt”) as well as *casando* [ka'zãndu] (“marrying”) and *caçando*

[ka'sãndu] (“hunting”) are minimal pairs, i.e. the use of the voiceless sibilant here leads – in Portuguese – to different semantics of the phrase:

(117) *Homem tem que ajudar as coisas [koĩsas] da casa ['kasa] (A9 (42) m. 1994 D. ZL: 13.11).*

(118) *Depois de cinco anos elas acabaram casando [ka'sando] (A7 (32) f. 1994 S. Ce: 2.57).*

The realization of *casa*, of course, could also be explained as the employment of the Castilian word *casa*, because orthographically it is a full cognate.

The influence of the graphemes shown by Akerberg (2002) for the perception and realization of /z/ and /s/ by Spanish learners of Portuguese cannot be confirmed. According to my data, various Portuguese graphemes, namely {z}, {s}, {ç}, {c} or {ss}, are all realized as [s] by my Bolivian informants:

(119) *Fazendo [fã'sendu] a limpeza [lim'pesa] de casa ['kasa] criança [kriãnsa] tudo aquilo (A9 (42) m. 1994 D. ZL: 13.28)*

(120) *Mas tem que ver que se eu posso fazer [põsu fã'ser] uma coisa [koĩsa] como agora estou passando essas classes [pa'sandu ,esas 'klases] eu posso fazer [põsu fã'ser] melhor (A14 (28) m. 2006 S. ZN: 7.00).*

(121) *Depois vai fazer [fã'ser] a::: as leis brasileiras que::: documentos pedem para trabalhar certinho [ser'tʃĩnu] né (A13 (33) m. 2000 D. ZL: 5.23).*

Only in the speech of some younger speakers with academic education and one *dono*, /z/ occurs as [z]:

(122) *Não, mas eu acho que preciso sim [pre,sizo 'sĩ] (A20 (27) f. 2006 P. Gu: 5.06).*

(123) *Muita gente sabe que o estrangeiro ((0,6s)) tem muito mais capacidade para fazer [fã'zer] para desenvolver en la área (A17 (31) m. 2007 P. Ce: 1.13).*

(124) *Da segunda até a sexta a gente fica em casa ['kasa] trabalhando não sai muito né mas [...] a gente vai pra fazer [fã'zer] compras alguma coisa ['koĩza] importante né (A15 (28) m. 2000 D. Ce: 6.30).*

But even for them I find that the distinction between /s/ and /z/ is not continuously given, as shown by (124), where *casa* is again realized with the voiceless sibilant. Thus, this phenomenon concerns all speakers the same way, irrespective of social factors. It therefore is a clear case of L1-transfer, probably due to the only allophone character of the /s/-variation in Castilian, and the difficulties of perceiving sounds that are very similar to those the speakers already have in their repertoire as described by the SLM.

#### 6.2.8 Elision of final /s/ and absence of number agreement

The elision of final /s/ is known for rather norm-distant varieties in both countries. For CP, it is mostly described as one of the allophonic realizations of /s/, [Ø], and hence as a phonological pro-

cess. As plural in Castilian is mostly marked exactly by the final *-s*, this phonological process has also morphological consequences, but scholars do rarely describe it as leading to the absence of number agreement.

For PP, though, studies confirm that the elision of /s/ has to be discussed as a morphophonological process because it occurs mostly in the nominal syntagm when omitting the morpheme indicating the number agreement by final /s/.

Carvalho (2006: 162) confirms that the loss of nominal agreement in PP is attributed to a morphological and not a phonological process, because it concerns not only the dropping of the final /s/, but of the entire plural morpheme in words requiring irregular plurals, e.g. *-es* as in *mulher – mulheres* or *-ões in conexão – conexões*, even though the “plural stripping” (cf. l.c.) is more typical for the regular forms:

(125) *Leva o almoço para minhas irmãØ que trabalhavam nas fábricaØ e os meuØ irmãoØ* (Niehoff 2007: 47)

(126) *Vieram meu pai, minha mãe e os dois irmãoØ* (l.c.: 66).

Different studies (cf. e.g. Pereira Scherre 1991, 1998 and 2001) have shown that phonological factors like stress (if the final syllable of the singular form is stressed and the singular/plural opposition contains more material than final *-s*, the overt plural marker on all elements of the phrase is more probable), as well as further structural factors like the position in the nominal syntagm and its grammatical class also play an important role for the variation. Castilho (2010) states that in PP, there is only the

*[m]anutenção da concordância quando há saliência fônica entre a forma do singular e a forma do plural* (l.c.: 208).

In contrast to the unsystematic absence of number agreement in bilinguals’ CP, in PP the phenomenon indicates a simplifying pattern<sup>133</sup>: Instead of realizing the plural marker on all the units, as required by the norm, speakers reduce this redundant indication of plural to only one marker placed on the first element of the phrase, typically the determiner, as follows from the examples above.

This pattern is also valid for *Fronterizo* as Carvalho (2006) shows:

*[T]he application of a Portuguese rule pushes toward the loss of nominal agreement already common in Spanish, primarily caused by the loss of the /s/ phoneme* (l.c.: 163).

---

<sup>133</sup> Pereira Scherre (2001) and Pereira Scherre and Naro (1991) emphasize that “the variation in noun phrase plural agreement [...] cannot be explained by the principle of linguistic economy [...]” (Pereira Scherre 2001: 469), because “marking tends to occur precisely in those contexts in which it is most highly redundant and could therefore be discarded with no loss of information” (Pereira Scherre and Naro 1991: 23). This fact is explained by a “universal tendency to repeat similar grammatical forms” (Pereira Scherre 2001: 466).

In many cases, it is not possible to unambiguously identify if the missing /s/ in the speech of the Bolivians is an expression of phonological elision as it occurs in CP or the morphological process frequent in PP. I therefore present the examples that are clearly phonological first.

To be unambiguously identified as the result of a phonological process, the final /s/ may not represent an entire morpheme. This means that the examples concern either word classes other than nouns as in (127) to (129), or the nouns in question require a different plural morpheme, for example *-es* as in (130) and (131).

(127) *Nomás* [no'maØ], marido e irmão dele (A3 (34) f. 2004 SA ZL: 6.31).

(128) *Passamos as cosas a menor que deseamos* [de'samoØ] né *mas* [maiØ] a gente de de ciudad nunca se vai deixar (A14 (28) m. 2006 S. ZN: 2.11).

(129) *Depois* [depoiØ] vai fazer a::: as leis brasileiras que::: documentos pedem para trabalhar certinho né (A13 (33) m. 2000 D. ZL: 5.23).

(130) *Éh::: o curso são dois vezes* [veseØ] a semana [...] de quarta e::: terça e quarta feira (A22 (27) f. 2008 P. ZS: 6.03).

(131) *Só dois* [doiØ] cinco seis *meses* [meseØ] [...] (A9 (42) m. 1994 D. ZL: 4.59).

(132) *Con ese::: miedo ... por eso que as vezes* [veseØ] *nosotros* [no'sotroØ] só *moramos* [mo'ramoØ] dois *anos* [anoØ] depois vamos embora depois outro volta va volta va así (A1 (23) m. 2006 SA Ce: 2.46).

In (132), the first three phenomena are identified unambiguously as due to the phonological process; the fourth could also be morphologically conditioned. Because of the linguistic context of continued elision of final /s/ before, this example also seems to point in the phonological direction, though.

In addition, I regard those examples as phonological ones where the final /s/ indeed constitutes the plural morpheme but the elision concerns the first element of the nominal syntagm. For PP, many studies have shown that there are complex rules commanding the probability of the absence of number agreement, which elements of the syntagm must have the plural marker, which need not, etc. Recent studies agree, however, that

*a primeira posição do SN (sintagma nominal, sn) é a mais marcada e as demais posições mostram um índice baixo de marcas (Souza Campos and Rodrigues 2002: 111).<sup>134</sup>*

<sup>134</sup> The functionalist approach explains this fact by the redundancy of the plural markers that can be left out when the first element is already marked. Other approaches claim that this explanation is not enough, because not only the first position, but all “elements occurring to the left of the head are marked more frequently than the elements to the right of the head. For the head itself, the left most category also favors overt plural marking and the other positions disfavor it” (cf. e.g. Pereira Scherre 2001: 4). For my aims, though, the finding that the marking of the first element(s) is more probable than its non-marking, combined with the marking of latter elements that most theoretical approaches agree on, is sufficient.

Thus, it is very rare for the overt plural marker /s/ to occur at the second or third element while the first element is unmarked. As I find this combination in my corpus, (133) to (137) are also considered examples of the phonological elision of -s.

(133) *As* [aØ] *mulheres são muito bom* (A8 (27) f. 2005 R. Ce: 7.04).

(134) *Aqui em essa pequena parte de São Paulo donde mora muito boliviano yo m... ((0,3s)) yo oyo que hay ah às* [aØ] *vezes muita discriminação* (A22 (27) f. 2008 P. ZS: 19.27) .

(135) *Nao tem muitos* ['muʃtuØ] *bolivianos aqui igual o Brás* (A11 (20) m. 2002 S. Pe: 14.12).

(136) *Deveriam ser mais unidos ter muitas* ['muʃtaØ] *coisas os bolivianos* [,boli'vjanoØ] (A9 (42) m.1994 D. ZL: 20.45).

(137) *Eu tenho muito poucos* ['pokuØ] *amigos bolivianos* (A2 (18) f. 2005 SA ZL: 5.14).

Of course, these examples could also represent cases of hypercorrection, meaning in this case that language learners generalize the PP rule “absence of number agreement” and apply it to all elements of the syntagm without having realized the syntactic structure underlying it. This would explain why in (136), which contains various nominal syntagms in one sentence, the elision of final /s/ affects different elements of the syntagm: Regarding the first syntagm I find no final /s/ on the first element *muita*, but on the second *coisas* I do; in the second nominal syntagm, though, the first element *os* receives the final /s/, but the second element *bolivianos* does not. In CP, the absence of number agreement is found at all elements in the nominal syntagm without this regular pattern. Thus, the examples could also be due to structural transfer from CP.

The realization of *poucos* as [pokuØ] in (137) is ambiguous, but not with regard to the plural morpheme, but because it could also be attributed to the misinterpretation of *pouco* as an adverb in analogy to *muito* right before it. The fact that both words are employed one after another makes this interpretation rather plausible.

Rather ambiguous cases are represented by (138) to (143), because the nominal agreement only depends on the final /s/. This means that it could be either structural L1-transfer or accommodation toward PP.

(138) *Maioriamente com as criança*Ø (A2 (18) f. 2005 SA ZL: 2.13).

(139) *Ele se enriqueceu né aos recurso*Ø *de nos* (A6 (34) m. 1999 S. ZL: 6.44).

(140) *Aí eu fiquei lá:: aprendí a mexer com as máquina*Ø *né* (A12 (39) m. 1990 D. ZL: 1.29).

(141) *Então o que falo pra eles nó se conformar com aqueles trabalho*Ø (A9 (42) m. 1994 D. ZL: 4.56).

(142) *Aquele colorido que têm os índio*Ø (A23 (45) f. 1990 Ph ZO: 9.17).

(143) *Cê vai ver nas praça*Ø *todo mundo estudando* (A19 (42) m. 1988 P. Ce: 30.21).

It is easy to see that all the examples work according to the same pattern. In the linear nominal-phrase, the first element, mostly the article, is marked with final /s/, while the second one receives no

final /s/. Similar cases represent the phrases where the plural is marked by a quantifier preceding the unmarked noun:

- (144) *Yo ya trabajé **dós año**Ø um año he trabajado (A24 (25) m. 2004 S/SA Ce: 2.32)<sup>135</sup>*  
(145) *A gente só é **três irmão**Ø (A11 (20) m. 2002 S. Pe, S. 9.24).*  
(146) *Depois fiquei quase **dez ano**Ø no Bom Retiro (A7 (32) f. 1994 S. Ce: 1.51).*  
(147) *Imagina fazer o documento em **quinze dia**Ø (A9 (42) m. 1994 D. ZL: 2.01).*  
(148) *Van van ser já quatro anos. [...] **Quatro ano**Ø é (A16 (32) f. 2004 P. ZL: 0.18).*

For more complex structures, I observe that there is number agreement regarding the first elements of the syntagm (left of the head), while the last element(s) (right of the head) is (are) unmarked.

- (149) *Mas é um pouco complicado porque se exigem **as quatro oito cursos traduzido**Ø e **legalizado**Ø (A17 (31) m. 2007 P. Ce: 3.32).*  
(150) *A minha esposa fazia **uns bolos gostoso**Ø (A12 (39) m. 1990 D. ZL: 6.07).*

The regularity of this pattern might already indicate that the process underlying the absence of nominal agreement in the speech of my informants is not alone motivated by the Castilian elision of final /s/.

This finding is confirmed by cases where the plural morpheme contains more phonic material, like *-es* employed for the plural marking of nouns ending on [s], or *-ões* which is obligatory for most nouns ending on *-ão*, because these occurrences constitute an unambiguous indicator of plural stripping. As pointed out above, in PP, the absence of plural marking with these nouns is less common, because of the phonic saliency of the markers (cf. e.g. Carvalho 2006: 163 or Castilho Souza Campos and Rodrigues 2002: 108 f.). Accordingly, my corpus shows only two examples for *-ão/ões*:

- (151) *Como eu não tinha **essas informação** (A6 (34) m. 1999 S. ZL: 31.14).*  
(152) *Como era dois era **dois portão** (A11 (20) m. 2002 S. Pe: 18.05).*

Regarding *-es*, though, my corpus contains many examples showing the deletion of the complete syllable:

- (153) *Ya faz quase um ano **cinco mês**Ø **seis mês**Ø aqui (A4 (21) m. 2006 S. ZL: 3.15).*  
(154) *Todo mundo chega come **às vez**Ø brinca [...] ah sim **às vez**Ø, **às vez**Ø elas saem (A8 (27) f. 2005 R. Ce: 4.54).*  
(155) *Vou ficar um mês **dois mês** despues eu volto (A5 (28) f. 2000 D. ZL: 3.59).*  
(156) *Porque só foram **quatro mês** só corrido né para a USP (A6 (34) m. 1999 S. ZL: 19.09).*  
(157) *Eu estava mais ou menos **quatro mês** acho que **oito mês** por aí no Brás (A7 (32) f. 1994 S. Ce: 1.40).*  
(158) *Eu brigo **muitas vez**Ø com ele (A9 (42) m. 1994 D. ZL: 11.56).*

---

<sup>135</sup> As this example is in Castilian, the deletion of final /s/ is probably motivated by the phonological process.

(159) *Sim olhei que foram discriminados só **unas vez**Ø aqui en una parada de ônibus (A22 (27) f. 2008 P. ZS: 20.58).*

(160) *Tenho uma filha de:: seis e uma de um ano e **cinco mês** (A18 (30) m. 2004 P. Ce: 6.55).*

(161) *Aí eu mh:: só vim passar férias por **três mês** mas acabei ficando (A19 (42) m. 1988 P. Ce: 0.54).*

At first sight, these examples corroborate the finding that the absence of plural agreement is due to the morphological process of plural stripping.

The absence of plural agreement is widely spread among the informants. Just three speakers, two academics and one *dono*, always apply the plural markers in the nominal syntagm. The two speakers of (151) and (152), in contrast, both seamstresses in São Paulo's periphery, almost seem to follow the rule "no number agreement". In (162), for example, there is the self-"correction" of the speaker, changing his sentence from PC-conforming *meus irmãos* to PP-conforming *meus irmão*Ø when the interviewer did not understand him:

(162) - (Ni): *Eu tinha minha oficina lá com **meus irmãos**. - (I): Como? - (Ni) Com **meus irmão**Ø (A11 (20) m. 2002 S. Pe: 2.59).*

It seems, thus, as though the PP rule pushing toward the absence of number agreement is easily accepted. This could be explained by various processes that might even all work together: Accommodation toward rather norm-distant Portuguese, simplification during language learning, and also L1-transfer: the phonological process of final /s/ deletion or the morphological reduction of nominal agreement caused by the influence of the indigenous languages on CP.

### 6.2.9 Gender attribution and absence of gender agreement

---

In CP, apart from the absences of number agreement shown above, the absence of gender agreement occurs as well (Mendoza 2008: 226):

(163) *Nunca no se arregla **esa** problema.*

(164) *Por **ese** razón estamos seguros.*

(165) *Está **controlando** toda esta situaciones.*

(166) *Incluso las plazas son muy **hermoso** por allí.*

The examples show an irregular pattern where either the feminine or the masculine gender, or either the singular or the plural form is favored. This phenomenon is attributed to the absence of gender (and number agreement) in the indigenous languages. Consequently, Mendoza (l.c.) confirms that these are interlanguage phenomena of bilingual speakers:

*Esta discordancia tiene character de rasgo transitorio y está presente más que todo en el habla de bilingües. Por tanto es inexistente en el habla de monolingües (Mendoza 2008: 226).*

As the attribution of a different gender and the absence of gender agreement are not described for Brazilian Portuguese, I associate this with the influence of a rather norm-distant PC heavily influenced by the indigenous languages. At a closer look, though, I can differentiate linguistic contexts that foster gender agreement and its absence without staying in direct contact with the influence of Aymara or Quechua.

The most obvious influence of Castilian are *false gender cognates*, i.e. nouns that exist in both languages and have the same semantics, but require different gender in the two languages. These examples are rare in my corpus, but in (167) and (168), I find *a terminal* and *o passagem*, obviously cases of direct transfer from Castilian *la terminal* and *el pasaje* that result in the – from the Brazilian view incorrect – attribution of gender in Portuguese, where *terminal* is masculine and *passagem* feminine.

(167) *Mas eu cheguei lá cheguei à terminal e meu deus no céu não gostei nada (A18 (30) m. 2004 P. Ce, 15.08).*

(168) *Además do que faz dua dois mês que tem que trabalhar de graça para pagar o passagem (A4 (21) m. 2006 S. ZL, S. 18.13).*

(169) and (170) must also be considered separately because the absence of gender agreement in the linguistic context of *dois* is quite frequent. Instead of associating these examples with the indigenous languages, the use of only the masculine form *dois* even in connection with a feminine noun and not vice-versa points to the fact that while in Portuguese the number receives gender marking, in Castilian *dos* – the equivalent to Portuguese *dois/duas* – is indeclinable.

(169) *Éh:: os cursos são dois vezes à semana [...] de quarta e:: terça e quarta feira (A22 (27) f. 2008 P. ZS: 6.03).*

(170) *Porque aqui têm digamos dois costuras né (A14 (28) m. 2006 S. ZN: 6.37).*

Similarly, there are many cases of absent gender agreement in nominal syntagms containing a possessive adjective. As the singular forms in Castilian are also indeclinable, but receive gender marking in Portuguese, the variation regarding these syntagms seems rather probable. In contrast to *dois*, though, there is not only the masculine singular form in connection with feminine nouns as in (173) and (174), but also feminine possessives together with masculine nouns as in (171) and (172).

In the case of (171) and (174), the choice of the feminine or masculine possessive despite the masculine *nome* and the feminine *oficina* could reflect the semantics that they as a woman (171) and a man (174) are in the possessing position, emphasizing this aspect at the expense of the morphological agreement. This interpretation is further confirmed by the again not agreeing *próprio* in (174) and *próprio* in (175), both masculine forms used to emphasize the role of the speaker as owner, although

even the natural sex of *pareja* (Castilian) and *namorada* (Portuguese) clearly requires the feminine morpheme.<sup>136</sup>

- (171) *Éh **mia** nome é Nancy Nancy Limatol (A26 (25) f. 2005 S. ZL: 0.06).*  
 (172) *Eu desistí porque **mia** gente **mia** mãe **mia** pai viram que: ((0,6s)) solo mio nó era estudar né era fazer ((1,0s)) desordem (A14 (28) m. 2006 S. ZN: 8.38).*  
 (173) *Mas éh.: donde yo moro allá en Santo Amaro ((0,3s)) éh.: nó tem muito boliviano entonces eu **meu** família yo acho que es minha família nó yo acho que es assim en essa setor únicos bolivianos que moram lá (A22 (27) f. 2008 P. ZS: 19.45).*  
 (174) *Como eu já tive **meu oficina próprio** [...] (A11 (20) m. 2002 S. Pe: 7.24).*  
 (175) *Eso me há revelado [...] mi aislamiento de mí con mi **propio pareja** con **mi propio namorada** (A24 (25) m. 2004 S/SA Ce: 14.07).*

So far, similar examples are also discussed for Spanish learners of Portuguese whose native variety does not contain the absence of gender agreement (cf. Akerberg 2002: 24). In my corpus, though, the absence of gender agreement and *incorrect* gender attribution are found in further linguistic contexts that can not be explained by the characteristics of their Castilian counterpart:

- (176) *La **no estação** de trem mesmo nó do metrô (A11 (20) m. 2002 S. Pe: 11.56).*  
 (177) *Nós damos assim **atenção personalizado** como se pagassem ((0,4s)) realmente uma consulta (A21 (62) m. 1964 E. Ce: 2.04).*  
 (178) *Ele faz o desconto de acordo como eles querem me pagar ((0,9s)) mensal em **quantas** meses (A9 (42) m. 1994 D. ZL: 4.26).*  
 (179) ***Nos ONGs** por exemplo (A9 (42) m. 1994 D. ZL: 3.32).<sup>137</sup>*

On the contrary, the corresponding Castilian nouns *estación*, *acentuación* and *mes* carry the same gender as in Portuguese. Thus, transfer from Castilian is obviously not the only process underlying the speech of the Bolivians, but the actual weakening of the noun's gender qualities. In the examples above, the endings *-ão* and *-es* as well as the acronym are rather opaque with regard to their gender, maybe a complicating factor.

Correspondingly, the group of masculine nouns ending on *-a* is often found in combination with feminine articles and adjectives, as the following examples show:

- (180) *No conozco yo casi **muchas idiomas** (A1 (23) m. 2006 SA Ce: 0.59).*  
 (181) *E esse é **a problema** do boliviano (A9 (42) m. 1994 D. ZL: 11.42).*  
 (182) *Não tive **problema nenhuma** (A13 (33) m. 2000 D. ZL: 1.48).*  
 (183) *E eles às vezes têm **muitas problemas** têm **muitas problemas** (A22 (27) f. 2008 P. ZS: 7.48).*

<sup>136</sup> In most sentences containing a possessive adjective, the Castilian forms *mi* and *su* are used:

- a) Las filhas eran pequeninas eran (?) pequenina como su filha (A25 (60) f. 1989 S/SA ZL: 0.18).  
 b) Mis melhores amigos ((0,5s)) pra te falar verdade são os brasileiros (A13 (33) m. 2000 D. ZL: 2.42).

<sup>137</sup> The acronym ONG ['ɔŋi] stands for *organização não-governamental*, a fact the speaker actually might be unaware of.

It seems, thus, as though the attribution of gender according to Castilian norms is actually somewhat weakened in CP, particularly because (180), a completely Castilian sentence, contains the same phenomenon. Therefore, the finding that the absence of gender agreement also occurs in nominal syntagms where the feminine elements end on *-a* while the masculine ones end on *-o*, is not surprising:

(184) *A bebida alcoólico* (A24 (25) m. 2004 S/SA Ce: 5.25).

(185) [*O espanhol é q]uase **uma língua completo** de você tem que pronunciar mesmo ((0,8s)) todas as palavras certas (A18 (30) m. 2004 P. Ce: 4.08)*

Interestingly, in (184) and (185), the syntagms are divided in terms of the gender. While the first elements, the article and noun, show agreement, the following adjective gets no feminine marker. Whether the obvious parallelism between this structure and PP's rule regarding the absence of nominal agreement discussed before is due to similar underlying processes (e.g. the reduction of the redundant gender marking on further elements or the non-marking of elements right to the head of the NP) cannot be answered in this study. As I know of no studies exploring the absence of gender agreement as a regular structure, it remains a desideratum for further research.

Additional cases underscoring the fact that gender is a weakened category in the Portuguese of Bolivian immigrants are (186) and (187). Here, two men refer to themselves by the feminine adjectives *sozinha* and *tranquila*.

(186) *Agora se eu ficar ou estiver **sozinha** ah já não* (A18 (30) m. 2004 P. Ce, 14.27).

(187) *Tô **tranquila** nó pois é claro que o estrangeiro é diferente mas é ((1,1s)) assim* (A15 (28) m. 2000 D. Ce, 8.21).

In generalizing expressions, the absence of gender agreement is also widely spread. A reason for this finding could also be that the natural sex of the persons the speaker is referring to, overrules the morphological gender. In (188), for example, the male informant speaks about himself being *brasileiro*; for (189), though, the context shows no clear female reference explaining the use of *refinada*.

(188) *A gente ya mora mais que vinte anos a gente ((0,3s)) queira ou não **a gente é brasileiro** também né* (A19 (42) m. 1988 P. Ce, S. 15.07).

(189) *É **o pessoal** mais **refinada** éh sabe* (A17 (31) m. 2007 P. Ce, S. 2.56).

(190) also points toward the same direction as the previous examples, but at the same time it shows that for anaphoric constructions the absence of gender agreement is often observable. The so called *loísmo* – rather common in Castilian varieties – denotes the phenomenon that speakers employ the clitic *lo* referring even to an indirect object or a feminine direct object instead of the clitics *le* and *la* required by standard variety. While (190) could be attributed to *loísmo*, (191) and (192) show that

this feature is also found for different anaphoric contexts – even when natural sex and morphological rule coincide as in (191).

(190) *Es como discriminación sería usar a **una** persona y éh no pagarlo bien de no tratarlo bien (A24 (25) m. 2004 S/SA Ce: 9.31).*

(191) *Vim com **três filhas uno** de quinze años **uno** de nueve años éh ocho años (A25 (60) f. 1989 S/SA ZL.: 0.23).*

(192) *Nosso **carnaval** é muito diferente (?) se ((0,4s)) colocar **os dois** você vai ver que nó tem nada ver **uma** com **a outra** né (A19 (42) m. 1988 P. Ce: 12.30).*

Additionally, in (192) the absence of gender agreement not only occurs in nominal syntagms with feminine nouns and the rather less marked masculine article or morpheme *lo* or *-o* , but also vice-versa in syntagms with masculine nouns and feminine anaphors:

In another linguistic context, though, the absence of gender agreement and also of number agreement is due to simplification, i.e. the masculine singular form is employed in any case of predicative adjectives, even if referring to feminine or plural subjects:

No number agreement:

(193) *Então **homem e mulher** têm que ser **igual** (A9 (42) m. 1994 D. ZL:13.07).*

(194) *[Eles] são bem **alto** (A23 (45) f. 1990 Ph ZO: 6.02).*

No gender agreement:

(195) *A gente ya mora mais que vinte anos a gente ((0,3s)) **queira** ou não **a gente** é **brasileiro** também né (A19 (42) m. 1988 P. Ce, S. 15.07).*

(196) *A **situação** tá **complicado** (A2 (18) f. 2005 SA ZL: 8.40).*

(197) *É que **mão-de-obra** é **barato** (A19 (42) m. 1988 P. Ce: 7.52).*

(198) *As **mulheres** são muito **bom**. (A8 (27) f. 2005 R. Ce: 7.04).*

Thus, it seems as though the predicative adjectives are analyzed and reinterpreted as indeclinable adverbs, resulting in syntagms where either gender or number agreement is absent. In norm-distant PP, the examples of no number agreement (cf. (193) and (194)) are acceptable. Even (195) to (197) could be heard by native Brazilians, as – in context with *é* or *tá* –, the adjectives are also interpreted as indeclinable adverbs. (197) and (198) though, could not be accepted even in most norm distant varieties of PP, i.e. at the norm-distant end of the PP-PC continuum. Particularly revealing is (198), where both agreement rules are violated by the predicative adjective *bom*.

This phenomenon is extremely far spread among my informants. Most of my informants show occurrences of *incorrect* gender attribution and no gender agreement. The only exceptions are women working as physicians outside the textile industry. In their speech, I find very few or no examples at all.

Hence, the general variation regarding number and gender agreement cannot be attributed unambiguously, either to L1- transfer, or to accommodation. Instead, various factors come together here: Undoubtedly, there is a certain influence of the weakening of gender attribution and agreement rules in CP, but there is also the weakening in particular of number agreement rules in PP. These factors and simplification processes common in the course of language learning add to a system of interlanguages among the Bolivian immigrants where the agreement in nominal syntagms is not obligatory anymore.

#### 6.2.10 Personal pronouns and verbal paradigm

The pronominal system of the varieties varies considerably due to the changes the pronominal system of PP has undergone in the last decades, resulting in

*fortes consequências na estrutura sintática da lingual. [...] [A] reorganização do quadro dos pronomes repercute nos demais pronomes, na morfologia verbal, na concordância verbal e na estrutura funcional da sentença (Castilho 2010: 477).*

In this study, I concentrate on the subject pronouns as they are rather revealing for L1-transfer or accommodation, because the varieties in contact differ strongly in this regard.<sup>138</sup> In Bolivian Castilian, there is the following pronominal system:

Table 6.3: Pronominal system of Bolivian Castilian (cf. Fontanella de Weinberg 2000: 1404 f.)

Person	Subject	Complement
1 <sup>st</sup> pers. sg.	Yo	me, mi
2 <sup>nd</sup> pers. sg.	tú/vos; usted	te, ti, vos
3 <sup>rd</sup> pers. sg.	él/ella	le, lo/la, se, si
1 <sup>st</sup> pers. pl.	nosotros/as	nos
2 <sup>nd</sup> pers. pl.	ustedes	os/as, los/las
3 <sup>rd</sup> pers. pl.	ellos/ellas;	les, los/las, se, si

According to Fontanella de Weinberg (2000: 1405), in Bolivia, *voseo* and *tuteo* alternate. This means that depending on trust and intimacy between speaker and hearer, either *tú* or *vos* are used, both either with the original verb morphology of the 2<sup>nd</sup> ps. sg. or of the 2<sup>nd</sup> ps. pl., e.g. *tú tienes* or *tú tenís* and *vos tienes* or *vos tenís*. Placencia (2001) shows that this distinction is also employed for social discrimination: Studying the treatment of indigenous speakers in public institutions, she notic-

<sup>138</sup> A thorough study of the complements employed by the Bolivians cannot be conducted here, but remains a strong desideratum.

es that indigenous Bolivians are addressed by *tú* and *vos* despite the formal, little intimate situation, while people of other ethnics are treated by *usted*.

Table 6.4: Pronominal systems of Brazilian Portuguese (formal and informal) (cf. Castilho 2010: 477)<sup>139</sup>

	Formal Brazilian Portuguese		Informal Brazilian Portuguese	
Person	Subject	Complement	Subject	Complement
1 <sup>st</sup> pers. sg.	eu	me, mim	eu, a gente	eu, me, mim
2 <sup>nd</sup> pers. sg.	tu, você, o/a senhor/a	te, ti	tu, você, (o)cê	(v)ocê, te, ti
3 <sup>rd</sup> pers. sg.	ele/ela	o/a, lhe, se, si	ele/ela, ei, Ø	ele/ela, lhe, Ø
1 <sup>st</sup> pers. pl.	nós	Nos	a gente	a gente
2 <sup>nd</sup> pers. pl.	vós, os/as senhores/as	Vos	vocês/ocês/cês	vocês/ocês
3 <sup>rd</sup> pers. pl.	eles/elas	os/as; lhes, se, si	eles/elas, eis	eles/elas, eis

Table 6.4 gives an overview over all varieties of Brazilian Portuguese without considering the diatopic particularities of São Paulo. In this general picture, the differences between the two varieties primarily concern the 1<sup>st</sup> pers. pl. The formal *nós* is more and more replaced by *a gente*. Ilari et al. (2002) even show that *a gente* is used more often than *nós* by *Paulistanos* taking part in NURC (l.c.: 81):

*No geral [...] a variedade culta mostrou selecionar com mais frequência o sujeito nós, mas a gente aparece como um seríssimo concorrente (Ilari et al. 2002: 91).*

It is particularly interesting that by this change, the semantic of *a gente* (Portuguese) and *la gente* (Castilian) became different as the sentences *a gente come muita carne* (Portuguese) and *la gente come mucha carne* (Castilian) show: While the first sentence includes the speaker as meat-eating, the second one expresses a generalization excluding the speaker. To express this semantics in Brazilian Portuguese, instead of *a gente* other words like *o pessoal* or *o povo* are used: *O pessoal/o povo come muita carne*.

Additionally, the forms of the 2<sup>nd</sup> pers. given for formal Portuguese are very rare in PC, the formal variety of São Paulo. The use of *vos* is restricted to ritualized communicative situations as i.e. in church it was replaced by *vocês* several decades ago:

<sup>139</sup> Again, I want to emphasize that this does not mean that in formal situations, speakers never use e.g. *vocês*. The distinction made in this table, only resorts to the **theoretical concept** of two contrasting varieties. This explains why the NURC data differ from Castilho's (2010) table: As NURC is about oral speech, not every parameter of communicative distance is fulfilled. Therefore, the variety of NURC is not represented by this pole of formal Brazilian Portuguese presented in Table 6.4, but is placed on the continuum between formal and informal language.

*Os gramáticos registraram já na década de 70 essa alteração de uso: [...] se emprega “vocês como o plural de tu” por ter caído o pronome vós em desuso (Ilari et al. 2002: 85).*

According to NURC, *tu* in rather formal speech is common only for the South (Porto Alegre) of Brazil (cf. Ilari et al.: 77 ff.). In São Paulo, *tu* does not occur in PC, as was confirmed by NURC (cf. e.g. Ilari et al. 2002: 77 ff.). Here, *você* and the cliticized forms *ocê* and *cê* have substituted *tu* almost entirely. The latter is rather rare and mostly socially stigmatized, particularly when accompanied by the 3<sup>rd</sup> pers. sg. as e.g. in *tu vai atrás disso*. While this expression is also rather common in the PP of further areas of Brazil like Rio de Janeiro, in São Paulo, it is mostly associated with the *Nordestinos*, due to their presence in São Paulo. Similar syntagms that reveal no agreement between the verb and the subject are quite common in PP. Dores Nicolau (2000) argues convincingly that in the case of 3<sup>rd</sup> pers. sg. and pl. the absence of verbal agreement as in *as pessoas fala* (Subj: 3<sup>rd</sup> pers. pl. + Verb: 3<sup>rd</sup> pers. sg.) constitutes a “regra variável, cuja aplicação depende de determinados fatores, estruturais e não estruturais.”<sup>140</sup>

The change of the pronouns has led to a profound change of the verbal paradigm from the ancient norm of Brazilian Portuguese (today still spoken in most of Portugal’s varieties) with six distinguishing forms to the general employment of the 3<sup>rd</sup> pers. sg. in the norm-distant varieties in Brazil. The use of the 3<sup>rd</sup> pers. sg. was expanded to the 2<sup>nd</sup> pers. sg. because of the substitution of *tu* by *você*, a form that is used with the 3<sup>rd</sup> pers. sg. as in *Você gosta de arroz* instead of *Tu gostas de arroz*, the ancient norm that does not occur in São Paulo’s varieties. A similar effect showed the substitution of *nós* by *a gente* because the latter also requires 3<sup>rd</sup> pers. sg. instead of 1<sup>st</sup> pers. pl. as in *A gente gosta de arroz* instead of *Nós gostamos de arroz*. In a generalizing process the 3<sup>rd</sup> pers. sg. is also expanded to further persons resulting in the absence of verbal agreement for all plural forms and the 2<sup>nd</sup> pers. sg. as Table 6.8 shows:

---

<sup>140</sup> Apart from these factors, Dores Nicolau (2010) enumerates stress on the final syllable, the grade of the phonic difference between the plural and singular forms, but also sociolinguistic factors as age and social group, demonstrating that the loss of verbal agreement is most common for elder people of low social level (cf. l.c.: 136 ff.).

Table 6.5: Verbal paradigm of simple present in ancient PC, current formal and informal Brazilian Portuguese.

Person	PC (ancient)		Formal Brazilian Portuguese (current)		Informal Brazilian Portuguese (current)	
1 <sup>st</sup> pers. sing.	Eu	falo	Eu	falo	Eu	falo <sup>141</sup>
2 <sup>nd</sup> pers. sing.	Tu	falas	(Tu)/você	fala(s)	Tu/você	falaØ/fala
3 <sup>rd</sup> pers. sing.	Ele/ela	fala	ele/ela	fala	ele/ela	fala
1 <sup>st</sup> pers. pl.	nós	falamos	Nós/a gente	falamos/fala	Nós/a gente	falaØ/fala
2 <sup>nd</sup> pers. pl.	vós	falais	vocês	falam	vocês	falaØ
3 <sup>rd</sup> pers. pl.	Eles/elas	falam	eles/elas	falam	eles/elas	falaØ

Rodrigues (2000: 59) states that in São Paulo, the employment of the singular verb in connection with a plural subject as in the following example is quite frequent like in (199). The absence of personal agreement like in (200), though, is rather perceived as “characteristic of the variety spoken by an adult *Paulistano* who is an analphabet or has received only little school education” (my translation, sn).<sup>142</sup> Ferreira Coelho (2006: 47 ff.) underlines this finding but adds a stylistic dimension, explaining that for example hip-hop-artists from São Paulo use the form nós + 3<sup>rd</sup> pers. sg. to create their identity as one social group.

(199) *As pessoaØ falaØ, falaØ, mas não resolveØ nada. (Castilho 2010: 208),*

(200) *Nós não querØ vale nós querØ dinheiro. (Niehoff 2007: 48)*

This change of the verbal paradigm has also consequences on the employment of subject or object pronouns: While the subject pronoun can regularly be omitted as long as the verb forms are distinguishable, the subject nouns must be employed if the verb paradigm is reduced to the 3<sup>rd</sup> pers. sg. to give unambiguous information about the subject: *A gente/ele/você gosta de arroz.*

Considering the complement, in informal varieties of Brazilian Portuguese, the clitic forms *o/a* and *lhe* are rarely used. *Lhe*, for example, is only employed when referring to the 2<sup>nd</sup> pers. sg. in alternation with *te* as in *eu lhe mato/eu te mato* (cf. Castilho 2010: 207). In spoken Brazilian Portuguese, the cliticized form required for the position of the direct object is either omitted as in *Eu ví Ø*, or expressed by the use of the subject form *ele/ela* as in *Eu ví ele*.<sup>143</sup>

<sup>141</sup> The construction of 1<sup>st</sup> pers. sg. (subject) with 3<sup>rd</sup> pers. sg. (verb) *Eu fala* is discussed in the literature, but not for urban varieties.

<sup>142</sup> The original quote: “tende a constituir traço característico da língua falada pelo paulistano adulto analfabeto ou de baixa escolaridade” (Rodrigues 2000: 59).

<sup>143</sup> For an extensive discussion of the factors influencing the employment of clitics and pronouns, cf. e.g. Reich (2002).

Regarding the complement of the 3<sup>rd</sup> person in Bolivian Castilian, *lo* as well as *le* are discussed as invariable complements as they are used even where the Castilian norm would demand no complement at all, as in *Cerrámelo las ventanas!* (Mendoza 2008: 226 f.) or the indirect object *le* although the direct object would be required by the norm as in *Le habían pegado bien grave* (l.c.).

Interestingly, the phenomenon of absence of agreement in the verbal phrase is not reported for Bolivian Castilian, not even for the norm-distant varieties.

#### 6.2.10.1 Personal pronouns

---

The description and analysis of the personal pronouns will focus on the characteristics of the pronominal system of Brazilian Portuguese that have influenced the verbal paradigm, i.e. *você* and a *gente* substituting *tu* and *nós*. As regards the other personal pronouns, mostly the Portuguese pronouns are used. Only *yo* is often found instead of *eu*, probably due to the phonological closeness.

Regarding the second person singular<sup>144</sup>, there are two speakers who use *tu*:

(201) *Você en la área de trabajo que tu faz (1,7) buscas trabalho (A24 (25) m. 2004 S/SA Ce: 12.30).*

(202) *Tu não é daqui? (A8 (27) f. 2005 R. Ce: 8.10).*

(203) *Onde tu mora? (A8 (27) f. 2005 R. Ce: 4.18).*

Interestingly, both pronouns, *você* and *tu*, are found in (201), but the regular 2<sup>nd</sup> pers. sg. morpheme *-s* is only realized for *buscas*, while *fazer* has no additional morpheme, i.e. the 3<sup>rd</sup> pers. sg. In (202) and (203), *tu* also occurs with the 3<sup>rd</sup> pers. sg. inflection. Because of the verb form, I would interpret these examples not as Castilian but PP. As mentioned above, *tu* is rare in PC (except for the south of Brasil), but used for example in the northeastern dialects. As the speaker of the latter two sentences enthusiastically narrates her living in a *favela* in São Paulo, where many people originating from the northeast of Brazil live, it seems safe to attribute these examples to their influence. Similarly, in (201) the speaker pronounces *tu faz* as [tu 'fais]; exactly in this context, the employment of *tu* is quite salient in spoken Portuguese of São Paulo.

In the other cases of direct addressing in the 2<sup>nd</sup> pers. sg., *você* or the reduced form *cê* occur:

(204) *Você vai ficar boca aberta (A9 (42) m. 1994 D. ZL: 0.02).*

(205) *Entonces falamos 'você tem isso você tem isso você não tem que preocupar-se' (A22 (27) f. 2008 P. ZS: 17.18).*

(206) *Cê pode ter visto nossos pamfletos nós atendemos também às outras pessoas de Paraguai (A21 (62) m. 1964 E. Ce: 1.13).*

---

<sup>144</sup> As there are very few examples of the use of 2<sup>nd</sup> pers. pl., generating no further insight, I claim that basically the same observations are valuable as for 2<sup>nd</sup> pers. sg..

(207) *Aí **cê** quer saber como eu me integrei dentro da sociedade brasileira? (A19 (42) m. 1988 P. Ce: 0.24).*

(208) *Aí está pior estava ... **cê** viu né minha casa lá (A11 (20) m. 2002 S. Pe: 2.48).*

(209) *Tava estudando aí e ele falou pra nós né '**cês** nó querem ir están a fim de ir para Brasil assim trabalhar **cê** vai ganhar lá em dolares' (A6 (34) m. 1999 S. ZL: 1.52).*

Additionally, *você* and *cê* are also used for generalizing declarations as in (210) to (212), broadening personal experiences of the speakers to become an objectively comprehensible fact.

(210) *É o dono o patrão que dá a vivenda e a comida mas isso desconta do salário que **você** tem (A4 (21) m. 2006 S. ZL: 8.00).*

(211) *Onde **você** vai **cê** mora nos Estados Unidos **cê** vai falando todo dia inglês vai aprendendo [...] mesma coisa (A15 (28) m. 2000 D. Ce: 2.49).*

(212) *Então final quando **você** comenza **você** não sabe **cê** está utilizando o término corretamente ou se já está éh misturando com com a língua materna né (A18 (30) m. 2004 P. Ce: 2:13).*

In both contexts, *você* is widely spread<sup>145</sup>; *cê*, though, is only found in the speech of male informants having spent more than eight years in São Paulo and one ambiguous example in the speech of a young male physician (cf. (212)): As *cê* and the impersonal *se* normally used for this type of generalizations are homophone – at least for speakers who do not reduce *se* to [si] –, it is impossible to decide if the speaker intended to use *cê* or *se*. Only because of the context with *você*, however, I transcribed it as *cê*. It seems as though the reduction of *você* to *cê* represents a feature that learners integrate into their speech rather late and that women tend to avoid.

Regarding the 1<sup>st</sup> pers. pl., I find very few cases of *nosotros* in my corpus:

(213) ***Nosotros** fornecemos a eles os medicamentos (A22 (27) f. 2008 P. ZS: 15.22).*

(214) *Con ese::: miedo ... por eso que as veze **nosotro** só moramo dois ano depois vamos embora depois outro volta va volta va así (A1 (23) m. 2006 SA Ce: 2.46).*

As the speakers of (213) and (214) are the informants having spent least time in São Paulo, this insertion of *nosotros* can be attributed to duration of stay.

Apart from the Castilian influence, the variation between *nós* and *a gente* in Brazilian Portuguese has to be considered, too. As was shown before, *a gente* is frequently used, on the one hand as a personal pronoun, but also in the original Castilian semantics referring to other persons without including the speaker.

The use of *a gente* with this generalizing meaning has declined in Brazilian Portuguese in favor of expressions like *o pessoal* or *todo mundo*. Those are also found in my corpus, particularly in the

---

<sup>145</sup> On informants with very little contact to Brazilians I have no data, because neither *você*, *cê* nor other pronouns of the 2<sup>nd</sup> pers. sg. are found in their speech examples.

speech of informants having a high level of Portuguese proficiency and working outside the textile industry:

- (215) *Em Santa Catarina o pessoal é mais éh::: mais educado mais simples. Aqui o pessoal já é um::: ya stá todo misturado né mas em geral o brasileiro é bom como pessoa é bem legal (A20 (27) f. 2006 P. Gu: 4.10).*
- (216) *São de Cochabamba não. Tem:: pouco pessoal de de Cochabamba aqui (A17 (31) m. 2007 P. Ce: 0.31).*
- (217) *O pessoal pode ir lá e estudar o que quiser (A19 (42) m. 1988 P. Ce: 5.42).*
- (218) *Todo mundo hablava saludava éh [...] todo mundo asociable a gente muito bom (A8 (27) f. 2005 R. Ce: 1.00).*

Considering only the pronominal use of *a gente*, all informants whose speech samples contain the 1<sup>st</sup> pers. pl. use *a gente*, as (219) to (222) illuminate:

- (219) *Mas **a gente** quiere né pero::: ((1,4s)) **nós** vamos acho que vamos ficar aqui (A28 (28) m. 2002 S. ZL: 3.15).*
- (220) *Entendo completamente. Só que **a gente** nó fala (A3 (34) f. 2004 SA ZL: 0.47).*
- (221) ***A gente** vai pra fazer compras alguma coisa importante né [...] fora do trabalho não dá para sair também né (A15 (28) m. 2000 D. Ce: 6.35).*
- (222) *Yo traba... éh aqui: **a gente** está trabalhando em outra coisa (A23 (45) f. 1990 Ph ZO: 1.40).*

Examples for the pronoun *nós* are only found for some speakers, although the form is phono- and morphologically closer to Bolivian Castilian *nosotros*. Interestingly enough, the occurrences concern almost only men who are either sewers who have spent little time in São Paulo or have had very little contact with Brazilians ((223) and (224)), or the eldest Bolivians having spent most time in São Paulo and having the most elevated social position ((225) to (227)):

- (223) *Aqui **nós** moramos assustados porque tem::: ((0,4s)) nó es como en allá en allá você pode sair ((0,4s)) altas horas da noite nao acontece nada com você (A1 (23) m. 2006 SA Ce: 2.22).*
- (224) *Tem gente bra... de brasil que ele elas eles falam que **nós** ganamos muita muito dinheiro aqui (A4 (21) m. 2006 S. ZL: 12.17).*
- (225) *Mas aqui é diferente **nós** temos que respeitar as costumes que tem ele porque estamos **nós** como estrangeiros (A9 (42) m. 1994 D. ZL: 10.52).*
- (226) ***Nós** temos essa herança do espanhol por causa deles (A19 (42) m. 1988 P. Ce: 12.49).*
- (227) *E:: se vê a quantidade de remédios que **nós** temos que sao doações e **a gente** também faz doação para as pessoas que precisam (A21 (62) m. 1964 E. Ce: 2.47).*

I find only one woman using *nós*; she is *dona* and one of the woman having spent most time in São Paulo (8 years):

- (228) *Você sabe quando **nós** conversamos com boliviano boliviano mesmo pais... paisano assim patrício meu nó es tímido não (A5 (28) f. 2000 D. ZL: 8.46).*

This unequal division could indicate that the *nós* of the first group is due to L1-transfer because of little language proficiency, while for the second group, the more formal *nós* and more informal *a gente* vary freely, employing them for stylistic variation. The speaker of (227), for example, seems to employ *nós* first and foremost when he speaks about giving general advice to other Bolivians. As he works also as a pastor in an evangelical church, the use of *nós* could be interpreted as part of his pastoral identity, offering advice to his fellows.

All in all, the pronominal system of São Paulo's Brazilian Portuguese has been adopted thoroughly by the Bolivians. Although the forms differ strongly from Bolivian Castilian, *você* and *a gente* are employed by all speakers using personal pronouns in these persons – a clear sign for accommodation by a socially highly salient marker. The differentiation inside the group regarding *cê* and *nós* can be attributed various social factors such as gender, duration of stay, language proficiency and social status.

#### 6.2.10.2 The absence of verbal agreement

---

In 6.2.10.1, I have already given the only examples of the absence of verbal agreement in the case of *tu* constructed with 3<sup>rd</sup> pers. sg. (cf. (202) to (204)). As these are the only examples for this person, I concentrate my analysis on verbal syntagms where the subject stands in 1<sup>st</sup> pers. sg. and pl. or 3<sup>rd</sup> pers. pl. but the verb in 3<sup>rd</sup> pers. sg. As was explained above, verbal syntagms containing 1<sup>st</sup> pers. pl. or 3<sup>rd</sup> pers. pl. subjects and the verb in 3<sup>rd</sup> pers. sg. are frequent, but socially rather stigmatized. There are few occurrences of *nós* in my corpus. I find only two examples of the absence of verbal agreement as it regards *nós* together with the 3<sup>rd</sup> pers. sg. Interestingly, in (229) the following verbal form is even an insertion of a Castilian form; hence, language proficiency might have had influence here. Both speakers had come to work in the textile industry still belonging to a low social class:

(229) *Nós quiere ficar-nos aqui nomás* (A25 (60) f. 1989 S/SA ZL: 4.14).

(230) *Por o hecho que nós trabalhavamos ela pensava que nós ganhava muito dinheiro* (A4 (21) m. 2006 S. ZL: 12.54).

There are far more examples of the absence of verbal agreement regarding 3<sup>rd</sup> pers. pl. in combination with 3<sup>rd</sup> pers. sg. for men, women, seamstresses, *donas* and physicians:

(231) *E a gente [...] tá::: bien ligado a todas las cosas que tá acontecendo aqui* (A26 (25) f. 2005 S. ZL: 4.42)

(232) *Eles falava inglês* (A2 (18) f. 2005 SA ZL: 2.07)

(233) *Não era português que eles falava né* (A6 (34) m. 1999 S. ZL: 28.05)

- (234) *O preço que **eles ganha** né, é pouco.((1,6s)) Porque **eles deixa se deixa** explorar com o coreano (A9 (42) m. 1994 D. ZL: 22.10)*
- (235) *A minha esposa fazia uns bolos gostoso aí falei 'vamo tentar na feirinha porque **os patrício** também **gosta**' né de comer bolo (A12 (39) m. 1990 D. ZL: 6.07).*
- (236) *Eles vêem sua liberdade que **eles não tinha** lá ((1,0s)) e que:: eles têm o dinheiro que **eles não tinha** lá (A18 (30) m. 2004 P. Ce: 11.21).*
- (237) ***Falta** ainda para aposentar **falta cinco anos** mais (A23 (45) f. 1990 Ph ZO: 3.01).*

Remembering the Bolivians' difficulties with the realization of the nasal diphthong [ã̃] in the context of the 3<sup>rd</sup> pers. pl., at first sight, it only seems logical to associate the many examples of absent verbal agreement with the phonological process of denasalization. This reason also seems plausible for similar cases where [ẽ̃] is realized as [e] in (238) and (239), although (240) could be due to *hypercorrection* involving the palatization of *-ti/dĩ*<sup>146</sup>.

- (238) *Têm **brasileiros** que nó **sabe** trabalhar né (A24 (25) m. 2004 S/SA Ce: 12.55).*
- (239) *E acaba sempre pegando nó pode ser éh:: nó pode ignorar também **os problemas** que **acontece** aqui né (A19 (42) m. 1988 P. Ce: 15.17).*
- (240) *Eu ouví que **muitos** éh se **sente** discriminados (A20 (27) f. 2006 P. Gu: 8.00).*

In (236), though, the speaker realizes a perfectly nasalized [nã̃] right before the seemingly denasalized *tinha* [ˈtĩɲa]. Additionally, the denasalized plural morphemes are regularly realized as [an], corresponding to their Castilian equivalents. Thus, it seems safe to argue that for my corpus, the phonological process of denasalization plays a certain role, but the absence of verbal agreement exceeds a purely phonological explanation.

Further arguments for this cause are presented by (241) to (244), where irregular verbs would show more phonic differences between 3<sup>rd</sup> pers. sg. and 3<sup>rd</sup> pers. pl. than [a] and [ã̃], e.g. *vai – vão, é – são, morreu – morreram*. Still, there are many examples, particularly of male speakers, where the 3<sup>rd</sup> pers. sg. verb forms result in the absence of verbal agreement in the syntagm:

- (241) *Aí vai aos poucos aos poucos **vai** baixando **as forças** (A11 (20) m. 2002 S. Pe: 10.47).*
- (242) *Como lá não tem muito trabalho pelo problema da população que é muita [...] e que **os profissionais** **é** muitos (A18 (30) m. 2004 P. Ce: 1.49).*
- (243) *Até **as cadeias públicas** **é** muito diferente (A19 (42) m. 1988 P. Ce: 24.10).*
- (244) *Né ((0,5s)) que ((0,2s)) são **meninos** novos que migraram de um país e às vezes os pais ni sabem que **morreu**. (A21 (62) m. 1964 E. Ce: 12.26)*

Another influencing factor could be the absence of number agreement resulting in examples like (245) to (247), where only the first element of the nominal syntagm has the plural morpheme.

<sup>146</sup> Cf. the previous discussion on this in section 6.1.5

(245) *As pessoa me ensinava eu aprendendo* (A2 (18) f. 2005 SA ZL: 1.32).

(246) *Mas mis filho no quer* (A3 (34) f. 2004 SA ZL, S. 2.35).

(247) *Se eu contasse as coisa que aconteceu* (A9 (42) m. 1994 D. ZL, 21.15).

Interestingly, the absence of verbal agreement is also found in a Castilian context:

(248) *Cuatro está en Bolivia* (A27 (35) f. 2006 SA ZL: 5.06).

It seems, thus, that various factors such as the weakening of nominal agreement in CP and PP and phonological processes, result in the generalization of the verb form of 3<sup>rd</sup> pers. sg. for the 3<sup>rd</sup> pers. pl.

Regarding the construction of this verb form in combination with 1<sup>st</sup> pers. sg., examples like (249), could be due to phonological variation between /e/ and /i/ common in CP, even if the speaker here is a monolingual woman belonging to La Paz' upper class:

(249) *Algumas coisas ((0,5s)) dentro de mia profissão teve que aprender* (A23 (45) f. 1990 Ph ZO: 7.16/7.51).

However, there are further examples like (250) to (254) that are not explicable by phonological processes because they concern verb forms that differ more from each other than only by *-e/i* or *-o/u*. Particularly frequent are cases where the speaker seems to adopt the verbal form used in the question of the interviewer without realizing the person change from 2<sup>nd</sup> pers. sg. to 1<sup>st</sup> pers. sg.:

(250) - (I): *Você saiu da costura mesmo?* - (A8): *Saiu.* (A8 (27) f. 2005 R. Ce: 2.14)

(251) - (I): *Você conhece outros bairros aqui?* - (A3): *Conhece não muito não conhece* (A3 (34) f. 2004 SA ZL: 6.07).

(252) - (I): *Nossa quantos anos cê tem?* - (A26): *Eu tem* *veinticinco* (A26 (25) f. 2005 S. ZL: 0.16).

(253) - (I): *Há quanto tempo cê tá aqui?* - (A10): *Eu tá* *aí quase há dois dois anos já* (A10 (22) m. 2006 S. Pe: 0.56).

(254) - (I): *Quando você veio você veio já pra cá pra Bom Retiro ou?* - (A7): *Não eu veio* *por Parque Novo Mundo* (A7 (32) f. 1994 S. Ce: 0.58).

As these examples mostly concern speakers with little language proficiency in Portuguese, they could constitute a learner's technique of just imitating the linguistic material offered by the interacting person.

Apart from these examples, though, I also find sentences where the absence of verbal agreement between the subject in 1<sup>st</sup> pers. sg. and the verb in 3<sup>rd</sup> pers. sg. does not depend on the preceding question, but seems to constitute a step further into the direction of the generalization of the 3<sup>rd</sup> pers. sg. verb form for all persons:

(255) *Mas eu gosta de Brasil* (A1 (23) m. 2006 SA Ce: 3.02).

(256) *Eu tem* *una filha casada que tiene la su filinha* (A25 (60) f. 1989 S/SA ZL: 4.17).

(257) *Mas eu ouviu a festa a festa era bonito (A10 (22) m. 2006 S. Pe: 7.51).*

(258) *Primeiro bolivianos depois ya **trabalhou** a fora também (A15 (28) m. 2000 D. Ce: 3.58).*

(259) *Eu aprendí porque são seis anos que **eu serve** deus (A9 (42) m. 1994 D. ZL: 3.35).*

Instead of attributing these examples unambiguously to accommodation, it rather seems as though various phonological and morphological factors together with simplification and possible accommodation toward a norm-distant PP strengthen each other until the different processes result in the weakening of standard agreement rules and the stabilization of the general use of the 3<sup>rd</sup> pers. sg. for all persons and for all tenses<sup>147</sup>. The fact that the examples presented above come from almost all speakers, regardless of sex, age, duration of stay or profession, confirms this thesis.

### 6.2.11 Prepositions

---

The prepositional inventory of Spanish and Portuguese coincide almost completely including identical prepositions like *a*, *por*, *para* and others that differ at most phonetically e.g. *en* and *em*, respectively. Therefore, I focus on their use in specific linguistic contexts, e.g. whether my informants employ the contraction of *por*, *de* and *em* with the article as the Portuguese norm requires, or if traces of the variation regarding the prepositional system discussed for bilingual speakers of CP are found, e.g. the confusion or omission of prepositions. Hence, they are rather revealing for the influence of CP-transfer.

One characteristic often discussed for Bolivian Castilian is the construction of *en* + *locative* (cf. e.g. Mendoza 2008: 224), expressed by *en* + deictic terms like *aquí*, *acá*, *ahí*, *allí* and *allá*.

(260) *En allá estaba hablando con tu hermano.*

(261) *En ahí están jugando.*

Phrases like **Fehler! Verweisquelle konnte nicht gefunden werden.** and (261) are common for informal situations, although

*[e]n castellano, los deíticos [...] ya indican por sí mismos la relación especial por lo que no requieren de una preposición locativa”(l.c.).*

While this construction is mostly explained as a direct imposition of Quechuan or Aymaran structure (cf. Mendoza 2008: 224), Pfänder (2002: 235) holds the view that it represents a generalization of a structure already present in Castilian, following the lines of “en la mesa” that permits the distinction between place and direction even for deictic terms.

---

<sup>147</sup> Although I didn't focus on this aspect, there are obviously cases of present indicative, future conjunctive and perfect indicative.

In my corpus, there are various examples like (262) to (264), mostly by speakers having spent little time in São Paulo or having little contact with Brazilians.

(262) *En<sup>148</sup> alf fazia alguma roupa pra mim (A14 (28) m. 2006 S. ZN: 3.02).*

(263) *Mas en aqui nó es assim né (A25 (60) f. 1989 S/SA ZL: 11.12).*

(264) *Nó es como en allá en allá você pode sair ((0,4s)) altas horas da noite não acontece nada com você (A1 (23) m. 2006 SA Ce: 2.22).*

In contrast to these findings, my corpus contains many examples like (265) to (267), where the locative receives no preposition at all, although the toponym would require *em* to complete the prepositional syntagm.<sup>149</sup>

(265) *Fui a morar Ø Ibateringa, fui a morar Ø Itaquera (A8 (27) f. 2005 R. Ce: 1.46).*

(266) *Porque depois de Cambucí foi morando Ø Bom Retiro lá (A12 (39) m. 1990 D. ZL: 2.44).*

(267) *Que é o dialeto que se:: utiliza lá:: Ø Cochabamba (A18 (30) m. 2004 P. Ce: 1.22).*

But not only toponyms are constructed without the necessary preposition. The same phenomenon is observed in other contexts as in (268) and (269), even though it heavily complicates the comprehension of the phrases.

(268) *Brasileiros que trabalham eu ví Ø costura mas ((0,5s)) só em poucas oficina né (A14 (28) m. 2006 S. ZN: 11:21).*

(269) *Trabalho Ø outro consultório pela Vila Maria (A18 (30) m. 2004 P. Ce: 8.11).*

Interestingly, this feature is not limited to speakers originating from rural areas in Bolivia and working in the textile industry, but also used by monolingual academics.

This is rather surprising, because for CP, the strong variation concerning the use of prepositions is attributed to the nonexistence of prepositions in Aymara and Quechua and therefore only found for bilingual speakers of CP (Mendoza 2008: 228). According to Mendoza (2008), L1 transfer from the indigenous languages results in the confusion of prepositions, e.g. *de* and *con* as in

(270) *Sufrió un percance de mis herramientas (l.c.),*

the addition of a preposition, e.g. *de* as in

(271) *A ver pues vemos de poco a poco (l.c.)*

or even the complete omission of prepositions, e.g. *en* as in

---

<sup>148</sup> The preposition in these examples is pronounced as [en] and thus transcribed accordingly, but in another context the transcription could be different, reflecting that the denasalized pronunciation of *em* could also sound like [en].

<sup>149</sup> To distinguish the absent preposition in the examples, I apply Ø to the position where the preposition would be expected.

(272) *Porque Ø cualquier momento pueden decir que no hay (l.c.).*

In my corpus, I find similar examples for all speakers, not just bilingual ones: not only *em* is omitted, but also *a* or *para* in syntagms describing a direction, e.g. (273) and *de* as in (274).

(273) *Acaba mi trabalho a gente vai embora Ø su casa para descansar (A24 (25) m. 2004 S/SA Ce: 8.43).*

(274) *Pelo motivo Ø trabalho sim (A18 (30) m. 2004 P. Ce: 16.52).*

Furthermore, the omission of prepositions also occurs in verbal dependency, i.e. the verb requires the preposition *de* as in (275) and (276) and *em* as in (277). The first of these examples could still be interpreted as difficulties due to the different construction of the verb *gustar* in Castilian (Clitic + verb + subj.) and *gostar* in Portuguese (Subj. + verb + de), but (277) can only be attributed to the weakening of the prepositional system because the verbs *entrar* requires the same preposition in Castilian and in Portuguese:

(275) *Não gosto Øo lugar. (A23 (45) f. 1990 Ph ZO: 8.04)*

(276) *Gostei Øaqui né (A6 (34) m. 1999 S. ZL: 24.39)*

(277) *Estoy fazendo curso do português e agora estou faz... voy entrar Ø um curso de ultrassom (A22 (27) f. 2008 P. ZS: 0.8).*

The fact that the prepositional paradigm shows strong unstructured variation similar to the processes described for CP is underlined by the following examples illustrating the *confusion* between different prepositions. In (278) and (279), for example, *por* is used instead of *para*, while (280) and (281) display *com* instead of *por* and *de*, (282) *em* and (283) *sobre* instead of *com* / *em* respectively.

(278) *Se emigrar **por** um país sin:: sin ter um apoio moral [...] sin ter ajuda é difícil (A17 (31) m. 2007 P. Ce: 7.14).*

(279) *E quando [...] tem que viajar **por** Jabaquara ou **por** Penha **por** outros lugares ((0,4s)) nó sólo aqui nomás eu conoce. (A25 (60) f. 1989 S/SA ZL: 14.34)*

(280) *Porque eles deixa se deixa explorar **com** o coreano. (A9 (42) m. 1994 D. ZL: 22.13)*

(281) *Fica perto **com** Corumbá. (A1 (23) m. 2006 SA Ce: 0.43)*

(282) *Ele saiu **nos** quinze anos. (A23 (45) f. 1990 Ph ZO: 1.11)*

(283) *Em Oruro ((0,9s)) e:: ((0,8s)) vou ajudar **sobre** tua pesquisa né (A19 (42) m. 1988 P. Ce: 0.13).*

The examples (284) and (285), where *em* is used although no preposition would be needed, complete the rather chaotic picture. It is impossible to identify a regular pattern, be it linguistic or social factors explaining the variation.

(284) *A gente falava **no** outra língua né (A15 (28) m. 2000 D. Ce: 2.13).*

(285) ***Na** Bolívia está dividida agora (A17 (31) m. 2007 P. Ce: 7.53).*

Almost all speakers show either omission or confusion of prepositions. Although I find a higher frequency for speakers having spent little time in São Paulo, there are also many examples by women and men who have worked outside the textile industry in Brazil for many years.

The contraction of the prepositions *a*, *de*, *em* and *por* with the following article *o/a/os/as* when used in sequence, e.g.: *a + o = ao*; *de + a = da*; *em + os = nos*; *por + as = pelas*, etc., required by normative grammars of PC are a further aspect that is regularly described as problematic for Spanish learners of Portuguese, because in Castilian, the only contraction that takes place between preposition and article is the vowel reduction of the prepositions *a* and *de* when followed by *el*: *a + el = al* and *de + el = del*.

This is confirmed by a Spanish learner describing her difficulties with this construction:

*[T]he combination of the preposition /por/ and the definitive articles /o/, /a/, /os/ and /as/ to make /pelo/, /pela/, /pelos/, pelas/ [...] were extremely difficult for me to process and accept given that [...] I had to work extremely hard to block out the articles in Spanish /la/, /el/, /las/, /los/ and their sounds” (Cowles and Pires 2008: 246).*

In my corpus, there are also various examples where the contraction does not take place. It concerns mainly the preposition *de* ((286) and (287)), but also *em* ((288) and (289)) and *por* ((290) to (291)).

(286) *Meu pai me fala que eu tenho que pregar a palavra **de ele** (A9 (42) m. 1994 D. ZL: 6.07).*

(287) *Conhece aqui a outros barrios **de aqui** perto (A28 (28) m. 2002 S. ZL: 8.51).*

(288) *Às vezes **em um** ponto de ônibus eu falo com com o pessoal (A17 (31) m. 2007 P. Ce: 2.27).*

(289) *Aqui **em essa** pequena parte de São Paulo donde mora muito boliviano yo m... ((0,3s)) yo olho que hay a a veces muita discriminação (A22 (27) f. 2008 P. ZS: 19.27).*

(290) *Mas que português es quasi castellano mal falado parece; porque aí os bolivianos que moram **por os** cantos falam como tipo português (A5 (28) f. 2000 D. ZL: 6.39).*

(291) ***Por a** situação política económico ficou ruim allá na Argentina eles imigraram em massa (A21 (62) m. 1964 E. Ce: 3.29).*

Yet, the contraction rule is also not obligatory in spoken varieties of Brazilian Portuguese. Therefore this feature is not unambiguously attributable to L1-transfer.

Interestingly, I find also hypercorrections where the article and the preposition are contracted although only the simple preposition is required. However, in these cases, phonological assimilation processes certainly play an important role as well.

(292) *Eu vinha **da avião** (A2 (18) f. 2005 SA ZL: 6.19).*

(293) *Ah foi foi mesmo ((1,8s)) nó é fácil acostumar à ((0,4s)) tradição **da aqui** (A15 (28) m. 2000 D. Ce: 1.37).*

Similarly, there are rare cases of *pelo* / *pela*, but only by speakers who have either spent many years in São Paulo or are working in academic professions:

(294) Trabalho outro consultório **pela** Vila Maria (A18 (30) m. 2004 P. Ce: 8.11).

(295) Que inclusive aqui **pelo pelo** ((0,7s)) **pela** lei do Lula hoje em dia né (A19 (42) m. 1988 P. Ce: 3.48).

A further phenomenon that could rather be attributed to accommodation is the phonetic reduction of the preposition *para* that mostly results in the omission of the first vowel *para* > *pra* as in:

(296) O vizinho que deu um pouco **pra** mim, **pra** mim não (Kewitz 2009: 630).

If a vowel follows the preposition, the second vowel is substituted by this vowel, e.g. *para* + *o* > *pro*: *deu pro muleque*. Further examples show also the substitution of the second vowel, while the first vowel is maintained as in *para ocê* > *pa[ru]cê* or diphthongization as in *pra o rapaz* > *[praw]* *rapaz* (Kewitz 2009: 605 f.). However, for the latter I find no examples in my corpus. For *pra*, though, there are some occurrences:

(297) Então eu: falei com meu cunhado então vamo vamo **pra** Brasil (A5 (28) f. 2000 D. ZL: 3.32).

(298) Eu fui **pra** San Mateus (A4 (21), m. 2006 S. ZL: 9.05).

(299) Mis melhores amigos ((0,5s)) **pra** te falar verdade são os brasileiros (A13 (33) m. 2000 D. ZL: 2.42).

For *pro*, I only find one example:

(300) Falando sobre a ((0,6s)) especificamente sobre ((0,7s)) a imigração do boliviano **pra** cá **pro** Brasil (A19 (42) m. 1988 P. Ce, S. 6.54).

However, more common are examples where *para* is not phonetically reduced:

(301) Então eu vou sempre viajando **para** esses lugares (A7 (32) f. 1994 S. Ce: 14.24).

(302) Eu solicitei **para** um vereador um político **para** ver se nos poderia ajudar [...] **para** conseguir um:: terreno [te'reno] **para** nós poder construir e ter um lugar próprio né (A21 (62) m. 1964 E. Ce: 2.37).

(303) Venho aqui **para** ajudar a la a gente [ʒentʃi] boliviana (A22 (27) f. 2008 P. ZS: 0.45).

It seems though as if the substitution of *a* by *para* as described by Kewitz (2009) is also taking place; however, there are still more contexts in my corpus where the Bolivians use the preposition *a* than in the Portuguese spoken in São Paulo:

*Em relação às ocorrências da preposição a no Português Popular, alguns casos referem-se a formas mais cristalizadas, como 'pedir graças a Deus' ou 'de sete às dez horas'. São nesses contextos que a preposição a ainda resiste – ao menos nos dados do Português Popular em São Paulo (Kewitz 2009: 630).*

This finding is mostly explained as a result of morph-phonological processes like the external sandi, being “mais favorável ao apagamento de *a* do que de *para*” (l.c.: 606).

Considering the prepositions in the context of a relative clause, in Brazilian Portuguese “a norma padrão prescreve piedpiping” (Tarallo 1996: 86) as in (304):

(304) *E um deles foi esse fulano aí, com quem eu nunca tive aula.*

Tarallo, however, demonstrates in various studies (1983, 1996 a, b) that this construction only occurs in written Portuguese. In oral contexts it is either realized as in (305) or (306):

(305) *E um deles foi esse fulano aí, que eu nunca tive aula com ele.*

(306) *E um deles foi esse fulano aí, que eu nunca tive aula.*

According to Mollica (2006), today, the norm-distant construction without pied-piping is even found in written Portuguese. In (305) the gap existing in the “norm”-following sentence is filled by a resumptive pronoun (cf. Tarallo 1996a: 41). This indicates that the relative pronoun, normally uniting the functions of a conjunction and a demonstrative or possessive, has lost the second function and only conserves the first connecting function. Thus, the relative is not perceived anymore as a nominal syntagm of the introducing phrase (cf. Ilari and Basso 2006: 116). In (306), a construction very common not only in PP but also in more formal varieties, the preposition is omitted completely resulting in a generalization of *que* that also includes the diminishing of other relative pronouns as *cujo*, *onde* and *quando* (cf. Castilho 2010: 202/342). This generalization of *que* instead of e.g. *quando* is also found in CP:

(307) *[E]n la tarde que te encuentras con los cuates [...] (Pfänder 2009: 150).*

In my corpus, cases of pied-piping as in (308) occur, but in general there is the tendency to omit prepositions like *com*, *de*, *em* or *a* entirely:

(308) *Graças a deus a gente **com a que** yo moro tenha auto (A22 (27) f. 2008 P. ZS: 18.25).*

(309) *Éh como a pessoa Ø **que** eu trabalhava aí não falava português só falava a nossa língua [...] não tinha problema (A7 (32) f. 1994 S. Ce: 7.20).*

(310) *Então é uma coisa Ø **que** também estamos cuidando (A21 (62) m. 1964 E. Ce: 12.24).*

(311) *Essa ONG Ø **que** eu estou atualmente né ((0,9s)) é isso contra isso que a gente luta né (A6 (34) m. 1999 S. ZL: 35.22).*

(312) *Mas têm bolivianos que deixam que é esse Ø **que** eles acostumam quando têm bolivianos que **eles** deixam (A3 (34) f. 2004 SA ZL: 5.00)*

As these examples show, the omission of the preposition also leads to the generalization of *que* as the only relative pronoun. (312) is particularly interesting because it shows the resumptive pronoun *eles* as described above. A further example for this phenomenon (without preposition) is also found in (313):

(313) *Não tem como los Parguayos que eles têm dois idiomas (A14 (28) m. 2006 S. ZN: 1.24).*

These last findings concern almost all speakers regardless of age, sex, duration of stay or educational level. The prepositions constitute one more phenomenon where the linguistic variation is huge. As

the similarity of the prepositions in CC and PC does not explain this amount of differences, omission, confusion and hypercorrection, I estimate that these phenomena are more common in CP spoken by **monolinguals** than it is discussed in literature. Accordingly, the occurrences in my corpus could be due to the influence of an already weakened system that gets even more instable in the language learning situation. Additionally, accommodational processes like the phonetic reduction of *para* are found but not very salient – this might be due to the little social saliency prepositions have in comparison with e.g. personal pronouns.

### 6.2.12 Word order

Generally spoken, in Brazilian Portuguese and CC, the regular word order is subject – verb – object (SVO). However, in CP, “the verb can be found in first, second or final position” (cf. Pfänder 2009: 103, own translation, sn). For the latter, Pfänder (l.c.) gives the following examples:

- (314) *De eso harta pena **tengo** (l.c.: 104)*  
 (315) *Rocío con su mamá **ha entrado** (l.c.).*

This free word order though, is attributed mainly to bilinguals and transfer from the indigenous languages, because e.g. Aymara and Quechua prefer SOV (Mendoza 2008: 231).

Therefore, findings of differing word order in my corpus could be interpreted as L1-transfer of CP. However, there are very few examples. In (316) and (317) the typical Castilian and Portuguese word order SVO seems to be dissolved with regard to the position of the object. Instead of its positioning right behind the verb, the examples show the insertion of an adverbial expression, a prepositional syntagm in the first example and a deictic in the second.

- (316) *Dos meninos vão começar fazer **na semana que vem ya** informática né (A9 (42) m. 1994 D. ZL: 5.16).*  
 (317) *Eu vim a procurar **aqui** trabalho (A8 (27) f. 2005 R. Ce: 2.08).*

In (318) to (320), the verb is realized in final position. In (319), there is even the word order OSV.

- (318) - (I): [...] *Vizinhos aqui são todos bolivianos também?* - (A27 (35) f. 2006 SA ZL): *Todo puro boliviano **es**.*  
 (319) *Quando eso uno **fala** ((0,3s)) eles pensam que você quiere humiliar a eles (A14 (28) m. 2006 S. ZN: 9.02).*  
 (320) *Evo Morales bastante **ajudou**. (A11 (20) m. 2002 S. Pe: 18.34).*  
 (321) *Mas tem brasileiro que [d]a gente **compra** (A3 (34) f. 2004 SA ZL: 6.39).*  
 (322) *Él quer su dinheiro **ganhar** e pronto (A6 (34) m. 1999 S. ZL: 10.53).*  
 (323) *Não deixava sozinho **ir** (A6 (34) m. 1999 S. ZL: 4.25).*

Although there is no direct object in (320) and (321), Portuguese norm would expect *bastante* and *[d]a gente* in final position after the verb. In (322) and (323), the full verbs *ganhar* and *ir* should be placed right after the auxiliary verbs *quer* and *deixava*, respectively.

As these examples are very rare and only found for persons working in the textile industry, this phenomenon could indeed be due to the influence of CP on their Portuguese. However, it seems as though duration of stay and language proficiency are not the decisive factors here, as the examples found in my corpus were realized by speakers who have spent at least three years in São Paulo and have rather high language proficiency.

### 6.2.13 Discourse marker

---

Discourse markers are meta-pragmatic linguistic signs (Traugott 1997: 3) typically occurring in oral communication that help constructing cohesive and coherent discourse and that refer not only to the intra-textual but also the interactive level between speaker and hearer:

*[F]uncionam como articuladores não só das unidades cognitive-informativas do texto como também dos seus interlocutores, revelando e marcando [...] as condições de produção do texto, naquilo que ela, a produção, representa de interacional e pragmático (Marcuschi 1989: 282).*

There are verbal markers including lexical units like *sabe?* and non-lexical elements like *ahn*, *éh*, as well as prosodic marker like pauses and prolongation and even paralinguistic markers like gestures and mimics (cf. Urbano 2003: 99). For my analysis, only lexical markers are considered. Non-lexical markers as well as the enumerated prosodic markers are considered for the transcription, but not for the analysis due to the limitations of this study. Similar to Marcuschi (1989), Castilho (2010: 229 f.) distinguishes interpersonal markers that are oriented toward the interlocutor serving the structuring of the conversational turns and textual markers serving the topical negotiation and development of the discourse. The latter are used to induce, accept, organize, model and end the topic while the interpersonal markers are employed to show the interlocutor the participation of her counterpart in the communicational act.

In CC, there are markers like *bueno*, *así*, *mira*, *entonces*, *pues*, *probablemente*, *todavía*; interpersonal markers include *¿no?*, *¿no vé?*, *¿ya?*, *mira*, *oye*, *hombre*, *pero* (cf. Prada 2001: 49 and 86). In studies about CP, *pero*, *pues* (and its reduced form [ps]), *nomás* and *ya* are discussed as *modal particles* that are used as “verdaderas posposiciones muy diferentes al uso que se les da en el castellano estándar” (Mendoza 2008: 228). According to Mendoza (l.c.), they reflect “matices pragmáticos de énfasis,

incitaci3n o atenuaci3n” as in *Ven nom3s* or *No me entiendes pues*. Pf3nder (2009), though, emphasizes their use for the textual cohesion and coherence (cf. l.c.: 117) and also exemplifies their frequent employment, for example as turn-taking signals (l.c.: 119 to 136):

(324) *Tengo que irme, comadrta, 3ya?*

(325) *3No te gusta no ve?*

(326) *Ay3dame pues.*

(327) *3Me puedes preparar que comer nom3s?*

There are very few examples of Castilian discourse markers; I interpret the use of *entonces*, *nom3s* and *ya* in almost every sentence as in (328) to (331) as isolectal characteristics of the individual speakers because only they show these markers in high frequency. While this could be due to a low level of language proficiency for the first two speakers – the first having been in S3o Paulo for only three months, and the other one having very little contact to Brazilians despite her having been in S3o Paulo for 20 years – the third one, having advanced Portuguese proficiency, also frequently uses the Portuguese equivalents *j3* and the idiomatic *despois*<sup>150</sup>.

(328) **Entonces** *t3m mais confianca com:: com gente que fala espanhol, [...] entonces 3h ajudamos muito essa gente (A22 (27) f. 2008 P. ZS: 1.34).*

(329) *S3 tamb3en elas trazem pouquinhos **nom3s** tamb3en n3: monte de ni camionada (A25 (60) f. 1989 S/SA ZL: 3.20).*

(330) - (I): *A senhora a sua fam3lia pretende ficar aqui? - Sim **entonces ya ya** n3s quiere ficarnos aqui **nom3s** (A25 (60) f. 1989 S/SA ZL: 4.11).*

(331) **Ya** *fiquei **ya** acostumando daqui **ya** (A5 (28) f. 2000 D. ZL: 5.46).*

Similarly, there are only few occurrences in my corpus containing a Bolivian Castilian discourse marker concluding the comment. Only *n3* and *n3 v3* as in (332) to (334), but these are exceptional cases:

(332) *O Ju3n es primo tu primo **n3 v3**? (A10 (22) m. 2006 S. Pe: 2.20)*

(333) *N3s tem bolivianos que n3 falam muito correto **n3 v3**? (A3 (34), f. 2004 SA ZL: 3.43)*

(334) *Depois a3 na ((0,6s)) f3vela 3st3 perta da USP **n3 v3**? (A8 (27), f. 2005 R. Ce: 0.45)*

In Brazilian Portuguese, typical textual markers are, e.g., *bom*, *assim*, *ent3o*, *e a3*, *provavelmente*, *sei la*, *valeu*, *3 isso a3*; interpersonal markers are, e.g., *olha*, *mas* in the beginning of the turn, *3*, *t3*, *t3 entendendo* to signal agreement during the discourse, and *n3?*, *viu?* or *t3?* to conclude the comment (cf. Castilho 2010: 229 f.). In a quantitative study, Urbano (2002) shows that in S3o Paulo *sabe?*, *viu?* and particularly *n3?*, are very common as turn-taking signals (l.c.: 215 and 218-19). *N3* repre-

<sup>150</sup> *Despois* is a mixed form of Portuguese *depois* and Castilian *despu3s*.

sents also the only marker that is employed by all speakers, being used idiosyncratically by some speakers. (cf. Urbano 2002: 219 and 235).

In my corpus, for most of the speakers, I find *então* and *mas* at the beginning of the turn as in (335) and (336), but also *bom* as in (337).

(335) *Então né ele nos trouxe falando maravilha né* (A6 (34) m. 1999 S. ZL: 1.25).

(336) *Mas a gente ya está:: se prepara: no prepara: no né* (A13 (33) m. 2000 D. ZL: 5.04).

(337) *Bom graças a deus com minha pessoa nunca tive discriminacón* (A13 (33) m. 2000 D. ZL: 6.48).

During the discourse, the form of the interaction being an interview determines that interpersonal markers to signal agreement come almost exclusively from the interviewer, making it irrelevant for my study. Discourse markers concluding the comment are for example *entendeu*, *viu* and *sabe*, although they are only employed by some speakers:

(338) - (A12): *Aí fui ficando.* - (I): *Aha.* - (A12): *Entendeu* (A12 (39) m. 1990 D. ZL: 1.02).

(339) *Falava tudo errado, entendeu* (A2 (18) f. 2005 SA ZL: 2.24).

(340) *Mas é geral viu ((0,3s)) que inclusive aqui pelo pelo ((0,7s)) pela lei do Lula hoje em dia né [...] você nao pode reprobar [...] né* (A19 (42) m. 1988 P. Ce: 3.44).

(341) *É o pessoal mais refinada éh sabe* (A17 (31) m. 2007 P. Ce: 2.56).

The most common concluding marker is *né* as (335) and (336) have already shown. It is employed by almost all speakers; even in linguistic environments that are dominated by Bolivian Castilian as (342) and (343) show:

(342) *Aquí nó se puede hacer né* (A25 (60) f. 1989 S/SA ZL: 10.51).

(343) *Mas en aquí nó es así né* (A25 (60) f. 1989 S/SA ZL: 11.12).

As *né* is mostly used to signal the conclusion of the comment, it is often followed by turn-taking transitions, i.e. the interviewer responds or at least signals comprehension:

(344) - (A2): *Mais ou menos. Dá pra mim virar né.* - (I): *Ah com certeza* (A2 (18) f. 2005 SA ZL: 2.31).

(345) - (A12): *Meus tios tinha uma oficina de costura né.* - (I): *Mhm* (A12 (39) m. 1990 D. ZL: 1.24).

(346) - (A18): *Porque um fica estressado chega em casa não tem com quem falar ((1,2s)) éh ((0,3s)) o stress continua né.* - (I): *Sim* (A18 (30) m. 2004 P. Ce: 14.33).

There are many examples, however, where I find *né* even at the beginning of the turn as in (347) and (348). These could either be interpreted as a deferred turn-taking-signal, because the interviewer was too quick to give her rapport, or the informant wants to confirm once more what he had said, getting himself time to think about what to say next. The breaks after the *né* in both sentences might confirm the second interpretation:

(347) - (A18): *Eu acho que é o maior problema que tem o boliviano.* - (I): *Mhm.* - (A18): **Né** ((0,5s)) *acho que pelo fato de ser humilde que ele não cresce* (A18 (30), m. 2004 P. Ce: 9.58).

(348) - (A19): *Dezesseis horas por dia eles trabalham né [...] por produção.* - (I): *Sim.* - (A19): **Né** ((0,6s)) (A19 (42) m. 1988 P. Ce: 7.33).

Additionally, *né* occurs at the beginning of the turn in connection with other markers as *mas* or *aí*. The speaker of (349) even uses the combination *desculpa né mas* as a discourse marker to introduce a new topic.

(349) **Desculpa né mas** ((0,4s)) *es a diferença né la gente do campo nó entende muitas coisas* (A14 (28) m. 2006 S. ZN: 8.57).

(350) **Aí né** agora tô aqui ((1,6s)) *sin saber que que eu faço* (A11 (20) m. 2002 S. Pe: 2.41).

In other examples, *né* is found in every position of the sentence: in the beginning of the turn as in (351), in the middle of the verbal syntagm positioned between the verb and the additional proposition + infinitive as in (352), or between noun and adjective in the middle of the nominal syntagm as in (353).

(351) *Tudo né tudo muda no outro país* ((A15 (28) m. 2000 D. Ce): 1.43).

(352) *Minha esposa fazia uns bolos gostoso aí falei 'vamo tentar na feirinha porque os patrício também gosta né de comer bolo'* (A12 (39) m. 1990 D. ZL, S. 6.07).

(353) *A minha mae deixou né tudo os documentos né assinados tudo* (A11 (20) m. 2002 S. Pe, S. 2.13).

(354) *A gente né ficou né::: éh impressionado né do que ele falou né* (A6 (34) m. 1999 S. ZL: 2.11).

In (354) *né* is used after almost every second word. As only the second occurrence in (354) can be interpreted as an instrument of deferral – the long vowel and the following *é* point in this direction – I suggest that my informant uses *né* not only to signal the conclusion of the comment but also to sound as Brazilian as possible. This interpretation might be valid for further Bolivians, too. The overly frequent employment is only found for men, particularly younger men claiming to have regular contact with Brazilians. Interestingly, the only informants who do not employ *né*, or any other Brazilian discourse marker concluding the comment, are women having a low level of education with almost no or very little contact to Brazilians, despite their duration of stay of two years or more.

In São Paulo's PP, the use of *meu* as interpersonal marker is quite common and considered as typical for the CoP of young people in an informal situation, i.e. it has a high degree of social indexicality as (355) underlines:

(355) *Oi, meu! [...] Escute, meu. [...] É mole, meu, só tem parwana e mina na jogada* (Preti 2004: 90).

This discourse marker though, is found only once in the whole corpus:

(356) *Tem que ter uma foto meu* (A14 (28) m. 2006 S. ZN: 14.20).

This could either mean that the Bolivians do not employ this marker at all or that they do not use it in the discourse situation of the interview, with a woman as interlocutor recording the conversation. My observations would rather confirm the second option because I heard young male Bolivians use *meu* when speaking Portuguese with other men of their age.

#### 6.2.14 Common Expressions

---

Under Common Expression I subsume terms and fixed combinations of words which are employed for pragmatic reasons as (most of) their semantic content has been lost. Preti (2004) calls this vocabulary *gíria*<sup>151</sup> (swear word) defining it as

*“um recurso simples para aproximar os interlocutores, quebrar a formalidade, forçar uma interação mais próxima dos interesses das [...] pessoas que dialogam.”* (Preti 2004: 65)

This definition clearly describes the process of accommodation. Therefore, I would interpret occurrences of Common Expressions as signs for accommodation. As Common Expressions are interpreted exclamations like *graças a deus* or *gente no céu* employed to express relieve or surprise without really referring to the original religious semantics to *puta que pariu* or *filho da mãe*, swear words whose use is still socially restricted, or single words like *mano* or *cara* as forms of personal address employed to strengthen the group identity (cf. Ferreira Coelho 2006). Many Common Expressions that are considered as swear words still underlie a certain social tabooization. This means that they are limited to strictly orally conceptualized discourses and their use is generally associated with the style of young men. Women employing words like *foda* or *merda* are only socially accepted in certain groups where their speech style represents part of their identity-building as young modern women who do not care about social restrictions.

Thus, similar to the findings regarding the discourse marker *meu*, in my corpus, there are only three examples for this type of Common Expressions:

---

<sup>151</sup> According to Preti (2004: 66; cf. also l.c.: 111), “[...] *gíria* [é] um fenômeno tipicamente sociolinguístico, que pode ser estudado sob duas perspectivas: [...] *gíria de grupo* [e] *gíria comum*.” He elaborates that the difference between the two types lies in the divulgation in the population; while *gíria de grupo* is only used by a certain group, often youngsters or criminals. It becomes *gíria comum* when adopted by the rest of the society, also called *vocabulário popular*.

As these terms seem rather vague and explain little the mechanisms of how words become *gíria* in the first place, let alone *gíria comum*, I prefer the more neutral term *Common Expression* that underscores the importance of frequency. In my opinion, the reference to different social groups is better described by the additional characteristic of social indexicality.

(357) *Entó:::n você sempre **una foda** né teu amigos que fal... a **indo de merda** (A14 (28) m. 2006 S. ZN: 9.42).*

(358) ***Na porrada** gente de que tratou. Ele se enriqueceu né aos recurso de nos (A6 (34) m. 1999 S. ZL: 6.34).*

(359) *E o **filho da mãe** né fiz isso com nos (A6 (34) m. 1999 S. ZL: 6.44).*

Interestingly, they come from the two speakers who also used *meu* as a discourse marker (cf. (356)) and *né* in a particularly high frequency (cf. (354)).

(360) shows a Common Expression that Preti (2004) considers “um vocabulário gírio”, explaining at the same time that this expression

*já pertence ao léxico popular, integrou-se ao uso oral do falante comum e tornou-se, [...] normal na mídia (l.c.: 65).*

This means that the second order social indexicality of this expression, indicating a certain social group as the former examples, has been lost completely. Similar to this example there are occurrences regarding personal interaction like *malandro* or *cara* that are common in spoken Portuguese without indicating any specific social group, but are merely indicators of orality:

(360) ***Dei uma bronca** (A9 (42) m. 1994 D. ZL: 8.55).*

(361) ***O cara** passou direito ni falou ni cumprimentou (A16 (32) f. 2004 P. ZL: 8.39).*

(362) *Roubaram tudo **esse cara** aí (A11 (20) m. 2002 S. Pe: 5.16).*

(363) *Aqui em Brasil ((0,6s)) aqui em Brasil também ((0,9s)) éh tem muito muito muito **malandro** aqui (A10 (22) m. 2006 S. Pe: 4.19).*

Further Common Expressions that were once socially restricted as blasphemy but have long been integrated into the daily use of São Paulo’s speakers are expressions with religious references as found in (364) to (368):

(364) ***Se deus quiser** no ano que vem tô querendo meu... ter meu a minha própria [oficina] (A11 (20) m. 2002 S. Pe: 7.28).*

(365) *Então::: **se deus quiser** daqui para frente já todo legalizado ((0,6s)) mudar para o sul (A18 (30) m. 2004 P. Ce: 13.17).*

(366) ***Ah meu deus** (A23 (45) f. 1990 Ph ZO: 6.13).*

(367) ***Ixi**<sup>152</sup> **Maria** (A10 (22) m. 2006 S. Pe: 8.26).*

(368) ***Vix** ((0,3s)) eu comí ((0,8s)) éh lapin (A17 (31) m. 2007 P. Ce: 9.05).*

While *se deus quiser*, *meu deus (do céu)* and *Vixe (Maria)*, are only found for some speakers, *graças a deus* is common for many speakers, particularly those having a higher social status than the seamstresses and having been to São Paulo for more than five years or working as academics outside the textile industry:

---

<sup>152</sup> This interjection, known as *Vixe Maria* (originating from *Virgem Maria*) is regularly pronounced as [ˈviʃiˈma.ria], but most of the time abbreviated as [ˈviʃ(i)] or [iʃ(i)] like in my examples.

- (369) *Bom graças a deus com minha pessoa nunca tive discriminação (A13 (33) m. 2000 D. ZL: 6.48).*
- (370) *Ah teve que voltar abriu uma clínica com colega dele né ((1,2s)) graças a deus minha ((0,3s)) minha parte assim pessoalmente meus irmãos todos são profissionais (A19 (42) m. 1988 P. Ce: 5.21).*
- (371) *Graças a deus a gente con a que yo moro tenha auto [...] graças a deus (A22 (27) f. 2008 P. ZS: 18.25).*

In addition to these examples, there are many periphrastic expressions, particularly with the verb *dar*, that are highly typical for Brazilian Portuguese. This finding underscores my impression that it is a strategy of the Bolivian language learners to include the complete Common Expression into their speech because it makes it easier for them to avoid grammatical errors. Supporting this argument is the fact that the phonetic realization is also strictly Brazilian Portuguese: In (375) and (376), for example, there is the elision of the first vowel in *para* resulting in *pra* that is common for PP, but not for my corpus. In (376), the additional *um pouco* is inconsistent with the standard use of the expression – a further indicator for the hypothesis that the speakers insert the expressions mainly to accommodate toward Brazilian Portuguese, even though they do not have the knowledge how exactly to use them.

- (372) *E não deu certo não deu certo (A25 (60) f. 1989 S/SA ZL: 1.08).*
- (373) *Ay tomará que tudo dê certo lá acho que vai dar certo sim (A18 (30) m. 2004 P. Ce:13.23).*
- (374) *Agora o pessoal vem aqui ((0,3s)) outro se dá bem outro se dá mal. ((0,3s)) Graças a deus eu vou falar que não que me dei bem graças a deus (A19 (42) m. 1988 P. Ce: 17.45).*
- (375) *Mais ou menos. Dá pra mim virar né (A2 (18) f. 2005 SA ZL: 2.31).*
- (376) *Mas tem brasileiro que são ((0,7s)) gente boa e como eu falei gente ruim que não gostam de gente estrangeira de outro país, mas ((0,8s)) mas dá pra agüentar um pouco (A10 (22) m. 2006 S. Pe, S. 10.05).*

In contrast to the first Common Expression seen as socially indexical for young men I find these periphrastic expressions for almost every speaker, regardless of social variables like age, sex, profession or duration of stay.

### 6.2.15 Code-Mixing

In accordance with the terms of Muysken (2000) as presented in chapter 3.2.1.5, three different basic processes of code-mixing are distinguished: *insertion*, *alternation* and *congruent lexicalization* (cf. l.c.: 3). Alternation is the process regularly described by the term code-switching:

*In the case of alternation, there is a true switch from one language to the other, involving both grammar and lexicon (l.c.: 5).*

In intra-sentential code-mixing, the only type considered by us, alternation occurs mostly at the peripheral position of the clause and at major clause boundaries. In my corpus, there are only few examples for this process:

- (377) *Es una es una hente que:: e e falam eles que são explo... explorados (A22 (27) f. 2008 P. ZS: 8.04).*
- (378) *Como tanto trabalhando trabalhando eu ma: algunas vezes me siento muy mal también de mi (?) de mis pés (A25 (60) f. 1989 S/SA ZL: 3.51).*
- (379) *Porque eles chaman de macaquinho né [...] e vocês ((0,5s)) llaman de hermanito (A1 (23) m. 2006 SA Ce: 7.33).*
- (380) *Muita gente vem de La Paz né e nosotros siempre somos orgullosos de ser del Collasuy (A14 (28) m. 2006 S. ZN: 0.43).*
- (381) *Foi aumentando um pouco más a población boliviana y también de otros países (A7 (32) f. 1994 S. Ce: 12.04).*

Out of these few examples, only (377) starts with Castilian and terminates with Portuguese, the other examples, (378) to (379), show the reverse order. In the first two examples, the switch occurs after an elongated vowel in *que* and *ma* functioning similar to the break in (379) as the switch site. While the structure of the two first examples is interrupted at this point, in (379) structure is not interrupted but the alternation just occurs in the middle of the sentence between noun and verb, comparable to (381), where the verbal syntagm in Portuguese is followed by the nominal syntagm in Castilian. Here is no paratactic structure, though, but only one main clause in whose course the alternation takes place. This could indicate that even in this sentence, the underlying process is rather congruent lexicalization than alternation: As nasals are often realized according to the equivalent Castilian vocabulary by my informants, the second part of (381) could also represent Castilian vocabulary with Brazilian phonology.

The examples (382) to (384) also show ambiguous cases that could either be attributed to congruent lexicalization or to insertion. (382) might even be a case similar to (381), i.e. the idiomatic expression *no todo lo que brilla es oro* could constitute a case of alternation, accompanied by the PP discourse marker *né*. The various Castilian function words in the former sentence, however, point rather in the direction of congruent lexicalization while the idiomatic expression is inserted into the Portuguese matrix language.

- (382) *Entonce eles también de dessa parte devem ver no, s... ((0,7s)) devem olhar de uno mas ((0,6s)) yo acho que se:: lleve muito por las apariencias. Porque nó todo lo que brilla es oro né (A22 (27) f. 2008 P. ZS: 20.51).*
- (383) *Eu tá aproveitando lo que hay, lo que tem Brasil, lo que tá me oferecendo. Eu estou aprovechando el máximo (A26 (25) f. 2005 S. ZL: 0.22).*

(384) *Né então **yo estoy de acuerdo** então fazer uma coisa **bien** típica ((0,3s)) de nossa reg éh **región** né que es a Bolívia (A23 (45) f. 1990 Ph ZO: 9.52).*

(385) *A gente **viño mi pai mi mãe mi una** filha. Aqui **tuve** tres menino **ya** (A3 (34) f. 2004 SA ZL: 0.59).*

(383) and (384) show a similar pattern as (382). (385), though, is either another case of congruent lexicalization or the insertion of Portuguese nouns into the Castilian matrix language as it is clearly the case in (386) to (389).

Interestingly, in (386) and (387) the insertion occurs as a translation after the speakers already has given the Bolivian Castilian equivalent, i.e. *pareja* and *namorada* as well as *borracho* and *bebado*. Here, it seems that the speakers are conscious of using the other language, but only with regard to these two words.

(386) *Eso me há revelado lo más (?) mi relacionamiento de mí con mi propio **pareja** con mi propio **namorada**. [...] Un día le dije ‘te quiero y quiero llegar lejos contigo [...] (A24 (25) m. 2004 S/SA Ce: 14.08).*

(387) *Siempre se encuentra con la mala hierba como por exe mm **borracho** como aqui llama **bebado** (A14 (28) m. 2006 S. ZN: 7.48).*

(388) *Cuando el ladrón assalta una persona todo esta estas personas lo pegan al ladrón lo pe... **baten** chotan lo **baten** lo chotan hasta matan [...] En allá en **de noite** cuando assaltan de assaltan las casas ((0,9s)) a casas [...] lo pegan al ladrón ((0,5s)) lo ((0,3s)) lo matan lo penduran en un poste (A25 (60) f. 1989 S/SA ZL: 10.38).*

(388) is a further example of the insertion of Brazilian elements into the Castilian matrix language; here the prepositional syntagm *de noite*. Additionally, there is the insertion of the verb *bater*, although the morphology stays in Castilian. Interestingly, the speaker first uses the verb *pegar*, a well known so-called false friend, signifying *to beat* in Castilian, but *to catch* in Portuguese. In (387), it seems as though the speaker realizes this fact while she articulates *lo pegan* once more. She hence discontinues the word and substitutes it by the Portuguese equivalent *bater*.

Apart from the typical noun insertion in (389), there are further verb insertions in (390) and (391), also in the direction from Portuguese to Castilian. The latter example is rather important, because it constitutes one of the few words that the speaker utters in Portuguese; the rest of her responses are either in CP or not comprehensible, although she has spent more than two years in São Paulo.

(389) *Hay momentos que ya ((0,5s)) é:: éh dá **trocado** pero dá **trocado** de mala forma (A26 (25) f. 2005 S. ZL: 8.05).*

(390) ***Temos** pero nó está registrado no **temos** registrados (A28 (28) m. 2002 S. ZL: 9.45).*

(391) *Sólo yo vendo nada más no **trabalhei** aquí (A27 (35) f. 2006 SA ZL: 1.35).*

These cases of insertion into the Castilian matrix language are limited to some speakers who display least proficiency in Portuguese, although there are other informants who have spent less time in São

Paulo. All of them have come to Brazil for economic reasons and their contact to Brazilians is very rare.

Alternation, on the other side, is also found for speakers of high educational levels working outside the textile industry. One of them, though, had been in São Paulo for only three months at the time of the interview, while the other one has regular contact to Bolivia, returning there regularly.

For the other informants, though, I assume that the matrix language is not Bolivian Castilian but has changed to Brazilian Portuguese, displaying an “overall shift from one matrix language to another one” (Muysken 2000: 247.).

(392) to (394) show the insertion of a verb form as well as of nouns and an adverb. In a completely Portuguese context, only the alien vocabulary is inserted into the structure of the sentence without phonological adaptation toward Brazilian Portuguese.

(392) *O trabalho que eles fazem não **denigra** mas eu não gostaria que eles ficassem todo tempo na máquina (A9 (42) m. 1994 D. ZL: 5.02).*

(393) *Objetivo é eh trabalhar em equipe né, fazer um:: uma **associación** para poder nos ajudar. Por quê? ((0,6s)) Porque a gente um dia você está **arriba** outro dia pode cair né (A13 (33) m. 2000 D. ZL: 5.34).*

(394) *Foi totalmente uma experiência muito **horrible** né (A6 (34) m. 1999 S. ZL: 3.36).*

Despite these examples of alternation and insertion, congruent lexicalization clearly constitutes the process playing the most important role in the bilingual speech of my informants. According to Muysken (2000), this process

*is akin to language variation and style shifting: switching is grammatically unconstrained and characterizable in terms of alternative lexical insertions (l.c.: 221).*

Considering that in the case of congruent lexicalization, the two languages are simultaneously active (Muysken 2000: 252), the composition of (395) becomes clear: There are the PP personal pronouns *ele* and *a gente*, the discourse marker *né* and many elements that are equivalent in both languages contrasting with the Castilian adjective possessive *su*, the verb *buscar* (in the Castilian meaning *to look for*), and *pero* accompanied by the Portuguese equivalent *mas*. Particularly interesting is the verb form *fizo*, a combination of Castilian morphology and Portuguese vocabulary. Thus, in (395), Castilian structure and vocabulary overweighs the Portuguese elements.

(395) *Só que ele sabe **su** consciencia que ele está matando né ele tá **buscando** para **dialogar pero dialogar** a gente **va** sentar para **dialogar** ((0,4s)) **pero** mas de ele **lo que fiz**o comigo ele **nó** tem adiante. (A24 (25) m. 2004 S./SA. Ce: 10.30).*

Comparing (396) and (397), the latter seems to display the equality of both languages, sometimes even pointing in the direction of alternational code-mixing since they contain whole CC constituents like *también voy a ir hasta plaza Kantuta* or *el otro está cubriendo a este*. These constituents, though, are embedded in Portuguese dominated speech:

(396) *Éh:: yo quando vim pra aqui só em essa parte conheço bol... a los bolivianos y también voy a ir hasta plaza Kantuta mas más para frente quando yo vou pra minha casa graças a deus a gente con a que yo moro tenha auto (A22 (27) f. 2008 P. ZS: 18.12).*

(397) *Aqui agarra você ((0,7s)) coloca contra parede apuntado dois armas uno uno como posso falar. Una especificación la gente se ubica talvez (para isso?) no, mas ((0,4s)) uno está aqui ((0,5s)) y el otro está cubriendo a este porque se você reacciona ((0,5s)) nó te vai pagar atirar que está atrás de você mas que está a frente de você cê pode atirar (A14 (28) m. 2006 S. ZN: 4.07).*

For speakers with high educational level and high proficiency in Portuguese I only find single switches, mostly concerning only one word or morphological form as in (398) and (399).

(398) *Na Bolívia está dividida agora. [...] Sempre estuvo dividida. [...] Então acho que é a cultura boliviana mesmo o jeito de vivir o jeito de fazer as coisas [...] deles (A17 (31) m. 2007 P. Ce: 7.53).*

(399) *Como todo todo estrangeiro quando va para outro país toda pessoa quando va para outro país ((0,5s)) ele é o o branco das pessoas (A18 (30) m. 2004 P. Ce: 10.51).*

As these switches concern function words I rather regard them as congruent lexicalization than insertion.

Considering the last examples as well as e.g. (382) to (384), I suggest that congruent lexicalization is one of the most dominant processes regarding the speech of my informants. Because of the linear and structural equivalence between the varieties as well as the conformance of most of the vocabulary, many sentences could differ only with regard to their phonological appearance. In accordance, the phonological findings attributed to L1-transfer could also reflect code-mixing as was exemplified by the different realizations of the nasal diphthong [ãɹ] in analogy to the linguistic context of the lexical equivalent in Castilian. Therefore, in the case of two languages as narrowly connected as Spanish and Portuguese, congruent lexicalization could be the process that described L1-transfer most accurately. As congruent lexicalization is explicitly associated with typologically closely related languages sharing at least partly the grammatical structure, and even varieties of one language (cf. Muysken 2000: 122 ff.), the findings of my study confirm Muysken's classification.

### 6.3 Summing up

---

The corpus was analyzed with respect to 15 different features of the various linguistic fields. Summing up my findings, I will focus on the distribution of the features and the different linguistic processes that I attribute the different phenomena to. Some of them concern almost all speakers, others only small groups. Therefore, I will unite the features in one table, going from phonology to vocabulary and from more frequent phenomena to the rarer ones.

Table 6.6: Distribution of investigated features and corresponding linguistic processes

Linguistic feature	Distribution among the speakers	Underlying linguistic processes
Variation between open and closed /ɛ/ and /e/ resp. /ɔ/ and /o/	All speakers	L1-transfer / Congruent lexicalization
No phonemic distinction between /s/ and /z/	All speakers	L1-transfer / Congruent lexicalization
Variation between /e/ and /i/ resp. /o/ and /u/	All speakers	Accommodation to BP vs. L1-transfer / Congruent lexicalization
Variation concerning the nasal diphthong, realization according to Castilian equivalent	All speakers	L1-transfer / Congruent lexicalization
Diphthongization before final /s/	All speakers donos and (A28 (28) m. 2002 S. ZL) → hypercorrection	Accommodation to BP
Palatization of -ti/di	All speakers; Most frequent: academic women, male seamstresses (→ hypercorrection); Least frequent: academic men, some seamstresses	Accommodation to BP
Aspiration of /t/ in the syllable onset	All speakers; Most frequent: academic women, male seamstresses Least frequent: academic men, some seamstresses	Accommodation to BP
r-caipira in the syllable coda	Young physicians, A14	Accommodation to certain style / variety of spoken BP
Elision final /s/ and absence of number agreement	All speakers but A20, A21, A13. Most regular: A11, A6 (→ hypercorrection)	L1-transfer / Congruent lexicalization vs. accommodation to PP; simplification
Omission, confusion of preposition	All speakers	L1-transfer (CP) / Congruent lexicalization; simplification
Absence of verbal agreement, generalization of 3 <sup>rd</sup> pers. sg	All speakers but A20	Accommodation to PP; simplification
Gender attribution and absence of gender agreement	All speakers but A16 and A20	L1-transfer (CP) / Congruent lexicalization; simplification
<u>Pronomina</u>		

employment of <i>você</i> and <i>a gente</i>	All speakers	Accommodation to BP spoken in São Paulo
employment of <i>a gente</i> only for <i>we</i> , for generalizing expressions <i>o pessoal</i>	Academics but A22	Accommodation to BP spoken in São Paulo
<i>Nós</i> (instead of <i>a gente</i> )	Some seamstresses, A5  Male academics	L1-transfer / Congruent lexicalization  nonaccommodation to PP
<i>Cê</i> (instead of <i>você</i> )	Men who have been to SP more than six years, A17, A18	Accommodation to to BP spoken in São Paulo
<i>tú</i> (+ 3.pers.sg.)	A8, A24	Accommodation to PP
Word order	A25, A3, A14	L1-transfer (CP) / Congruent lexicalization
<u>Discourse marker</u>		Accommodation to PP
<i>né</i>	All but A26, A27, A8, A28. Most frequent: A14, A6, A11 (→ hypercorrection).	
<i>meu</i>	A14	
<b>Socially tabooed Common Expressions</b>	A14, A6	Accommodation to PP
<u>Code-mixing</u>		L1-transfer
<b>Insertion in Bolivian Castilian</b>	A27, A25, A24, A28	
<b>Alternation</b>	A1, A22, A23, A24	

The pattern underlying this distribution is quite complicated. There are features that all speakers realize despite the social differences among the group. Thus, they must be due to linguistic processes such as L1-transfer (probably caused by the difficulties of perceiving the difference between Castilian and Portuguese as described by the Speech Learning Model) and congruent lexicalization, or various factors operating simultaneously in the same direction. For example, the absence of number agreement: the phenomenon is present in CP, and also in CC as a reduction of the final /s/. It is also highly salient in PP; additionally, the reduction of redundant morphology constitutes a case of morphological simplification typical for language change in an oral context.

The other features, however, that vary strongly among the speakers must be due to social factors. While some can be attributed to L1-transfer of the different Castilian varieties, most of them can be classified as accommodation towards different social groups in São Paulo.

In order to prove that these differences are not due to gender, age, duration of stay or proficiency, I further present three examples comparing speakers who correspond with regard to these social features but still differ strongly with regard to their Portuguese:

Example I:

A27 is female, 35 years old, and has spent two years in São Paulo at the time of the interview. A20 is also female, 27 years old, and has also spent two years in São Paulo:

In the speech of the former, there is only one unambiguous Portuguese word, *trabalhei* (391), an insertion into the CP matrix language. Although she speaks no Portuguese apart from this word<sup>153</sup>, features like the absence of nominal number and gender agreement are found, which obviously also occur in CP.

In the speech of the latter, though, there is no code-mixing at all and almost no particularities in the lexical or morphosyntactic field. She is one of the few that realize all agreement rules according to PC. Only the phonological features common for all reveal that she is no Brazilian.

Thus, neither sex nor duration of stay explains the differences; the first one is a sales assistant in the Bolivian streets of Brás, the second one a physician studying at São Paulo's university. Correspondingly, at least the profession and level of education must be taken into account here.

Example II:

A28 is male, 28 years old and has spent six years in São Paulo working as a seamster in the textile industry. A14 is male, 28 years old and has spent only two years in São Paulo, also working as a seamster. Both come from La Paz/El Alto.

While the first shows the insertion of verbs into the Bolivian Castilian matrix language, employs no discourse marker *né*, rarely aspirates /r/, and palatizes *-ti/di*, the second one uses *né* overfrequently, and even employs *meu*, displays examples of hypercorrection with regard to palatization and fre-

---

<sup>153</sup> Naturally, this made the conversation with her almost impossible, particularly as it is held in a public place with a lot of noise in the background. The failure of the interview, however, appeared only with the transcription where I realized that she had told me to leave but I continued with the interview, obviously not having understood what she had said.

quent aspiration of /t/, although he also shows a high level of congruent lexicalization as well as transfer from CP, e.g. in the case of SOV word order.

Thus, the features that are associated with accommodation to PP are not connected with duration of stay, profession or level of education. I therefore suggest that the dissimilar translingual social networks, in this case formed by different attitudes, make the difference: While the first one is oriented toward a mainly Bolivian network, the second one tries to accommodate toward his Brazilian contacts even though they are rare.

Example III:

A25 is female, 60 years old, and has spent almost twenty years in São Paulo. She has worked as a seamstress and currently works at the Bolivian market selling goods with her family. A21 is male, 62 years old, and has spent more than 40 years in São Paulo. He is an engineer and a voluntary helper at the Bolivian organization ADRB. The former shows many examples of code-mixing, insertion, alternation and congruent lexicalization. Additionally, examples of SOV word order, CP discourse markers, and the absence of gender agreement can be found. But at the same time, she regularly realizes palatization and aspiration and uses only *a gente*. The latter shows no code-mixing at all, there are few cases of missing verbal and gender agreement, but none for the absence of number agreement. He rarely uses the palatized *ti/di* and realizes /t/ in the syllable onset as [R], though. Thus, the phonological accommodation is not necessarily connected with high level of proficiency or young age; again I would argue that the second speaker non-accommodates to the phonological features of PP because due to his social status he gets no pressure from his network to do otherwise.

Correspondingly, I deduce that there are roughly three different social groups whose speakers display more or less the same linguistic features:

The first group is constituted by those speakers for whom many examples of L1-transfer and code-mixing and only few cases of accommodation to the Brazilian varieties can be observed.

The second group represents those speakers who show high degrees of accommodation in terms of phonology, morphology and vocabulary, particularly to PP features. PP discourse markers and Common Expressions are particularly frequent and so are morphological simplification and generalization of standard distant forms.

The third group shows few transfer phenomena, particularly at the phonological level, but most standard morphological features; the absence of agreement is rather rare.

Therefore, I assume that the first group is rather oriented toward the Bolivian side of its translingual network. The second group is oriented toward the Brazilians, particularly PP, since they have no access to educational institutions but seek the approval of their contacts of the receiving society, obviously speakers of PP. The third group is oriented toward PC, because they learn Portuguese at universities; since their social status is high, the social approval is not necessarily dependent on linguistic accommodation. This would mean that different accommodation strategies by different socio-indexical features are employed by the various groups. Whether these strategies are successful must be evaluated by the receiving society.

Thus, to verify or falsify this interpretation, I asked the *Paulistanos* about their opinion regarding the Bolivians' speech, particularly about their linguistic and social classification.

## 7. The evaluation of the Bolivians' speech

---

In this chapter the results of the analysis of one hundred responses to an online questionnaire by *Paulistanos* are presented. The questions on the speech examples of three Bolivian immigrants concern the linguistic, social and personal evaluation of the speakers. The majority of the *Paulistanos* do not identify the speech of Bolivians, but most of them recognize S I and III as being influenced by Spanish or at least other languages. S II, however, is mostly estimated as representing the informal variety PP. The social characterization revealed that there is no negative stereotype connected with S I and S III. The evaluation of S II, though, displayed a clear correlation of the speech with poverty and little education. With regard to particular linguistic features identified by the participants as important for the linguistic and social evaluation, they enumerated the employment of discourse markers and Common Expressions by S II and the phonological markers reflecting the Castilian influence by S I and S III. While linguistic and social results are coherent, the personal evaluation differs from the other fields and indicates no clear tendencies. Correlation tests of the results according to sex, age group and social status of the participants reveal little significance. Thus, it seems that the evaluation of the speech examples is quite homogeneous in São Paulo's society.

---

As was mentioned in section 5.4, the distribution of the sample is not representative for São Paulo's society with regard to age, profession and knowledge of foreign languages, because the self-selection conditioned by the online-questionnaire obviously favored the participation of young academics knowing foreign languages. Thus, the results of the study are not valid enough to be generalized for the whole of *Paulistanos*<sup>154</sup>. Nevertheless, the findings should be interpreted as tendencies for the societal perception of Bolivian speakers, having in mind the biased distribution of the sample.

Based on Carvalho (2008), my study was also divided in three parts: The three speech examples were accompanied by questions regarding the linguistic, the social and the personal evaluation of the three speakers. The speech examples were selected because of their prototypicality concerning the different groups, differentiated after the qualitative analysis of the corpus:

---

<sup>154</sup> When distributing the questionnaire I asked for the participation of people which currently live in the metropolitan area of São Paulo and consider themselves Paulistanos, in order to exclude people living in São Paulo for only a short time or having moved there recently, although I have no direct means to obtain this information.

Table 5.1: Characteristics of speech examples used for the questionnaire

	S I	S II	S III
Age at time of recording	25	26	28
Duration of stay in SP	2 years	11 years	7 years
Profession	seamstress	seamstress	university student
Social Network	dominated by Bolivians, quite closed, multiplex	dominated by Brazilians ( <i>Zona Leste</i> ), rather open, rather uniplex	dominated by Brazilians (University), open, rather uniplex
Level of Portuguese	advanced	fluent	fluent
Linguistic characteristics according to the underlying processes as suggested in Chapter 6.	<p><b><u>L1-transfer / congruent lexicalization</u></b></p> <ul style="list-style-type: none"> <li>- Denasalation of nasal diphthongs: e.g. <i>leão</i> [le'ɔ̃]</li> <li>- Voiced sibilants pronounced voiceless: e.g. <i>preso</i> ['preso]; <i>ajuda</i> [a'ʃuda]</li> <li>- No aspiration of vibrants: e.g. <i>agarra</i> [a'gaɫa]</li> <li>- verb morphology: e.g. <i>vivir</i> (Ptg. viver).</li> <li>- Word order: <i>Aí o rato agarra éh leão deixa vivir um pouco más.</i></li> <li>- Insertion of Castilian vocabulary: <ul style="list-style-type: none"> <li>o <i>más</i> (Ptg. mais)</li> <li>o <i>mano</i> (Ptg. mão)</li> <li>o <i>caminar</i> (Ptg. caminhar)</li> <li>o <i>durmir</i> (Ptg. dormir)</li> </ul> </li> </ul>	<p><b><u>L1-transfer / congruent lexicalization</u></b></p> <ul style="list-style-type: none"> <li>- “false” gender attribution: <i>do árvore</i> (Ptg. da árvore).</li> <li>- <i>ratozinho</i> (Ptg. ratinho)</li> </ul> <p><b><u>Learner variety</u></b></p> <ul style="list-style-type: none"> <li>- “wrong” pronunciation: <i>galho</i> ['galɔ] (Ptg. ['gaɫɔ]).</li> </ul>	<p><b><u>L1-transfer / congruent lexicalization</u></b></p> <ul style="list-style-type: none"> <li>- nasal diphthongs: e.g. <i>leão</i> [le'ɔ̃].</li> <li>- voiced fricatives: e.g. <i>vida</i> ['bida].</li> <li>- verb morphology: e.g. <i>vivir</i> (Ptg. viver).</li> <li>- Insertion of Castilian vocabulary: <i>durmir</i> (Ptg. dormir)</li> </ul>

	<p><u>Accommodation to spoken Portuguese</u></p> <ul style="list-style-type: none"> <li>- Common Expression: <i>feliz da vida.</i></li> </ul>	<p><u>Accommodation to spoken Portuguese</u></p> <ul style="list-style-type: none"> <li>- Affrication of occlusive: e.g. <i>sorte</i> ['sortʃi].</li> <li>- Diphthongization: e.g. <i>mas</i> [mais].</li> <li>- Discourse marker <i>né</i></li> </ul> <p><u>Accommodation to PP</u></p> <ul style="list-style-type: none"> <li>- Common Expressions: <ul style="list-style-type: none"> <li>o <i>maluco</i></li> <li>o <i>pateta.</i></li> </ul> </li> <li>- Discourse marker <ul style="list-style-type: none"> <li>o <i>né</i></li> <li>o <i>meu</i></li> </ul> </li> </ul>	<p><u>Accommodation to PC</u></p> <ul style="list-style-type: none"> <li>- Hyperpronunciation of oral diphthongs: e.g. <i>poupar</i> [poʊ'par] (spoken Ptg. [po'par])</li> <li>- Use of subjunctive as negative imperative 3rd person: <i>Não me mate!</i> (PP <i>Não me mata!</i>)</li> <li>- Use of Future Subjunctive: <i>estiver precisando.</i></li> <li>- Clitic pronoun: <i>lhe.</i></li> <li>- Word order: <i>aproximou-se</i> (PP <i>se aproximou.</i>)</li> <li>- Common Expression <ul style="list-style-type: none"> <li>o <i>Certa vez na floresta.</i></li> </ul> </li> </ul>
--	---	--	---

Obviously, the example of Speaker I (S I) is highly influenced by L1-transfer. The others show only few examples of L1-transfer, although S III – despite the linguistic features that I would interpret as accommodation toward PC – displays some features that are also found in the speech of S I, e.g. the pronunciation of *leão* as [le'ɔ̃] or the use of *vivir* instead of *viver*. Considering Speaker II (S II), its main characteristic clearly lies in the concentration of colloquial regular expressions and discourse markers like *né*. As pointed out before, in literature, the expression *meu* is held as highly socio-indexical and associated with the colloquial style of young people. Speaker III (S III), on the contrary, differs from the other examples because of its formal style, particularly with regard to its morphology: The use of the 3<sup>rd</sup> person subjunctive to express the negative imperative and the clitic *lhe* are obviously correct. In Brazilian Portuguese, however, these forms are widely associated with written PC; in oral discourses they seem quite remarkable, as the pilot studies underscored. Therefore, S I was selected to represent the speakers having little language proficiency and displaying most direct influence of Castilian markers, for example the frequent insertion of Castilian vocabulary and phonological particularities like the voiceless realization of sibilants. S II represents the group that accommodates strongest to PP, i.e. he uses the discourse marker *né* and *meu*, Common Expressions and phonological markers like the diphthongization and the palatization discussed before. The direct influence of Castilian is small, although there are some particularities of a language learner. S III

shows more Castilian markers, particularly at the phonological level, but on the other hand, his speech example contains morphological structures like subjunctive or clitics typical for PC. As he studies at one of São Paulo's universities, he was selected to represent the group of academics who accommodate toward PC.

With these examples it should be tested if the *Paulistanos* recognized the speech of Bolivians as such and if social stereotypes existed in regard of the immigrants. Additionally, I aimed to know which linguistic markers led to the social evaluation of the speakers and if these correlated with the qualitative analysis of Bolivian immigrants, i.e. if the possible social indexicality of linguistic forms previously identified as typical for the different groups was confirmed.

### 7.1 The linguistic perception of the Bolivians

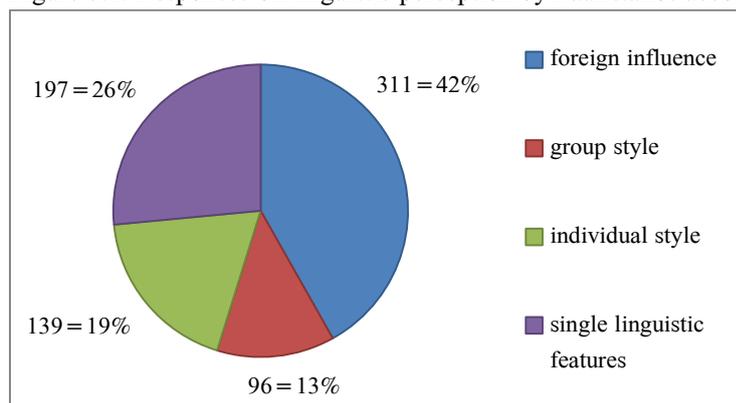
---

In the test, various direct and indirect questions referred to the linguistic perception of the speech examples. The question *Como você descreveria a língua que você está ouvindo? (How would you describe the language you are hearing?)* aimed for the classification of the variety as a whole. By asking: *Tem algum aspecto na fala da pessoa que chamou a sua atenção? (Is there any aspect of the speech of the person that caught your attention?)* I expected to obtain data about whether particular linguistic phenomena were identified by the participants. These open-response questions were complemented by more implicit items, e.g. the informants were asked to give the reason for their choice regarding place of origin, place of residence and – as long as this question was included into the questionnaire<sup>155</sup> – for the jobs they chose for the different speakers. The responses to these questions resulted in almost 750 comments on the linguistic perception of the three speakers by the *Paulistanos* as figure 7.1 shows:

---

<sup>155</sup> As was explained in 5.4.3, the question of subjective reaction in which the informants were asked to choose one job position for the speakers was excluded after the first weeks, because people refused to answer it or stopped answering the questions at precisely this question.

Figure 7.1: Responses on linguistic perception by Paulistanos according to language level.



Most of them concern the labeling of the language and the distinction between foreign languages and Portuguese, but also individual and group style that will be discussed in 7.1.2. 26% of the answers, though, named single linguistic features.

### 7.1.1 Linguistic features

It has already been mentioned that the data concerning particular features stem mainly from the first question accompanying each speech example, asking if any aspect of the example caught the attention of the informant. Naturally, the open-response questions allowed a huge variance of answers. While some people simply answered *não*, others actually referred to particular linguistic features even with reference to linguistic categories like phonology or syntax. Similar to these responses, the question to justify their choice with regard to place of origin or residence only led to the characterization of São Paulo's quarters in some cases, but many took the opportunity to explain which of the linguistic features they were conscious of having influenced their decision. Naturally, the emic perspective of non-linguists cannot result in detailed scientific analysis, but in order to investigate the second order social indexicality of linguistic markers, this view is very important, because it is a product of the perception of the entire society and not only linguists. 197 examples mentioning concrete linguistic markers prove that almost all of the participants are indeed conscious of linguistic features and have the means to express their perception of the particular features, even if not all comments can be interpreted unambiguously. Tested for correlations, it became clear that students rather refer to the speakers' style and offer less reflection about single linguistic features, while academics give most detailed information in this concern. The difference is not significant, though.<sup>156</sup>

<sup>156</sup>Chi-square-tests of *profession\*linguistic characterization (S I, II, III)* show that the residuals between expected and real values on single linguistic markers are always negative for students, while they are positive for academics and non-academics with a higher value found for the academics. Similar tests for *foreign language*

As these are the only open-response questions in the questionnaire and the meta-linguistic data are very important for this study, qualitative data will accompany the relevant statistics. Therefore, in table 7.1, some of the comments are collected to get an idea of how the phenomena are described by the participants:

Table 7.1: Excerpt from the responses: description of linguistic markers.

Linguistic features	Comments <sup>157</sup>
<b>Discourse marker</b>	<ul style="list-style-type: none"> <li>▪ <i>A forma como a pessoa conecta as "partes" da história. Ela repete muito "aí". (The way the person connects the "parts" of the history. It often repeats "aí")</i></li> <li>▪ <i>Presença das expressões "meu", "né". (Presence of expressions "meu" and "né")</i></li> <li>▪ <i>Pelos marcadores conversacionais próprios. (Through the actual conversational markers.)</i></li> </ul>
<b>Common Expressions</b>	<ul style="list-style-type: none"> <li>▪ <i>A expressão "dar uma mão" é bem típica nossa. (The expression "dar uma mão" is very typical for us.)</i></li> <li>▪ <i>"Maluco", "...o leão catô ele". Gírias e composições gramaticais "pobres". ("Maluco", "...o leão catô ele". Swear words and "poor" grammatical compositions)</i></li> <li>▪ <i>As expressões usadas. (The expressions used.)</i></li> <li>▪ <i>Uso de gírias, mano<sup>158</sup>. (Use of swear words, mano)</i></li> </ul>
<b>Insertion of Castilian vocabulary</b>	<ul style="list-style-type: none"> <li>▪ <i>Espanha pelo uso da palavra "viver". (Spain because of the use of the word "viver")</i></li> <li>▪ <i>Pelo final que usa a palavra espanhol mano. (At the end, as it uses the Spanish word "mano").</i></li> <li>▪ <i>Algumas vezes as palavras são em espanhol. (Sometimes the words are in Spanish.)</i></li> </ul>
<b>Intonation</b>	<ul style="list-style-type: none"> <li>▪ <i>A cadência da fala. (The cadence of the speech.)</i></li> <li>▪ <i>Tom aberto. (Open tone)</i></li> <li>▪ <i>O som de língua latina. (The sound of the Romance language.)</i></li> <li>▪ <i>Uma entonação muito forte, diferente. (A very strong and different intonation.)</i></li> </ul>
<b>Realization of certain phonemes</b>	<ul style="list-style-type: none"> <li>▪ <i>Dificuldade de falar o "ch", "nh" (caminhando), "j" ("axuda"), dificuldade com o "ão", entre outras. (Difficulties to pronounce "ch", "nh" (caminhando), "j" ("axuda"), difficulty with the "ão", among others.)</i></li> <li>▪ <i>A pessoa troca o e pelo i e vice-versa e substitui o s por ç e o ao por on. (The person changes e for i and vice-versa, and substitutes s (voiced, sn) for ç (voiceless, sn) or ao for on.)</i></li> <li>▪ <i>De novo a fonética, "bibir" "Leon". (Again the phonetics, "bibir", "leon")</i></li> <li>▪ <i>As vogais, além de serem abertas, nota-se que elas são mais longas, típicas do</i></li> </ul>

knowledge, profession, place of residence, age groups and gender\*linguistic characterization (S I, II, III) reveal no significant correlation.

<sup>157</sup> In order to give the most exact impression of the answers, the original quotes should be maintained as they are. I therefore do not correct them orthographically.

	<p><i>sotaque britânico. (The vowels, apart from being open, are noticeably longer – typical for the British accent.)</i></p> <ul style="list-style-type: none"> <li>▪ <i>Ele não pronuncia os R nos fins das palavras. (He does not pronounce the R at the end of the words.)</i></li> </ul>
<b>Verbal forms</b>	<ul style="list-style-type: none"> <li>▪ <i>A conjugação errada dos tempos verbais. (The false conjugation of the verb tempi.)</i></li> <li>▪ <i>A conjugação dos verbos. (The conjugation of the verb.)</i></li> </ul>
<b>Agreement</b>	<ul style="list-style-type: none"> <li>▪ <i>Geralmente há concordância. (Generally, there is agreement.)</i></li> <li>▪ <i>Troca a concordância de algumas palavras. (The change of the agreement of some words.)</i></li> <li>▪ <i>Mudar a concordância verbal algumas vezes.<sup>159</sup> (Change the verbal agreement sometimes.)</i></li> </ul>
<b>Further features<sup>160</sup></b>	<ul style="list-style-type: none"> <li>▪ <i>A confusão com pronomes. (The confusion with pronouns.)</i></li> <li>▪ <i>Emprego atípico de artigos ("poupar a vida ao ratinho" ao invés de "poupar a vida do ratinho"). (The atypical use of the articles "poupar a vida ao ratinho" instead of "poupar a vida do ratinho".)</i></li> <li>▪ <i>Emprego irregular dos adjetivos. (The irregular use of the adjectives.)</i></li> <li>▪ <i>Não constroi as frases direito. (He does not rightly put the phrases together.)</i></li> <li>▪ <i>Acordar a ele. Não é como falamos no Brasil. (Acordar a ele. We do not speak like that in Brazil.)</i></li> </ul>

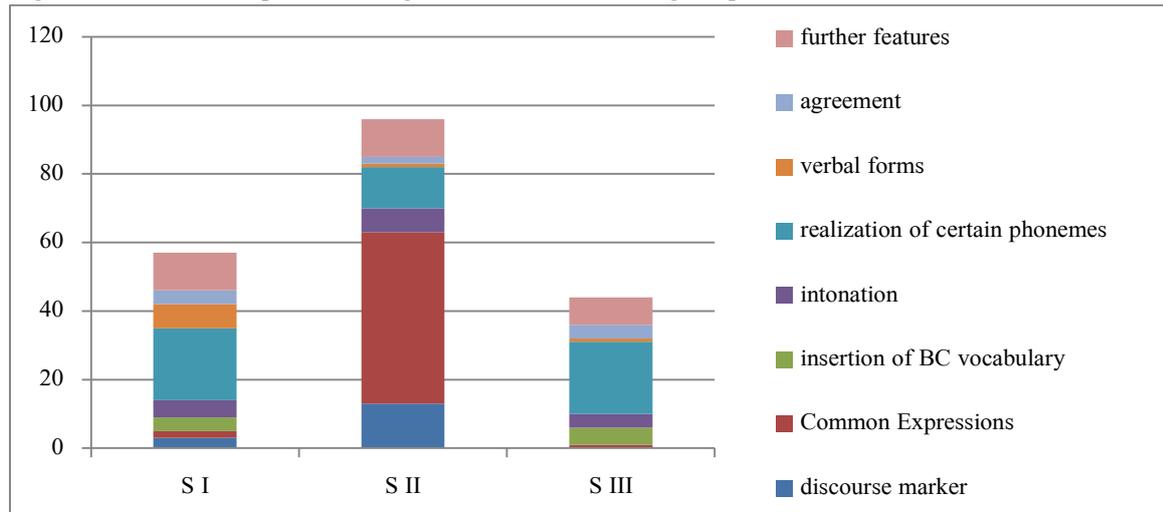
From this table one can deduce that the various linguistic features that were analyzed in the qualitative part are mentioned by the informants as well. They concern mostly Common Expressions, but also phonological and morphosyntactic features (cf. figure 7.2); for S II, comments about Common Expressions outnumber the other features by far,<sup>161</sup> (S I: 57, S II: 96, S III: 44). The reason for these differences regarding the numbers of comments is supposedly constituted by the fact that *Paulistanos* know Portuguese varieties and the second order social indexicality of certain linguistic features. For Portuguese spoken by foreigners, though, neither the stereotype nor certain linguistic features seem to be common knowledge of the *Paulistanos*. Thus, it is more difficult for the informants to name more concrete features than *sotaque* or *extranho* for the speakers they identified as foreign

<sup>159</sup> Unfortunately, it does not become clear to which nominal or verbal syntagm these comments refer, because they do not seem to describe a relation between various items of the syntagm as I would expect by the word *concordância*, but a change as the verbs *mudar* and *trocar* indicate. As agreement cannot simply change, I suspect that the use of exactly this word might be due to the fact that *concordância* is a feature that is often discussed in Brazilian schools. Thus, people use it to refer to various morphosyntactic particularities they simply have no name for. Since the survey format does not allow further inquiries with the informant, this question must remain unsolved.

<sup>160</sup> *Further features* unites various aspects regarding morphosyntax like pronouns or prepositions (in point two falsely identified as articles), or vocabulary that is only mentioned once.

<sup>161</sup>

Figure 7.2: Number of particular linguistic markers according to speaker.



The differences between the features of the speakers concern not only their quantity, but also their quality, particularly with regard to S II on the one hand, and S I and II on the other hand. While the features mentioned for S II concern mostly the employment of discourse markers and Common Expressions as well as phonology (intonation and certain phonemes). Interestingly, the use of *mano*, a Common Expression associated with the speech of young men in the periphery of São Paulo (cf. Ferreira Coelho 2006), is mentioned various times as reason for the choice of the *Periferia* or *Zona Leste* as place of origin and residence of S II, although he does not employ this feature at all. This finding does confirm the hypothesis that the perception of the hearers can differ from reality. Once a social stereotype is associated with a bundle of certain features, people deduce to hear the feature they attribute to the variety even if the speaker does not use it.

For S I and S III there are various comments about morphosyntactic features like agreement and verbal forms, as well as about the insertion of Castilian vocabulary. Those are rarely found for S II. Additionally, the other two speakers attracted most comments about the realization of certain phonemes. This type of features is mentioned 21 times each for S I and S III, but only twelve times for S II. A closer look at S I and S III's analyses reveals that they particularly concern phonemes described as examples of L1-transfer / congruent lexicalization, e.g. the nasal diphthong *-ão* or the voiceless realization of voiced sibilants. For S II, however, features like palatization or the realization of final vibrants as [Ø] are mentioned.

This constitutes a first hint that these features are most salient in the perception of the informants and hence the reason for the linguistic evaluation of the speakers' variety, particularly in connection with their social evaluation. This evaluation is clearly split along S II. While the analyses of S I and S

III's linguistic features mostly coincide quantitatively and qualitatively, S II displays another profile. This finding has consequences for the classification of the speech.

### 7.1.2 Linguistic classification of the speech

---

As was mentioned before, I asked for the classification of the examples by the direct question for the description of the language and indirectly by the question to choose the speaker's place of origin. The first question as well as the item asking for linguistic particularities resulted not only in the labeling of the language, but also in many comments about the style of the speakers. The limits between these concepts are fluent, of course; hence I only consider explicit language labels for the first question. Any further classification concerning the individual style or group references will be analyzed in the next step.

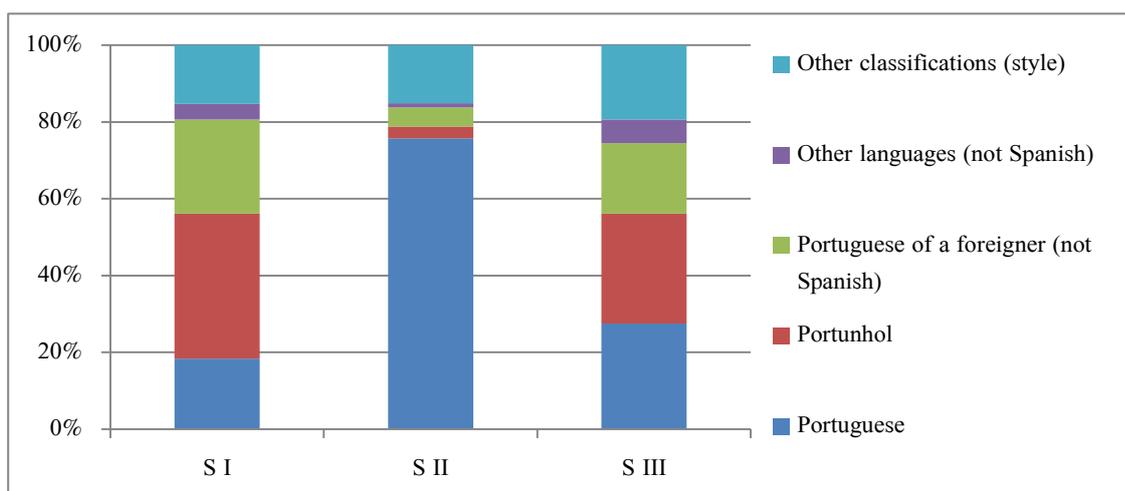
#### 7.1.2.1 Labeling the language

---

As this question was based on Carvalho's test of dialectal perception (2008: 164), I expected similar results for my examples, at least for the responses that explicitly named a language and did not refer to the style of the speakers. According to the author (l.c.), it will give answers "to the age-old question of where one language stops and another starts." The responses to this question, however, showed that almost all of the participants called the language they were hearing *Portuguese* despite the accent. Thus, my results differ greatly from Carvalho's (2008), because none of the informants labeled the language they were hearing as *Espanhol*, not even the first example that represents the group of least Portuguese proficiency and that had frequently received this label in the piloting group. This result might indicate that either the Portuguese of Carvalho's speaker with least language proficiency had a lower proficiency level than S I (her transcription does not suggest that, but as she offers no audio-material, I have no means to confirm it) or that the *Paulistanos* accept a variety as Portuguese more easily than the *Cariocas* of her study.

Instead, there are similar numbers for S I and S III, contrasting with S II like seen before:

Figure 7.3: Classification of the language ( $N_1 = 98$ ,  $N_2 = 99$ ,  $N_3 = 98$ )



Leaving aside the stylistic classifications that concern 15-20% of the responses for each speaker, the mode of the responses concerning S I and S III is *Portunhol*<sup>162</sup>, but Portuguese<sup>163</sup> in the case of S II uniting 75% of the answers. Only 9.2% labeled his speech as influenced by foreign languages. Regarding the other two speakers, there is the reverse tendency: While only 18.4% (S I) and 27.6% (S III), respectively, label their speech as Portuguese without foreign influence, more than 50% classify it as a foreign language or as Portuguese spoken by a foreigner<sup>164</sup>, among others *Portunhol*. Here, the difference between the language proficiency, that was a factor for the selection of the informants, is reflected by the fact that for S III 48%, but for S I only 33% do not mention the foreign influence. The label *Portunhol*, indicating the Spanish influence, was most attributed to S I (37.8%), clearly outnumbering the other labels in this case. With regard to S III, only 28.6% of the answers associated his speech with Spanish, but almost as many (24.5%) with other languages. Assuming that the speech examples are representative of the speech of Bolivian immigrants, this means, of course, that

<sup>162</sup> Including synonyms like

- c) português com sotaque espanhol.
- d) Parece um estrangeiro (cuja língua materna é espanhol) falando português.
- e) português misturado com espanhol.

<sup>163</sup> Including synonyms like

- f) português brasileiro
- g) português de São Paulo
- h) brasileiro colloquial

<sup>164</sup> Including synonyms like

- i) Uma mistura de francês e italiano.
- j) português com sotaque estrangeiro.
- k) português falado por um alemão.

despite the presence of many Bolivians and other Spanish-speaking people in São Paulo, their way of speaking is not recognized by the *Paulistanos*. While Carvalho (2008: 166) observed a significant relationship between the social factor “knowing other languages” and the responses to the similar question, none of the social factors that were tested (age, sex, profession, foreign language knowledge, place of residency), not even “living in quarters with strong Bolivian presence” brought any significant result. Thus, it seems that the tendencies having become apparent in the data – S II is perceived as the one with most language proficiency, while S I and S III are perceived as foreigners speaking Portuguese – are valid for all *Paulistanos*.

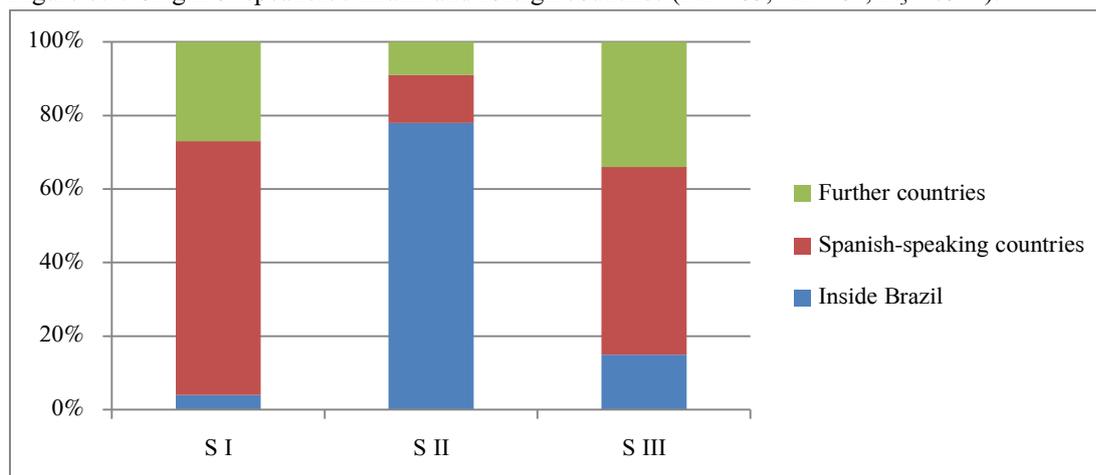
#### 7.1.2.2 The speakers' place of origin

---

Two questions were included that both aimed at a socio-geographic evaluation of the speakers, the first with regard to their place of origin, the second with regard to their place of residency. Despite of the apparent similarities between the two questions, *De onde vem o falante?* (*Where does the speaker come from?*) and *Onde mora o falante?* (*Where does the speaker live?*), they aim at different information. The first one offers another possibility to classify the speakers according to their foreign origin, the second one regards rather the social evaluation. This difference is reflected by the given scales the participants could choose from: The first only contains four items for the question (*São Paulo (capital) (City of São Paulo)*, *Interior de São Paulo (estado) (the interior of the state of São Paulo)*, *Fora de São Paulo (estado) (outside the state of São Paulo)*, *Fora do Brasil (outside Brazil)*), the second question contains a more differentiated range, particularly in terms of the region inside São Paulo state.

Thus, one would expect that the results concerning the language label are confirmed by the choice of the place of origin, i.e. that informants who chose *Portunhol* as label also chose a Spanish-speaking country as place of origin. This connection is confirmed until a certain point, as figure 7.4 demonstrates:

Figure 7.4: Origin of speakers / Brazil and foreign countries (N1 = 83, N2 = 84, N3 = 79<sup>165</sup>).



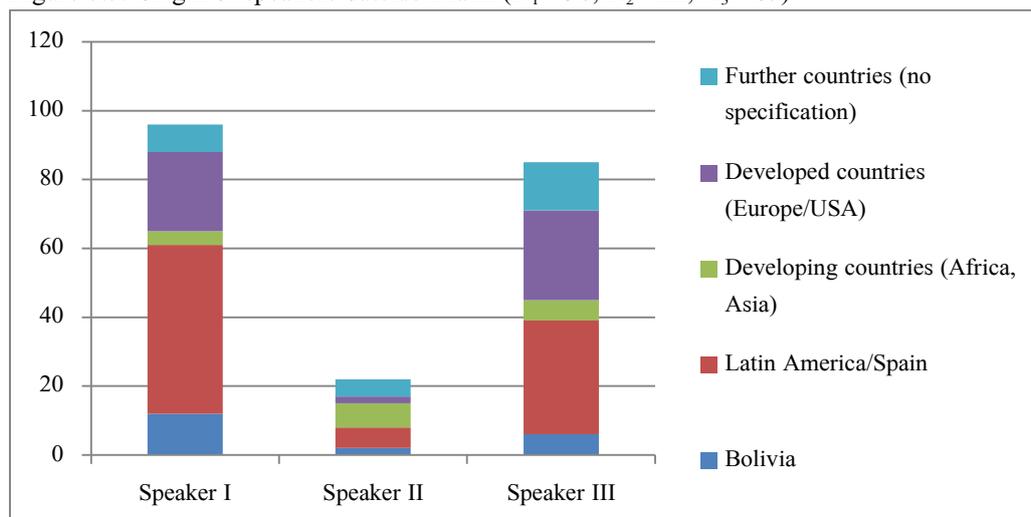
Comparing figure 7.3 and figure 7.4, the distribution for S II is almost the same, i.e. as many informants who have described his speech as *Português* and its varieties also think that he comes from Brazil. Interestingly, the numbers differ for S I and S III. Obviously, there are many more people who associate a Spanish-speaking country with the speakers than those who recognized the speech as influenced by Spanish. This lack of consistency might be due to the stylistic characterizations that can be attributed to any place of origin, or the fact that the Portuguese proficiency of the speakers is estimated high enough to be labeled as Portuguese but not as native Brazilian. Anyway, a 3x4 significance test<sup>166</sup> revealed that there is a significant relationship between the associated place of origin and the language labels of S I,  $\chi^2(6, 83) = 27.8, P = .001$ , and of S II,  $\chi^2(6, 84) = 73.0, P = .000$ .<sup>1</sup> The same procedure conducted for S III revealed no significant relationship between the place of origin and language labels attributed to him ( $\chi^2(6, 79) = 4.09, P = .684$ ). This result could be due to the fact that only fifteen informants named Brazil as his place of origin although 27 had labeled his speech as Portuguese, i.e. although people called his speech *Portuguese* they perceived him as foreigner. This result is confirmed by the many references to his *sotaque* in regard of the question about single features. Additionally, positive residuals of the *Spanish-speaking country\*Portuguese* relation (1.6) and negative residuals of the *Further countries\*Portuguese of a foreigner* relation (- 2.5) object the correlation.

Having a closer look at the places of origin outside Brazil (cf. figure 7.5), S I and S III reveal almost no differences, except for the fact that more people associated S I with Spanish-speaking countries (61, Bolivia + Latin America/Spain).

<sup>165</sup> For this analysis, stylistic characterizations were left out (=MISSING).

<sup>166</sup> Because of the size of the sample the contingency table contains expected counts < 5. Thus, I cannot conduct a chi-square-test with asymptotic p, but have to count the exact value (Janssen and Laatz 2007: 797 pp.).

Figure 7.5: Origin of speakers outside Brazil ( $N_1 = 96$ ,  $N_2 = 22$ ,  $N_3 = 85$ )



Interestingly, only 20 comments specified that the informants might come from Bolivia. This means that although the Bolivians constitute the biggest group of Latin American immigrants in São Paulo, they are not perceived as being different. In one of the open-response questions that will not be considered for this analysis, I asked the informants about their opinion on the recent immigration to São Paulo. Only one in hundred explicitly mentioned the Bolivian immigration, although many spoke about the “imigrantes latinos trabalhando na costura”. This might be another indicator for the fact that Brazilians simply do not distinguish the people coming from different Latin American countries. The narratives of Bolivians being discriminated as Columbian traffickers confirm this thesis.

For S II, in contrast to the other speakers, clearly more people associate him with developing countries<sup>167</sup> than with developed countries<sup>168</sup>. This also hints that the social evaluation of S II differs from the others.

---

<sup>167</sup> Including synonyms like

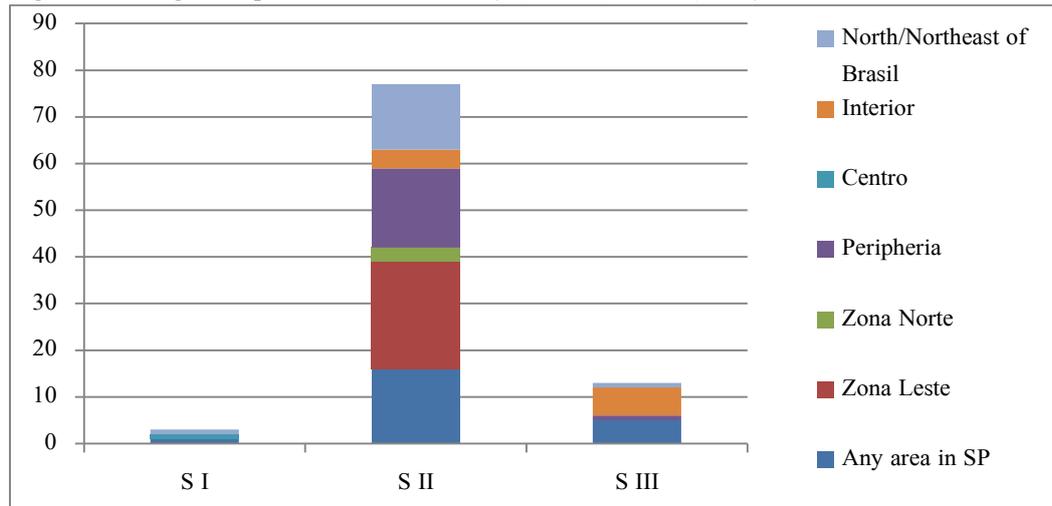
- l) países pobres
- m) da África
- n) Chinês.

Interestingly, nine informants said that S I was German (S II: 1, S III: 3). The fact alone that people explicitly mentioned German might be due to the link of the questionnaire leading to a German university. Why this led to such an unequal distribution among the speakers, however, cannot be explained.

<sup>168</sup> Including synonyms like

- o) países ricos
- p) países frios.

Figure 7.6: Origin of speakers inside Brazil ( $N_1 = 4$ ,  $N_2 = 78$ ,  $N_3 = 15$ )



Considering the distribution of the places of origin inside Brazil (cf. Figure 7.6), this estimation is confirmed by the association of S II with the North/Northeast of Brazil. The latter is the poorest area of Brazil, where many of São Paulo's *favelados* come from, who mostly live in the poorer areas of the city like *ZL* or the *Periphéria*. Interestingly, it is exactly these areas that are attributed to S II by 23 and 17 participants, respectively, when the informants specify their evaluation. *ZN* gets three votes, but none of the participants selects *ZO* or *ZS*, the socioeconomically stronger areas.

Apart from these data that already indicate that the style of S II is evaluated differently, the stylistic characterization of the speakers also underscores the discrepancy between S I and S III.

### 7.1.2.3 The speakers' style

As was argued before, group and individual style must be distinguished for any speaker, i.e. the (linguistic) style that defines the speaker's affiliation with a certain social group on the one hand and the individual style that refers more to the style of the narration without directly indicating a social group on the other hand. Naturally, this distinction is not completely unambiguous because comments like *ele fala mais formal* (*he speaks more formally*) do certainly refer to the individual style of the speaker but could also constitute a reference to PC and thus to the social group of well educated, well informed persons.

In the responses, there are many comments about the style of the speakers'. With regard to S I and S III, they are either about their individual style or – in case of comments on their group style – they refer to the level of education. The comments about S II are often explicitly connected with a social group as table 7.2 shows:

Table 7.2: Excerpt from the responses: Stylistic characterization of the three speakers<sup>169</sup> (own translation, sn)

	S I	S II	S III
Individual style	<ul style="list-style-type: none"> <li>▪ Confusa, meio insegura, raciocínio confuso. (<i>confuse, rather insecure, confuse thinking</i>)</li> <li>▪ A fala passa muita insegurança (<i>the speech communicates a lot of insecurity</i>).</li> <li>▪ Ele parece mais sério. (<i>He seems to be more serious.</i>)</li> <li>▪ Possivelmente esta pessoa está sob efeito de álcool (bebada). (<i>Possibly, the person is under the effect of alcohol (drunk).</i>)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Um brasileiro que fala um português mal falado (<i>A Brazilian who speaks a badly spoken Portuguese.</i>)</li> <li>▪ Pela fluência na língua. (<i>Through the fluency in the language.</i>)</li> <li>▪ Português extremamente colloquial, informal. (<i>Extremely colloquial, informal Portuguese.</i>)</li> <li>▪ o estilo mais jovial da linguagem (<i>The language style is more jovial.</i>)</li> <li>▪ Norma não padrão do português brasileiro. (<i>No standard norm of Brazilian Portuguese.</i>)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Boa conversa. (<i>Good conversation</i>).</li> <li>▪ Tem humor e descreve bem a situação. A lentidão de falar as palavras. (<i>He has humor and describes the situation well. The slowness of speaking the words.</i>)</li> <li>▪ Parece transformar a história em anedota. (He seems to transform the history in an anecdote.)</li> <li>▪ Parece que a pessoa está lendo. (It seems as though the person is reading.)</li> <li>▪ O falante parece ter cuidado ao escolher as palavras. (The speaker seems to attentively choose his words.)</li> <li>▪ Ele consegue se comunicar bem em português, a despeito do sotaque, e tenta fazê-lo da maneira mais correta possível. (<i>He achieves to communicate well in Portuguese, despite his accent, and tries to do it as correctly as possible.</i>)</li> <li>▪ português bem articulado (<i>Well-articulated Portuguese.</i>)</li> <li>▪ uma linguagem mais culta (<i>A more cultivated language style.</i>)</li> <li>▪ português correto, sem muitas gírias (<i>Correct Portuguese, without many swear words.</i>)</li> <li>▪ A fala é clara e pausada, além de apresentar-se bem</li> </ul>

<sup>169</sup> In order to give the most exact impression of the answers, the original quotes should be maintained as they are. I therefore do not correct them orthographically.

	S I	S II	S III
			próxima ao português culto. <i>(The speech is clear and with pauses; it also presents itself rather close to the Português Culto.)</i>
Group style: level of education	<ul style="list-style-type: none"> <li>▪ Parece um professor ou um aluno que dá ao texto bastante atenção. (He seems to be a professor or student who gives the text a lot of attention.)</li> </ul>	<ul style="list-style-type: none"> <li>▪ português falado por uma pessoa de pouca escolaridade <i>(Portuguese spoken by a person of little education at school.)</i></li> </ul>	<ul style="list-style-type: none"> <li>▪ uma pessoa estudada <i>(a person who has studied)</i>.</li> <li>▪ fala como professor <i>(he speaks like a professor)</i>.</li> </ul>
Group style: place	<ul style="list-style-type: none"> <li>▪ <i>Essa pessoa aprendeu a falar português escutando aos paulistas. (This person learned to speak Portuguese by listening to the inhabitants of the state of São Paulo.)</i></li> </ul>	<ul style="list-style-type: none"> <li>▪ Português com sotaque do Nordeste, por isso fala tão rápido e torna a compreensão bem difícil. <i>(Portuguese with accent of the Northeast, that is why he speaks so quickly and makes the understanding rather difficult.)</i></li> <li>▪ Tem um falar mais próximo do falar paulistano. <i>(He has a language style rather close to the style Paulistano.)</i></li> <li>▪ Português com influência regional paulista. <i>(Portuguese with the regional influence of São Paulo)</i></li> <li>▪ praticamente um dialeto variado do português. <i>(Actually a dialect deriving from Portuguese.)</i></li> </ul>	<ul style="list-style-type: none"> <li>▪ Dá pra perceber que ele é de São Paulo porque ele fala o “d” e o “t” do mesmo jeito que o paulista. <i>(One can perceive that he is from São Paulo, because he speaks the “d” and “t” the same way a Paulista does.)</i></li> </ul>

	S I	S II	S III
Group style: age and place		<ul style="list-style-type: none"> <li>▪ Porque na capital de São Paulo há algumas variantes e esta parece com a variante comum entre jovens na zona leste. (<i>Because in the city of São Paulo there are some varieties, and this seems like the variety common for the youngsters in the Zona Leste.</i>)</li> <li>▪ O popular “Manes” (fala como “Mano”), tradicional da Zona Leste entre os jovens. (<i>The popular “Manes (speaks like “mano”), traditional for the Zona Leste among the youngsters.</i>)</li> <li>▪ Acho que o sotaque informal me parece um jovem se identificando com o estilo periferia, mas não da periferia em si. (<i>I think that the informal accent seems to be a youngster who identifies with the style of the Periphéria, but is not from the Periphéria as such.</i>)</li> </ul>	
Group style: place and explicit social class		<ul style="list-style-type: none"> <li>▪ Não segue a norma culta, é uma pessoa da região mais pobre. (<i>He does not follow the norm, he is a person from a rather poor region.</i>)</li> <li>▪ Fala típica de favela. (<i>Typical speech of the favela.</i>)</li> </ul>	

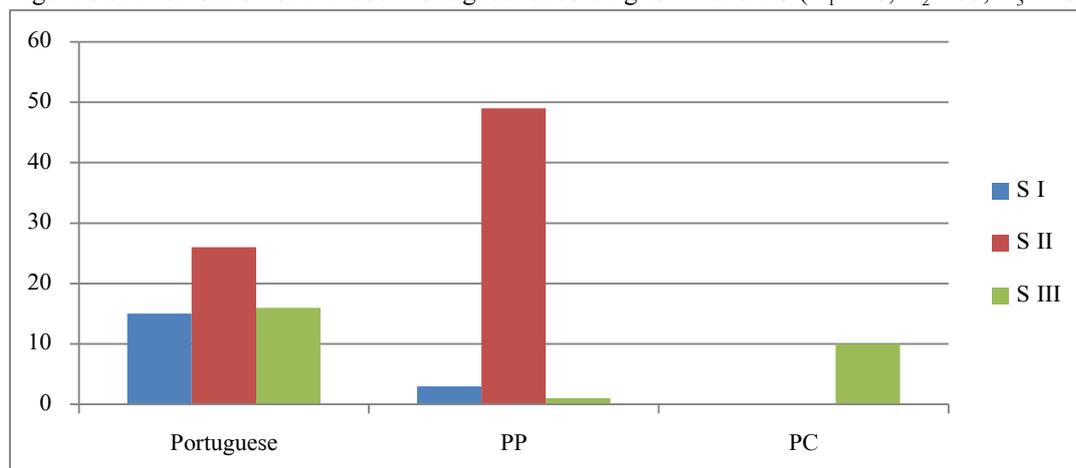
This excerpt shows that the stylistic evaluation of S II differs from the others, as it concerns not only the individual style or references to the educational level of the speakers and the places of residence or origin, but also age groups and even social class. Of course, this last dimension is also implicit in the geographic references, because the quarters mentioned by the informants, *ZL* and the *Periphéria*, are associated with poorer classes. According to the data, S II is seen as a typical young man who

belongs to a low social class and lives in the areas mentioned before, because he speaks PP. In contrast to his case, the style of S III is associated with social groups defined by their high educational level like professors or other academics, representing speakers of PC. This finding coincides with the comments about his individual style in which informants claim that he was reading.

S I, however, despite his “confuse” narration, is not associated negatively with a certain social group, but people only judge him on an individual level, thinking that he is drunk or insecure. This fact indicates once more that the *Paulistanos* do not connect a social stereotype with the speech they classify as foreign.

While the excerpt above stems from responses to the first question, almost the same data is revealed by a closer look at the responses to the language labeling question. Again, the speech of S II is mostly considered as Portuguese, but in addition there are many further differentiated comments, i.e. they include a group stylistic evaluation of the speech example that once more emphasizes the distance between the speakers, in particular between S II and S III (cf. Figure 7.7):

Figure 7.7: Distribution of the label Portuguese according to PP and PC ( $N_1 = 18$ ,  $N_2 = 75$ ,  $N_3 = 27$ ).



Apart from the cases where the speech is only labeled *português* or *brasileiro*, there are 49 comments that explicitly call the speech of S II *português colloquial* or *português informal*, subsumable under the label PP, and ten comments referring to S III as *português culto* or *português formal*, subsumable under the label PC.

In summary, mostly because of the discourse markers and Common Expressions, the vast majority of the participants think that S II’s speech is typical for PP spoken by young men coming from the poorer areas of São Paulo. The other two speakers are mostly recognized as foreigners because of the accent, but their speech is evaluated as rather positive. S III, in particular, is associated with PC.

## 7.2 Social evaluation

---

The social evaluation of the speakers was investigated by indirect and direct questions. The latter were constituted by three semantic differentials, asking how the informants perceived the speakers with regard to their prosperity, effort, and knowledge. Additionally, to meet possible reluctance of the participants to explicitly reveal (negative) social stereotypes, I inquired which place of residence the informant associated with the speakers. Because of the strong socio-geographic segregation of São Paulo, it is possible to deduce the social reasons underlying the selection without having to ask directly.

### 7.2.1 The speakers' place of residence

---

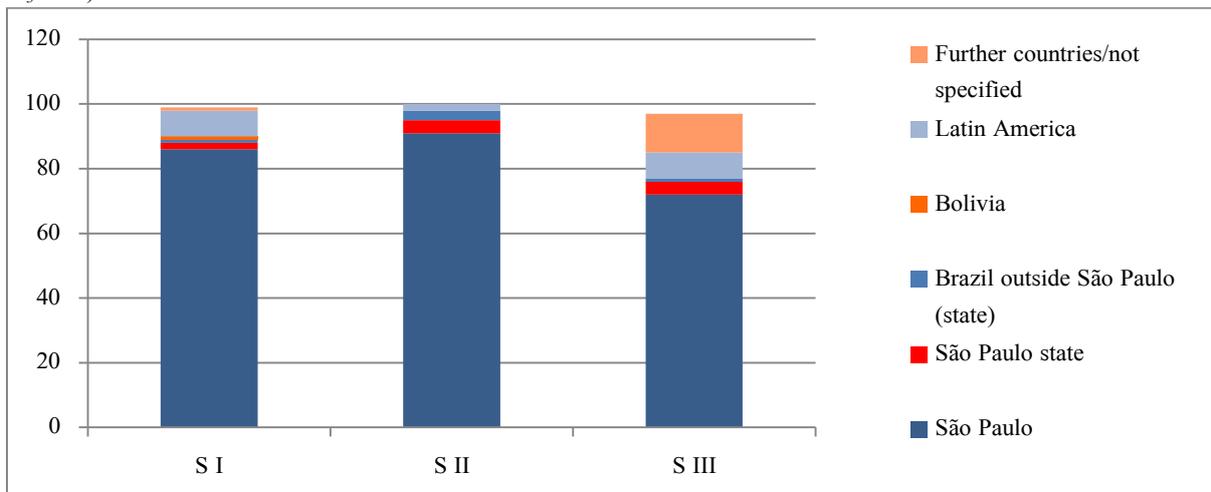
In the question about the probable place of the speakers' residence, the informants could choose among various quarters that each stands for one area of São Paulo, *Centro*, *ZL*, *ZO*, *ZS* and *ZN*, as well as the options *outside São Paulo* and *outside Brazil*. Additionally, the informants had the opportunity to add further quarters or specify their choice if they voted for the last two options. Looking at the distribution of the answers when uniting all quarters of São Paulo, the following pattern is revealed (cf. Figure 7.8): The majority of the informants think that all three speakers live in São Paulo.<sup>170</sup> For S I (10) and S III (20), some chose Latin American countries – only one mentioned Bolivia – but also European countries and the US for S III. The informants obviously thought that S III learned his Portuguese abroad and not in the natural environment of São Paulo. Compared to the place of origin there is a huge difference; many people think that S I lives in São Paulo but comes from another country.<sup>171</sup>

---

<sup>170</sup> This might also be influenced by the title of the questionnaire, *Como fala São Paulo?*; in the opinion, the social evaluation of the speakers would not be influenced by the information that the recordings were made in São Paulo, I chose this title to get attention in the social network platforms.

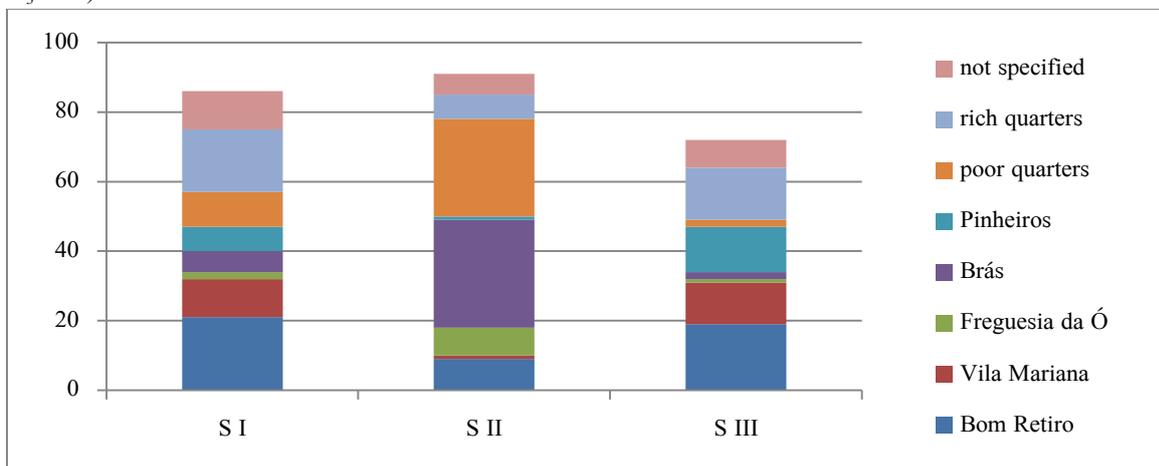
<sup>171</sup> A 3x3chi-square-test (place of origin\*place of residence: Inside Brazil, Latin America, Further countries), revealed no significant correlation.

Figure 7.8: Place of residence according to speakers, differences inside São Paulo excluded ( $N_1 = 99$ ,  $N_2 = 100$ ,  $N_3 = 97$ ).



These cases, however, allow little interpretation with regard to social factors. Considering the distribution inside São Paulo, though, I get a more detailed picture about the social evaluation of the speakers. Figure 7.9 shows that 80 responses connected their choice explicitly with the social structure of the area, by comments like *lugares de maior poder aquisitivo (places of a higher purchasing power)* or *uma região mais pobre (a poorer region)*, attributing mostly S II to poor areas (28), and S I (18) and S III (15) to rich areas.

Figure 7.9: Place of residence according to speakers, differences outside SP excluded ( $N_1 = 99$ ,  $N_2 = 100$ ,  $N_3 = 97$ ).



Apart from *Bom Retiro*, which is considered the typical immigrants' quarter in the center of São Paulo<sup>172</sup>, *Pinheiros* and *Vila Mariana* are the areas that are classified as financially most privileged<sup>173</sup>,

<sup>172</sup>This estimation is also confirmed by comments on Bom Retiro like

- q) Por ser um baixxo que reúne muitos migrantes é onde se concentram os latinos.
- r) Em razão de ser um bairro central com grande concentração de imigrantes provenientes de países da América Latina.

while *Brás* and *Freguesia da Ó* are seen as rather deprived areas.<sup>174</sup> The dispersion of the speech examples is unambiguous: The mode of S I and S III is *Bom Retiro*, the mode of S II is *Brás*. Including *Freguesia da Ó* and “poor quarters”, 67% of the informants associate S II with one of the socially unprivileged areas, only 9% with more prosperous quarters and also only 9% with the immigrants’ quarter (and actual home of the speaker), *Bom Retiro*. In contrast to S II, 30% think that S III lives in one of the wealthier areas. Only 5% associate him with poorer areas and 19% with his actual home, *Bom Retiro*. For S I, the distribution is similar, but not as unambiguous as for the other two, because 21% think he comes from *Bom Retiro*, 36% consider him an inhabitant of the richer areas, but 18% think that he comes from the poorer areas. Chi-square-tests of this evaluation and the social factors of the informants included in the questionnaire (sex, age group, profession, knowledge of foreign languages) revealed no significant relation, but a correlation analysis disclosed a negative correlation between the evaluation of S II’s residence according to social status and the informants’ place of residence according to Bolivian presence ( $r = -.25, p < 0.2$ ). This means that none of the informants living in areas with strong Bolivian presence associated S II with rather privileged areas, but nine informants living in areas with little Bolivian presence did, as table 7.3 illustrates:

Table 7.3: S II: Cross table Residence of informants according to Bolivian presence \* residence in SP (social status)

			S II: residence in SP acc. to social status			Total
			Bom Retiro	socially rather privileged areas	socially rather unprivileged areas	
Residence of informants acc. to Bolivian presence	areas with strong Bolivian presence	Count	4	0	38	42
		Expected count	4.4	4.4	33.1	42.0
	areas with little Bolivian presence	Count	5	9	29	43
		Expected count	4.6	4.6	33.9	43.0
Total		Count	9	9	67	85
		Expected count	9.0	9.0	67.0	85.0

s) Pensei nos bolivianos que dominam a região do Bom Retiro em São Paulo!

<sup>173</sup> This estimation is confirmed by comments like

t) Poderia muito bem ser um estudante estrangeiro, na USP (Pinheiros) ou na Escola Paulista (Vila Mariana). Ele deve ter recursos financeiros.

u) O gringo com grana mora na zona centro ou perto, zona oeste melhor

<sup>174</sup> This estimation is confirmed by comments like t) when choosing *Brás* and u) and v) when choosing *Freguesia da Ó*.

v) Moraria no brás, por ser um local com mais acesso ao transporte público e aluguel mais barato.

w) Devido à fala repleta de gírias típicas de falantes de classes sociais marginalizadas.

x) Por ser o bairro mais periférico entre as opções.

This could mean that the contact with Bolivian and other Latin American immigrants sensitized the informants to the poor living conditions of most of them. Another explanation could be the correlation between immigrants' areas and socially rather unprivileged areas<sup>175</sup> that became apparent by a further analysis ( $r = -.26$ ,  $p = .014$ ); i.e. people coming from socially unprivileged areas think that S II lives where they live. This pattern will be further checked by the semantic differentials concerning the ascription of social characteristics.

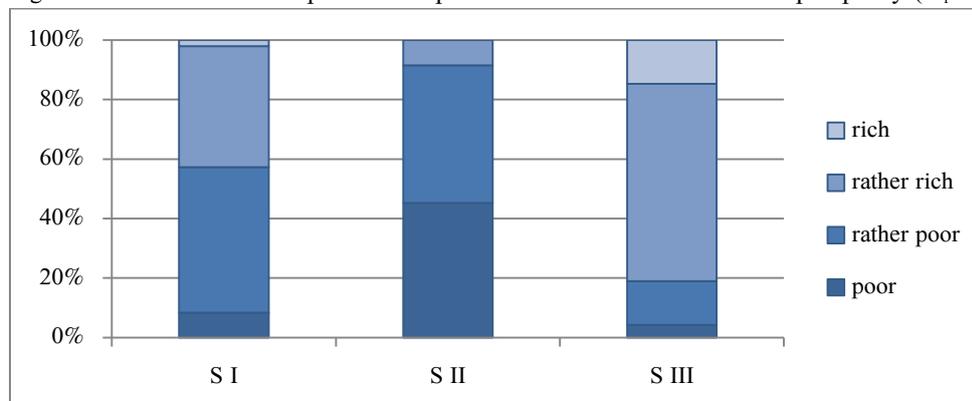
### 7.2.2 The ascription of social characteristics

According to Carvalho (2008), the social characteristics are measured by the semantic differential test to get data about the (unconscious) association of linguistic elements and non-linguistic characteristics, i.e.

o objetivo desta técnica é revelar as atitudes que os membros de uma comunidade de fala têm em relação às variedades lingüísticas, sem fazer-lhes perguntas diretas sobre o assunto (l.c.: 168).

The evaluation of the adjective pairs concerning prosperity, effort and knowledge by a scale containing four steps, disclose the following results:

Figure 7.10: Evaluation of speech examples: Semantic differential about prosperity ( $N_1 = 96$ ,  $N_2 = 95$ ,  $N_3 = 95$ ).



<sup>175</sup> The difference between the factors lies in the different classification of ZN, being a socially rather unprivileged area without Bolivian immigration.

Figure 7.11: Evaluation of speech examples: semantic differential about knowledge ( $N_1 = 95$ ,  $N_2 = 96$ ,  $N_3 = 97$ ).

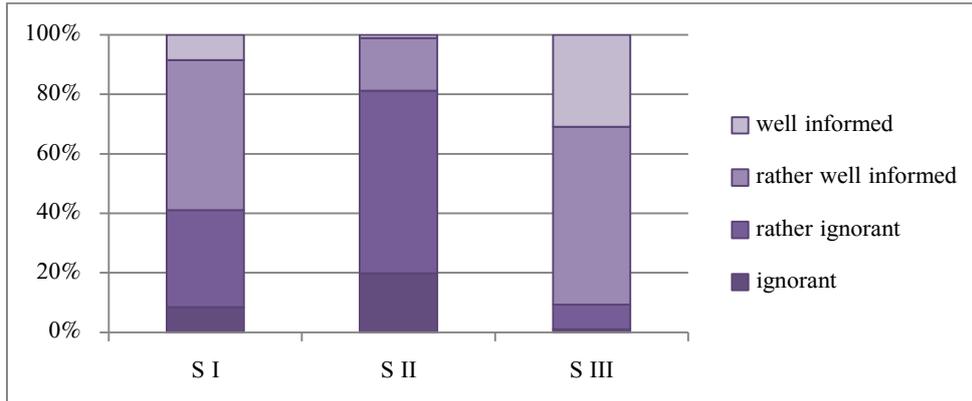
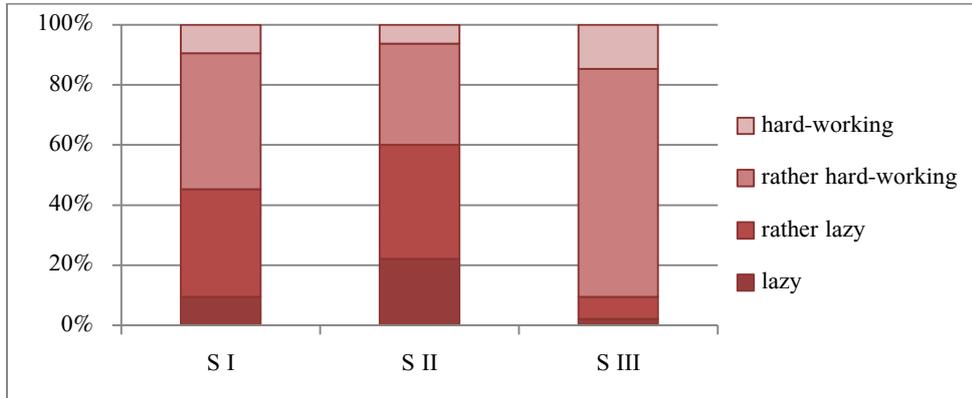


Figure 7.12: Evaluation of speech examples: semantic differential about effort ( $N_1 = 95$ ,  $N_2 = 95$ ,  $N_3 = 95$ ).



We immediately see that S II gets the least positive results of the three speakers (indicated by the lighter colors). This is confirmed by the observation of the extreme values (1) and (4): 6% think S II is hard-working, 21% think that he is lazy and merely 1% consider him as well informed while 19% consider him ignorant. None estimate that he is rich but almost 50% of all informants that he is poor. With regard to S III, I can observe almost the contrary: The vast majority assign positive values to him: Only 4% think he is poor but 14% that he is rich, 2% consider him lazy but 14% hard-working and just 1% think he is ignorant but 30% that he is well informed. The dispersion for S I is more ambiguous, displaying the same amount of both extreme values regarding effort and knowledge. Only with regard to S I's estimated prosperity, 8% evaluate him as poor and just 2% as rich. In comparison to S III, however, a clear distinction appears between the two speakers: While at least 80% of the informants associate the socially more positive scores with S III, at most 60% do so for S I.

As one would deduce from the obvious semantic similarity of the adjective pairs, a correlation analysis revealed that there is a positive correlation between the semantic differentials at the  $p < .01$  level (cf. table 7.4).

Table 7.4: Correlation between three semantic differentials according to speaker.

		S I			S II			S III						
		prosperity	effort	knowledge			prosperity	effort	knowledge			prosperity	effort	knowledge
r	prosperity	1.00	.30**	.67**	prosperity	1.00	.39**	.49**	prosperity	1.00	-.02	.44**		
N		96	95	95		95	94	94		95	94	95		
r	effort	.30**	1.00	.51**	effort	.39**	1.00	.40**	effort	-.02	1.00	.28**		
N		95	95	94		94	95	94		94	95	95		
r	knowledge	.67**	.51**	1.00	knowledge	.49**	.40**	1.00	knowledge	.44**	.28**	1.00		
N		95	94	95		94	94	96		95	95	97		

\*\**p* < .01 (2-sided).

Interestingly, I find the clearest correlation between prosperity and knowledge, while the correlation between effort and prosperity is rather weak, being not even significant in the case of S III<sup>176</sup>. This means that for the Paulistanos, prosperity and knowledge are directly connected, i.e. persons who are well informed are also rich but not necessarily hard-working, and someone who is poor is not well informed but can be hard-working anyhow. With regard to the speakers, these results clearly confirm the former findings that S III is seen as rather rich and hence a well-educated speaker of PC while S II is prototypical for speakers of little financial means and hence rather badly educated speakers of PP.

Checking possible social factors influencing the semantic differential test, I carried out various independent-samples-t-tests to compare the evaluations with regard to the informants' sex, social status of their place of residence, and their knowledge of foreign languages. For men and women there are significantly different scores only with regard to S I (prosperity and knowledge)<sup>177</sup>. A further test distinguishing between informants knowing no foreign language and the others knowing foreign languages, showed significant differences in scores also for S I but for S II<sup>178</sup>, in particular, indicating that the latter group evaluates S I's prosperity and all social characteristics of S II more positively. The differences in the means were very small though (eta squared < .01); i.e. this factor explains

<sup>176</sup> The negative value shows that there is even an inverted correlation between prosperity and effort, i.e. the richer the lazier. Because of the insignificant score, this result should not be overrated however.

<sup>177</sup> S I: poor (1) – rich (4): men (M=2.19, SD=.68) and women (M=2.53, SD=.62), *t*(94)=-2.56, *p*<0.5; ignorant (1) – well informed (4): men (M=2.40, SD=.77) and women (M=2.78, SD=.72), *t*(93)=-2.39, *p*<0.5, meaning that women perceived S I more positively than men. The differences in the means, though, were rather small (for both eta squared=0.6).

<sup>178</sup> S I: poor (1) – rich (4): further languages (M=2.48, SD=.63) and no further languages (M=2.04, SD=.68), *t*(94)=2.94, *p*<0.1.; S II: poor (1) – rich (4): further languages (M=1.76, SD=.65) and no further languages (M=1.28, SD=.46), *t*(93)=3.34, *p*<0.5; lazy (1) – hard-working (4): further languages (M=2.44, SD=.84) and no further languages (M=1.67, SD=.70), *t*(93)=4.03, *p*<0.1; ignorant (1) – well informed (4): further languages (M=2.10, SD=.66) and no further languages (M=1.72, SD=.54), *t*(94)=2.58, *p*<0.5.

at most 1% of the variance. The t-test I conducted to compare the results from people living in socially privileged areas and socially unprivileged areas also reveals that S II is judged more positively by the participants living in more privileged areas. Therefore, it seems like these informants employ more “political correctness”, while the others do not hesitate to judge someone they perceive to be from the same social background. This hypothesis is also confirmed by the result of a one-way analysis of variance indicating a significant difference between academics ( $M = 1.84$ ,  $SD = .68$ ) and students ( $M = 1.36$ ,  $SD = .49$ ),  $F(2,92) = 4.33$ ,  $p < .05$  with regard to the evaluation of S II’s prosperity. This means that academics perceive S II as more positive than students, although many informants consider his speech as typical for the youngsters living in the periphery. Because of the obvious correlation between students and age group, it comes as no surprise that significant differences are found for the different ages, too: one-way analyses of variance reveal that for the first age group (15-24) he is significantly poorer ( $M = 1.33$ ,  $SD = .48$ ) than for the other age groups ( $M_2 = 1.73$ ,  $SD = .69$ ;  $M_3 = 1.77$ ,  $SD = .62$ ),  $F(2,92) = 4.49$ ,  $p < 0.5$ . Additionally, the youngest differ from the oldest with regard to S III’s knowledge, as the 15-24 year old participants evaluate him as significantly more informed ( $M = 3.43$ ,  $SD = .50$ ) than the others ( $M = 2.94$ ,  $SD = .62$ ),  $F(2,94) = 5.31$ ,  $p < .01$ .

In summary, the social characteristics tested by the semantic differentials and the place of residence disclose a quite homogenous picture for all social groups represented by the sample. They all perceive S II as the one who is classified as socially least privileged, revealing a strong social stereotype connected with his way of talking. This confirms the results from the linguistic classification: While S II is perceived as representative of the young men living in poor areas, S III stands for the well-educated speakers of PC. As one informant puts it, his speech is “*the most correct of the three examples*”, although most of the informants still think that he is a foreigner. S I, though, is found between these two extremes, the opinions about him differ as much as the labeling of his speech as drunk on the one hand or as professor-like on the other hand. Overall, a clear social hierarchy of the speakers results from this analysis, seeing the apparent PP speaker as the one with least positive social status, the apparent Latin American as socially rather low, and the other foreigner (and PC-speaker) as the one with the highest social status.

### 7.3. Personal evaluation

---

Social psychology has demonstrated that evaluations of speakers fall into two broad categories: social status and solidarity. According to Edwards (1999: 102), the first concerns competence and intel-

ligence while the second combines integrity and attractiveness. In perception tests about linguistic varieties, the two categories often reveal an inverse relationship, i.e. although lower-class, minority, and “provincial” speech styles often have positive connotations in terms of integrity and attractiveness, their speakers are typically assessed as being less competent, less intelligent, and less ambitious than are those who enjoy some regional, social, or ethnic majority status (l.c.: 103).

Thus, one would expect that S II, whose scores with regard to the social status are rather low, gets more positive evaluation in the further questions about his attractiveness or sociability, while S III is judged rather negatively for representing high competence. As accommodation theory additionally claims that convergence is positively evaluated by the recipients, i.e.

will lead to high ratings for friendliness, attractiveness, and solidarity when recipients perceive [...] a match to their own communicational style [or] a match to a linguistic stereotype for a group in which they have membership (Gallois, Ogay and Giles 2005: 131), the described behavior is expected at least by the young men living in socially underprivileged areas, revealing significant differences between them and the academics who rather perceive S III as converging toward their speech style.

As this study is more interested in the linguistic and social aspects, the personal perception was investigated only by three adjective pairs. Akin to the semantic differentials analyzed above and again inspired by Carvalho (2008: 170), they were constituted by a scale of four proceeding grades between pouco simpático (little likeable) and simpático (likeable), pouco sociável (little sociable) and sociável (sociable) as well as between pouco atraente (little attractive) and atraente (attractive). Figures 7.13 to 7.15 illuminate the evaluation according to these characteristics.

Figure 7.13: Evaluation of speech examples according to likeability scale ( $N_1 = 96$ ,  $N_2 = 97$ ,  $N_3 = 98$ ).

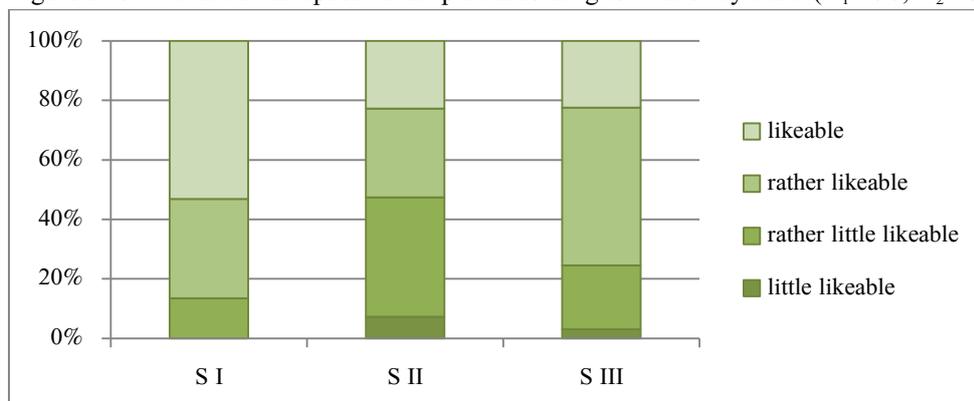


Figure 7.14: Evaluation of speech examples according to sociability scale (N<sub>1</sub> = 96, N<sub>2</sub> = 97, N<sub>3</sub> = 98).

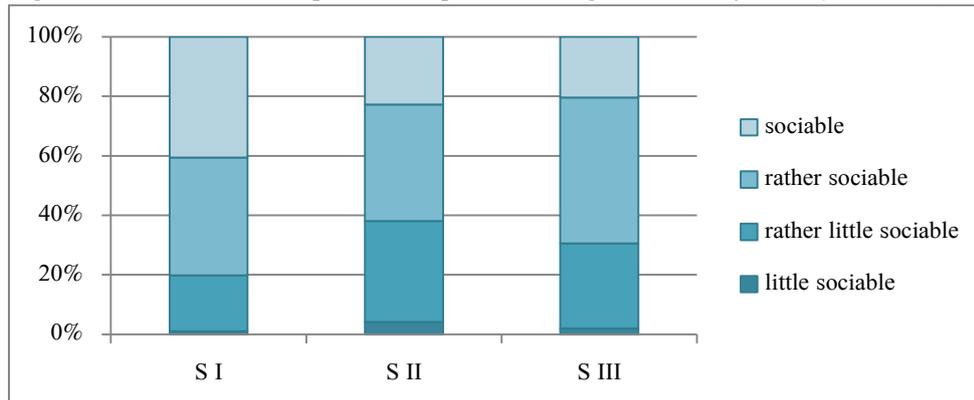
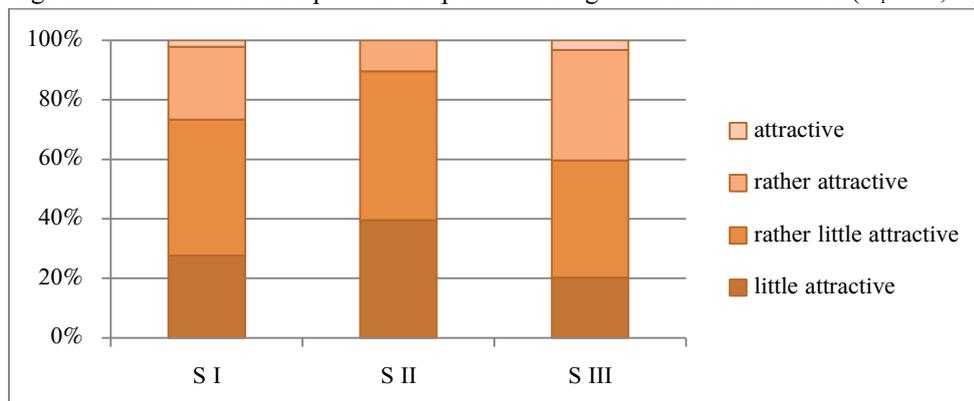


Figure 7.15: Evaluation of speech examples according to attractiveness scale (N<sub>1</sub> = 94, N<sub>2</sub> = 96, N<sub>3</sub> = 94).



Interestingly, the attractiveness scales clearly differ from the other characteristics, because all three examples are judged negatively (little attractive (3) to rather little attractive (4)) by the majority of the participants (S I: 73.4%, S II: 89.6%, S III: 59.6%), while they all get rather positive ratings in regard of sociability (rather sociable (2) and sociable (1); SI: 80.2%, S II: 61.9%, S III: 69.4%) and likeability (rather likeable (2) and likeable (1); SI: 86.4%, S II: 52.6%, S III: 75.5%). This is also confirmed by correlation analysis, revealing a positive correlation between S III's attractiveness and sociability ( $r = .42$ ,  $p < .01$ ) as well as between attractiveness and likeability ( $r = .30$ ,  $p < 0.1$ ). The correlation between sociability and likeability, though, is rather stronger; for this pair there are significant results at the  $p < .01$  level for all speakers ( $r_1 = .68$ ,  $r_2 = .55$ ,  $r_3 = .60$ ).<sup>179</sup>

In contrast to the social hierarchy established before, S III is evaluated most positively only with regard to attractiveness. Concerning sociability and likeability, S I is clearly judged more positively.

<sup>179</sup> This could be due to a slight different semantics of *atraente* having a sexual connotation that had already irritated particularly the men of the pilot group. A t-test revealed no significant differences between men and women, though.

Comparing the results with the semantic differentials of the social characteristics, S I and S II reveal a negative correlation<sup>180</sup> at the  $p < .01$  level between attractiveness and prosperity ( $r_1 = -.28$ ,  $r_2 = -.34$ ), attractiveness and effort ( $r_1 = -.37$ ,  $r_2 = -.28$ ), as well as between attractiveness and knowledge at the  $p < .01$  level for S I ( $r_1 = -.31$ ) and at the  $p < 0.5$  level for S II ( $r_2 = -.23$ ). Similar to these findings, likeability is negatively correlated with effort ( $r_1 = -.25$ ,  $r_2 = -.26$ ;  $p_{1,2} < 0.5$ ) and in the case of S I also with knowledge ( $r = -.22$ ,  $p < 0.5$ ).

These data signify that the inverse connection between social status and solidarity as found by psycholinguistic studies cannot be confirmed. On the contrary, the responses reveal that the lower the social status, the lower are also the scores for personal attractiveness, sociability and likeability, particularly in the case of S II. For S III, there are no significant data revealing a positive correlation between prosperity and likeability, and the frequencies observable in figure 7.13 rather point to a negative correlation.

In order to verify or falsify the findings of accommodation theory, that people tend to perceive those speakers as positive who speak similar to themselves, further tests were conducted to contrast the personal evaluation of the different social groups.

The independent-samples-t-tests comparing the results with regard to the informants' sex, social status of their place of residence, and their knowledge of foreign languages reveals no significant difference between men and women but indeed between people speaking foreign languages and the others, as well as between participants living in socially rather privileged areas and those living in rather unprivileged areas. Considering the knowledge of foreign languages as distinguishing factor, there are significantly lower scores for S III by the participants knowing further languages, i.e. he is perceived as significantly more likeable, more sociable and more attractive by this group.<sup>181</sup> For S I and S II, the scores regarding sociability differ significantly: The informants knowing further languages judge S I (and III) as more sociable, but they perceive S II as significantly less sociable than

---

<sup>180</sup> The poles measuring the personal evaluation are arranged in inverse order to the semantic differentials, i.e. (4) stands for *little attractive*, one of the most negative values of the personal ranking, but also for *richest*, one of the most positive values of the social ranking. Thus, the negative correlation means that low social values are interrelated with low personal values.

<sup>181</sup> S III: attractive (1) – little attractive (4): further languages ( $M = 2.59$ ,  $SD = .77$ ) vs. no further languages ( $M = 3.30$ ,  $SD = .70$ ),  $t(92) = -3.95$ ,  $p < .001$ ; sociable (1) – little sociable (4): further languages ( $M = 1.99$ ,  $SD = .74$ ) vs. no further languages ( $M = 2.50$ ,  $SD = .65$ ),  $t(96) = -3.13$ ,  $p < 0.1$ ; likeable (1) – little likeable (4): further languages ( $M = 1.88$ ,  $SD = .71$ ) vs. no further languages ( $M = 2.56$ ,  $SD = .65$ ),  $t(96) = -4.26$ ,  $p < .001$ ; the differences in the means, though, were very small (eta squared = 0.09 to 0.16), i.e. this factor explains only 9% to 16% of the variance in the evaluations.

the informants speaking no further languages.<sup>182</sup> This contrast could point into the direction described by CAT, because the *Paulistanos* who know foreign languages have usually studied more time and hence belong to a more privileged social class than the others. The magnitude of these differences, however, is very small (eta squared S I = .11, S II = .05), i.e. only 11% and 5%, respectively, of the variance in sociability can be attributed to the knowledge of further languages.

Therefore, further tests considering social factors like the place of residence and profession were conducted. An independent-sample-t-test revealed significant differences for the scores of S III's likeability and S I's sociability, showing that people living in areas with little Bolivian presence like the *ZO* or *ZS* evaluate S I and S III as significantly more sociable and more likeable, respectively, than the people living in areas of strong Bolivian presence like the *ZL* or the *Centro*<sup>183</sup>, but this factor only explains 7% (sociability S I, likeability S III) and 5% (likeability S I) of the variances. The same test distinguishing between socially privileged and unprivileged areas of São Paulo did not reveal new insights but confirmed the results of the Bolivian areas, although the difference with regard to S I's likeability was not significant in this case.<sup>184</sup> With regard to S II, there is no significant difference, although I would have expected a more positive evaluation by the informants living in the socially unprivileged areas because of his unambiguous classification as someone living in the *Periferia*.

To investigate the influence of the profession and the age of the informants, I conducted one-way analyses of variance.

Table 7.5: One-way descriptives: attractiveness of S I according to informants' profession.

	N	Mean	SD	SE	95%-Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper bound		
Academics	33	2.82	.917	.160	2.49	3.14	1	4
Students	23	3.43	.590	.123	3.18	3.69	2	4
Non-academics	38	2.87	.665	.108	2.65	3.09	2	4
Total	94	2.99	.783	.081	2.83	3.15	1	4

<sup>182</sup> S I: sociable (1) – little sociable (4): further languages (M = 1.65, SD = .74) vs. no further languages (M = 2.24, SD = .72),  $t(94) = -3.47$ ,  $p < 0.1$ ; S II: sociable (1) – little sociable (4): further languages (M = 2.31, SD = .83) vs. no further languages (M = 1.88, SD = .78),  $t(95) = 2.24$ ,  $p < .05$ .

<sup>183</sup> S I: sociable (1) – little sociable (4): few Bolivians (M = 1.60, SD = .71) vs. Bolivian neighborhood (M = 2.00, SD = .79),  $t(94) = 2.63$ ,  $p < 0.5$ ; likeable (1) – little likeable (4): few Bolivians (M = 1.45, SD = .62) vs. Bolivian neighborhood (M = 1.76, SD = .78),  $t(94) = 2.14$ ,  $p < 0.5$ ; S III: likeable (1) – little likeable (4): few Bolivians (M = 1.86, SD = .79) vs. Bolivian neighborhood (M = 2.24, SD = .68),  $t(96) = 2.63$ ,  $p < 0.5$ .

<sup>184</sup> S I: sociable (1) – little sociable (4): privileged neighborhood (M = 1.57, SD = .63) vs. Bolivian neighborhood (M = 1.98, SD = .84),  $t(94) = 2.65$ ,  $p < 0.5$ ; S III: likeable (1) – little likeable (4): privileged neighborhood (M = 1.86, SD = .68) vs. unprivileged neighborhood (M = 2.20, SD = .78),  $t(96) = 2.27$ ,  $p < .05$ .

In table 7.5 the scores of students and non-academics vary only between 2 (rather attractive) and 4 (little attractive). The academics also chose the first option (attractive), while the students' mean score is considerably higher than the others. The analysis showed that this difference is significant in the evaluation of S I's attractiveness among academics, students and non-academics,  $F(2, 94) = 5440, p < .01$ . The effect size was rather strong (eta squared = .10). Tamhane post hoc tests indicated that students judged S I as significantly less attractive than academics and non-academics ( $p < .05$ ), whereas the latter two groups did not differ from each other significantly.

Interestingly, this effect is not repeated by the analysis according to age groups. Here, only S III's scores show significant differences, as a further one-way analysis of variance revealed:

Table 7.6: One-way descriptives: S III: personal characteristics according to informants' age.

informants' age	N	Mean	SD	SE	95%-Confidence Interval for Mean		Minimum	Maximum	
					Lower Bound	Upper bound			
attractiveness	15 to 24 years	28	3.04	.793	.150	2.73	3.34	2	4
	25 to 34 years	35	2.46	.780	.132	2.19	2.73	1	4
	35 years and above	31	2.87	.763	.137	2.59	3.15	2	4
	Total	94	2.77	.809	.083	2.60	2.93	1	4
sociability	15 to 24 years	28	2.46	.793	.150	2.16	2.77	1	4
	25 to 34 years	38	1.84	.718	.116	1.61	2.08	1	3
	35 years and above	32	2.16	.628	.111	1.93	2.38	1	3
	Total	98	2.12	.750	.076	1.97	2.27	1	4
likeability	15 to 24 years	28	2.46	.838	.158	2.14	2.79	1	4
	25 to 34 years	37	1.86	.713	.117	1.63	2.10	1	4
	35 years and above	33	1.91	.579	.101	1.70	2.11	1	3
	Total	98	2.05	.751	.076	1.90	2.20	1	4

The statistics show that there are huge differences between S III's personal evaluation of the different age groups; the youngest informants judge S III most negatively, the scores of the older participants are most positive.

The one-way analysis discloses significant differences with regard to attractiveness,  $F(2,91) = 4.718, p < .05$ , sociability,  $F(2,95) = 6.199, p < .01$ , and likeability,  $F(2,95) = 6.657, p < .01$  with considerable effect sizes (eta squared<sub>attractiveness</sub> = .09, eta squared<sub>sociability</sub> = .12, eta squared<sub>likeability</sub> = .12). Tamhane post-hoc tests revealed significant differences between 15 to 24 year old and 25 to 34 year old concerning S III's attractiveness ( $p < .05$ ) and sociability ( $p < .01$ ). With regard to his likeability, the youngest differ significantly from the two older age groups,  $p < .05$ . As S III got very positive scores evaluating his social status, this could be a case of the inverse evaluation observed by psycholin-

guists: Speakers who get high scores for their social status are judged less positively with regard to solidarity, at least by those recipients by whom knowledge, effort and prosperity is not (yet) valued as much as by the older ones. It is surprising, though, that this effect is not visible for S II. Since the majority perceived him to speak like young men in the *Periferia*, one would expect that young participants identified significantly more with him than the other groups, particularly those living in socially unprivileged areas.

Thus, I conducted a two-way-analysis to observe if at least the evaluation of young people living in socially unprivileged areas differed from the others. While sociability and attractiveness revealed no significant difference, likeability did ( $p < .05$ ).

Table 7.7: Two-way analysis: Dependent Variable: S II: likeability (likeable (1) to little likeable (4))

Age	Place of residence	Mean	SD	N
15 to 24 years	SUA	1.82	1.015	17
	SPA	2.27	.786	11
	Total	2.00	.943	28
25 to 34 years	SUA	2.30	.979	20
	SPA	2.71	.772	17
	Total	2.49	.901	37
35 years and above	SUA	2.71	.772	17
	SPA	2.07	.799	15
	Total	2.41	.837	32
Total	SUA	2.28	.979	54
	SPA	2.37	.817	43
	Total	2.32	.908	97

Note: SUP: Socially unprivileged area, SPA: Socially privileged area.

The descriptive statistics show that, indeed, as the only group with a mean  $< 2.00$ , the first age group living in socially unprivileged areas like the *ZL* and the *Periferia* evaluate S II most positively.

Table 7.8: Tests of Between-Subjects Effects: Dependent Variable: S II: likeability (likeable (1) to little likeable (4))

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta-Squared
Corrected Model	10.248 <sup>a</sup>	5	2.050	2.709	.025	.130
Intercept	501.260	1	501.260	662.574	.000	.879
Age group	3.319	2	1.659	2.193	.117	.046
Neighborhood	.121	1	.121	.160	.690	.002
Age group * Neighborhood	6.005	2	3.003	3.969	.022	.080
Error	68.845	91	.757			
Total	601.000	97				
Corrected Total	79.093	96				

a. R-Squared = .130 (Adjusted R Squared = .082)

Table 7.8 shows that the differences according to age or neighborhood alone are not significant, but the combination of the two variables reveals a significant difference between the youngest living in the areas that were also associated with S II, and the other informants. The fact that people of the third age group living there, too, perceive him as rather little likeable, indicates that the older people do not approve of his style, while informants of the same age group living in more privileged areas are quite positive in their judgment.

Altogether, the personal evaluation of the speakers is less homogenous than the linguistic and social characterization. Various social factors influence the evaluation of the three speakers. However, the hypothesis that the listeners perceive those as positive who speak akin to their own style, can only partly be confirmed. The evaluation of S II by youngsters living in poorer areas pointed in this direction, but in summary, these are only tendencies, because the differences rarely reach significance. This result might be due to the fact that particularly S I and S III are perceived as foreigners and hence no one evaluates them as “someone who speaks like me”. Still, apart from the attractiveness, the speakers’ scores were rather positive.

Similar to the social evaluation before, S II is again evaluated most negatively. This could indicate that poverty and low social status are also connected with negative personal evaluation, exposing discrimination inside São Paulo’s society not because of ethnics but because of social class as was also commented by one of the informants: “Em São Paulo é bem comum o preconceito social e não necessariamente racial ou cultural.” (*In São Paulo, social prejudice is rather common, while racial or cultural prejudices do not necessarily exist.*) The inverse order of the other two speakers, seeing the presumably poorer S I more positive than S III, however, contradicts this interpretation. It seems, as though the social stereotype, the group style associated with S II’s speech, is seen as very negative by the entire society of São Paulo.

#### 7.4 Summing up

---

The perception test was conducted to see how the Bolivians’ different linguistic strategies identified by the qualitative analysis are perceived by São Paulo’s society, whether considerable differences between the speakers would be detected and whether social stereotypes based on their speech could be recognized. Further, I wanted to obtain explicit data about the linguistic features the *Paulistanos*

considered as decisive for their evaluation, i.e. whether the same were identified as socio-indexical for the different groups as I had found in the qualitative analysis.

On the basis of Carvalho's test (2008), the survey referred to three parts, the linguistic, the social and the personal evaluation. The linguistic classification contained the answers to various questions, concerning the labeling of the language, the estimated place of origin and explicit questions about the linguistic features that conditioned the informants' answers.

The analyses revealed that S I and S III were classified along the same lines: The majority judged that both speakers were foreigners coming from a Spanish-speaking country and called their language *Portunhol* or *Portuguese with foreign influence*. The difference between them: S I was evaluated more often as Latin American (or Spanish) speaking *Portunhol*, where S III got more marks labeling his speech as Portuguese or even another language. With regard to explicit linguistic markers, certain phonemes like the nasal diphthongs, the open vowels and the voiceless sibilants were named as well as the insertion of Spanish vocabulary and intonation.

The perception of S II differed strongly from the other two, because the majority of the *Paulistanos* evaluated him as a native Brazilian speaking Portuguese. For most of the informants, S II comes from one of the poorer, socially unprivileged areas of São Paulo or even the poorest area of Brazil, the Northeast. This evaluation was generally justified by his use of Common Expressions, in the informants' words *gíria*, and discourse markers. This means that these features have a strong second order social indexicality. They are obviously connected to a social stereotype as additional comments on S II's speech confirmed. For the other two, the opinions about their social status remained ambiguous at this point.

Regarding the explicit questions about the social classification of the speakers by three semantic differentials and their place of residence, the vast majority of the informants think that S II lives in socially rather unprivileged areas. S III, though, is seen by most as living either in a rather privileged area or the immigrants' traditional quarter, *Bom Retiro*. About S I, the opinions go in different directions, but more than 20% think that he comes from *Bom Retiro*, emphasizing once more that S I is most clearly associated with the recent immigration to São Paulo.

These results are further confirmed by the semantic differentials, displaying a strict hierarchy among the speakers: S II is the laziest, poorest and accordingly<sup>185</sup> the most ignorant speaker; S I is more

---

<sup>185</sup> The correlation between prosperity and knowledge is highly significant.

ambiguous but also rather lazy, poor and ignorant, but S III is considered as hard-working, rich and well informed. Differentiating the answers along the social groups of the informants, it is revealed that students and people living in socially unprivileged areas judge S II significantly poorer and more ignorant than the others, although (or because) they obviously perceive him as having their own social status.

The personal evaluation according to likeability, sociability and attractiveness revealed a different order of the speakers, seeing S I as most positively judged, S III as more or less negative and S II again as most negative. Accommodation theory's hypothesis that people tend to evaluate those positively who speak similar to them could partly be confirmed (S II, S III), as well as the findings of psycholinguists that social status and solidarity stand in an inverse relation (S I and S III). The overall negative evaluation of S II with regard to social status and to personal classification, however, reveals that the style S II represents is judged very negatively on every level, possibly pointing at social discrimination inside São Paulo's society.

Since S II is taken for a Brazilian and the others for foreigners, there is no indication for the association of the Bolivians' accent with negative social stereotypes and hence no ethnical discrimination, at least not based on speech.<sup>186</sup> On the contrary, foreigners are positively evaluated if they speak PC as S III does, associated with a high level of education and hence effort and prosperity.

---

<sup>186</sup> Discrimination because of the appearance of the Bolivians cannot be assessed, of course.

## 8. The Portuguese of Bolivian immigrants

---

This chapter discusses the results of the qualitative and the quantitative part of the study, resorting to the theoretical approach of network-oriented accommodation based on the sociological approach of assimilation. Because of the linguistic differences that were shown by the qualitative study, three groups were established that differ with regard to their degree of assimilation in terms of the four dimensions: cognitive, social, structural and identificational assimilation. As the educational background conditions the other dimensions, I have to distinguish between workers and academics. Additionally, assimilated and non-assimilated workers are discussed separately. The latter are strongly oriented towards the Bolivian contacts of their network: This means that linguistic accommodation to Portuguese, here cognitive assimilation, is of little importance to them, while L1-transfer is highly influential. The assimilated workers differ from this first group as they assimilate socially and cognitively, and hence show linguistic accommodation towards PP, particularly with regard to phonology and vocabulary. The academics are structurally assimilated because of their social status. Linguistically, this signifies accommodation towards PC, the standard Portuguese. On the other side, the social pressure to accommodate seems to be less prominent in their network. Therefore, non-accommodation can be observed with regard to phonology. The evaluation by *Paulistanos* confirmed this last result by accepting Castilian transfer for the speaker whose social status is perceived as high. Thus, not only my hypotheses about accommodation and social indexicality are verified, but also the estimation that the Bolivians aim for different linguistic varieties of Brazilian Portuguese. Accordingly, the view of *PC* as the only prestigious variety for the Bolivians is rejected.

---

As demonstrated in the last chapters, the linguistic analysis revealed a difference between speakers with academic grades working as physicians or engineers and the persons having come to work in the textile industry, even if they additionally work as sales assistants or have stopped working as seamstresses because they have become *donas* themselves. As described in Chapter 3, this bisection is obviously preconditioned by the socioeconomic status of the immigrants in Bolivia, i.e. their pull-factors for migration differ greatly from the second group of immigrants. Additionally, they have access to educational institutions in Brazil; the academics all frequented Brazilian universities.

Inside the group of workers<sup>187</sup>, the linguistic analysis revealed huge differences, even if at first sight they seem to constitute a very homogeneous group according to rough social categories: All of them live in the poorer quarters of São Paulo and work in the textile industry. Most of them came to São Paulo in their twenties because of economic reasons, mostly via contacts with other Bolivians already living in Brazil who brought them over; none of them took a language course at Brazilian universities or schools. Linguistically, though, some speakers showed much transfer from L1, alterna-

---

<sup>187</sup> As the sweatshops of the *donas* are small entities where they still sew themselves if they have too much work, I count them as workers, even though they have become entrepreneurs.

tion and insertion into the L1 matrix language, i.e. they obviously lean towards the Castilian-speaking contacts in their translingual network and show no signs of assimilation, neither social nor cognitive. Thus, I call them the ‘not assimilated workers’ (NAW). The other workers, though, have Brazilian-speaking strong-tie contacts in their networks. Since they lean towards these contacts linguistically and socially, they are called ‘socially<sup>188</sup> assimilated workers’ (SAW). The academics are united by the fact that they are structurally assimilated, in particular because of their professional position and their formal education. Thus, they are categorized as structurally assimilated academics (StAA).

In the following discussion I will examine how these dimensions of assimilation are connected with the Bolivians’ Portuguese. Resorting to my theoretical approach the linguistic features are attributed to the main linguistic processes of accommodation, transfer and simplification. Additionally, the influence of the social network is investigated. Here, I subsume factors which influence the network like social assimilation with regard to Brazilian strong-tie contacts, the immigrant’s identificational assimilation with regard to her attitude and her cognitive assimilation with regard to geographic mobility. The results of the quantitative survey about language regard will then be presented as the voice of the receiving society. The perception and evaluation of the linguistic assimilation strategies by the *Paulistanos* will be discussed in addition, focusing on whether they verify or falsify the results of the qualitative study, in particular my estimations about accommodation by socio-indexical features.

### 8.1 The not assimilated workers (NAW)

---

*Ele não sai não se comunica não se abre com [...] a cidade com o mundo;  
ele fica aí no seu cantinho (A18 (30) m. 2004 P. Ce: 9.31).*

As this quotation by one of the academics already shows, the NAW do not take part in São Paulo’s urban society, because they do not seek contact to Brazilians outside of their Bolivian networks. Because of the rapid growth of the Bolivian community to its current size, their “*cantinho*” however is pretty vast. The Bolivians already have shaped their own public places in São Paulo. Additionally, the NAW maintain close contact with their home country, going back to Bolivia to return to São Paulo a year later, forming actual transnational networks. The size of São Paulo enables them to live in the city without assimilating; they only add to the heterogeneity of the city.

---

<sup>188</sup> The following discussion will show if and how the other dimensions of assimilation are valid for these speakers, too.

I postulated in the theoretical approach that close-knit networks with many contacts from one's own community are prototypical for those who rather orient towards the Bolivian side of their transnational network. This hypothesis, as well as the suggestions for the linguistic characterization of this group, will be verified by the empirical data presented before.

### 8.1.1 The social network

---

The social network of the NAW is transnational, many of them plan to go back to Bolivia and maybe to return after some time:

*Eh:: eu vou sair a Bolivia fin de año [...] éh: talvez va voltar acá<sup>189</sup> (B11 (29) f. 2008 S. Pe: 2.37).*

A24, for example, had been in São Paulo for four years at the time of the interview and stated that he has already gone back to Bolivia three times: “Ah já voltei ya já [...] Já voltei três vezes ya” (A24 (25) m. 2004 S/SA Ce: 12.22) and B9 also confirms: “Me fui pra lá, eu volté, volté con él, es mi irmão”<sup>190</sup> B9 (24) m. 2005 S. ZL: 10.58). It is common that the children are left in Bolivia, even if they are born in São Paulo and thus have the Brazilian nationality. All these factors show that they are strongly oriented towards their home country. This close connection with Bolivia also reveals that there is no identificational assimilation among the NAW towards the Brazilian community. They know very little of Brazil and São Paulo because of the problems already mentioned: Their mobility is hindered by the extreme working hours, by the expensive public transport and often by fear of crime and discrimination or of being detected by the federal police if they are still illegal:

*Nó podia sair por medo que não tinha documento brasileiro. [...] Porque a gente aqui é criticado assim é discriminado assim (B6 (30) m. 2003 S. ZL: 2.30).<sup>191</sup>*

Thus, they move in the limited zone that is dominated by the presence of Bolivians, *Bom Retiro*, *Brás* and the historical center: “Venho de Bom Retiro a Brás, a centro, a Praça da Sé [...] a Vinte e cinco eu conheço” (B12 (21) m. 2005 S. Ce: 15.31).<sup>192</sup>

Therefore, additionally to their lack of structural assimilation with regard to legality, income etc., their cognitive assimilation in terms of geographic mobility is also highly restricted; they are well

---

<sup>189</sup> Paraphrasing in English: “I will return to Bolivia at the end of the year. [...] Maybe, I will return here [to Brazil, sn]” (own translation, sn).

<sup>190</sup> Paraphrasing in English: “I went there, I returned with him; he is my brother” (own translation, sn).

<sup>191</sup> Paraphrasing in English: “Out of fear that one had no Brazilian documents one could not go out. [...] Because here, we are criticized this way, we are discriminated this way” (own translation, sn).

<sup>192</sup> Paraphrasing in English: “I come from Bom Retiro to the Brás, to the city center, the Praça da Sé [...]; the 25 (25 de Maio; commercial street in the city center, sn) I know (own translation, sn).

informed, though, where they can make cheap calls to Bolivia, where they can book cheap travel home etc., underlining once more their orientation towards Bolivia.

With regard to social assimilation, i.e. interethnic social relations like friendship, marriage or further associations (cf. Bommers 2005: 20), the NAW are again not assimilated since they have very little contact with Brazilians. We just saw that B9 came back to Brazil with his brother. Since the dense migrational networks between São Paulo and Bolivia are used to attract workers, many of them have family members in Brazil with whom they work and live. Regarding relationships, not one of the NAW has a Brazilian boyfriend or girlfriend let alone husband or wife. B4, for example, confirms that an intercultural relationship is complicated, because he does not understand the Brazilian women's openness:

*Agora tô arrumando uma [esposa] [...] uma boliviana é mas melhor [...] porque brasileira acho que brasileiras são abertas, não? Então, não sei não consigo entender com elas [...] melhor é boliviana é com ela me intendo bem (B4 (27) m. 2003 S. Ce, S. 8.50).<sup>193</sup>*

If they take part in groups or associations, these are also Bolivian, e.g. folklore groups (B6 (30) m. 2003 S. ZL: 0.23). Two other speakers got to know each other in São Paulo and married there, although their families live in the same quarter in La Paz (A28 (28) m. 2002 S. ZL: 1.34).

Considering all these factors it is not surprising that most of the NAW answer the question if they have Brazilian friends with an outright *não* or *nó* (A26 (25) f. 2005 S. ZL: 5.50; B9 (24) m. 2005 S. ZL: 13.33) or like that: “Hay amigos só que não lembro muito bem” (A28 (28) m. 2002 S. ZL: 5.52).<sup>194</sup> If there is a Brazilian connection, it seems to be rather weak. When asked if he has Brazilian friends, B6 responded like that:

*Com brasileiro quase no, as amizade assim 'oi tudo bem tá legal como que vai' só assim. mas compartilhar assim só com boliviano (B6 (30) m. 2003 S. ZL: 9.08)<sup>195</sup>*

Thus, focusing only on the NAW's network in São Paulo without considering the connection to Bolivia, there is indeed a very close-knit, multiplex social network that consists almost entirely of Bolivian contacts:

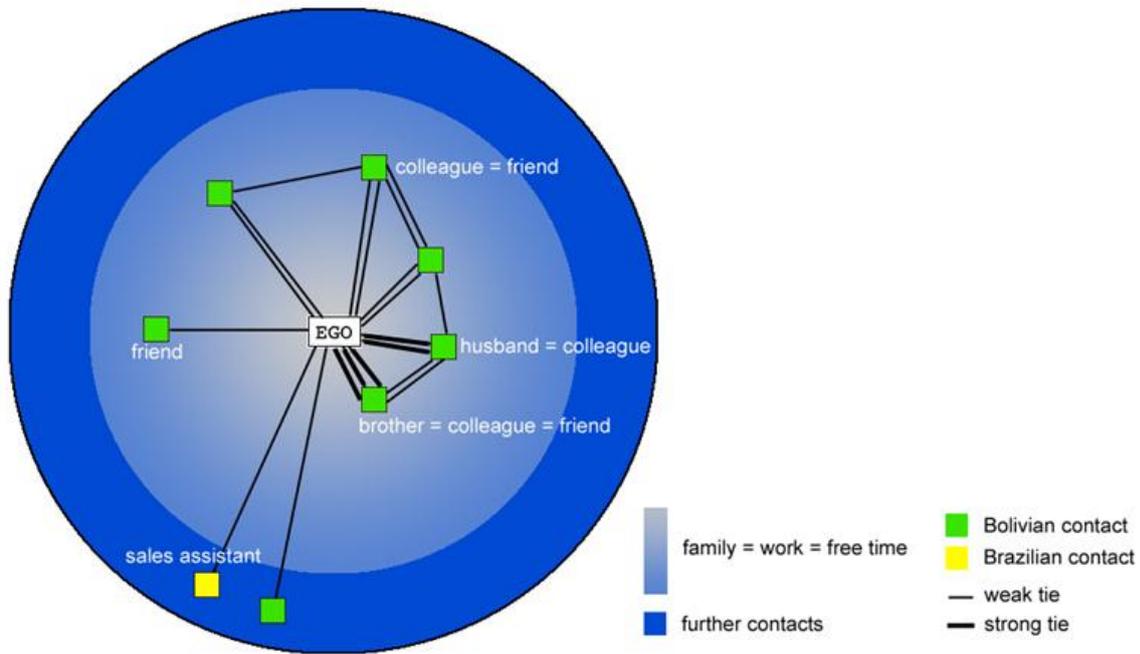
---

<sup>193</sup> Paraphrasing in English: “Now, I am looking for a wife [...]. A Bolivian wife is better [...], because I think that the Brazilian women are open-minded, right? So, I don't know, I cannot get along with them. [...] A Bolivian girl is better, with her, I get along well” (own translation, sn).

<sup>194</sup> Paraphrasing in English: “There are friends, I just do not remember them very well” (own translation, sn).

<sup>195</sup> Paraphrasing in English: “With Brazilians rather not, only friendships like that ‘hey, how are you, everything ok, how are you’, but actually sharing, only with Bolivians” (own translation, sn).

Figure 8.1: Prototypical network of one of the NAW (EGO-centered<sup>196</sup>).



Due to the working and living conditions of the NAW, the boundaries between the different areas of life and the different CoP like family, work and free time are blurred because the place where they work, live and spend their few hours of leisure time are the two rooms of the sweatshop. Of course, this leads to a reduction of the space where contacts and communication outside the sweatshop and hence outside the Bolivian community can occur. Consequently, the opportunity to form new contacts is reduced, while the relationships between the existing contacts become very complex: The same person is not only the husband of the EGO, but at the same time her colleague at work, her brother is the *dono* of the sweatshop, and all sleep together in one room.

Even the social contacts which happen by accident are mostly with Spanish-speaking persons, because in the quarters the Bolivians live in the neighborhood is dominated by compatriots, even bars or markets are led by fellow countrymen. Of course, contact with a Brazilian salesman or bus driver is possible, but this is almost completely negligible from a communicative perspective because communication can be avoided and being socially approved by the salesman or bus-driver is not important to the NAW. Without a single regular strong-tie contact with a Brazilian, there is obviously no social assimilation in case of the NAW.

<sup>196</sup> EGO-centered signifies that the network is built from the point of the individual, in this case the NAW, in the center of the network.

Thus, the postulated preconditions for the non-accommodation of the speaker are confirmed: the more influential members of the immigrants' original community the social network contains, the higher is the probability that the speaker non-accommodates.

Correspondingly, the NAW evaluate their own Portuguese proficiency as rather poor because communication happens almost exclusively with other Bolivians as A25 confirms when asked how he judges his own proficiency in Portuguese:

*Sei como aqui con colegas, en la casa con colegas, [...] con gente de nuestra terra [...] no más a gente fala [...] no aprende bien falar” (A25 (60) f. 1989 S/SA ZL: 9.12).<sup>197</sup>*

They even claim that they do not understand the Brazilians, and that they are not understood vice versa:

*Também hay alguns brasileiros que hablan rápido não dá pra intender (B9 (24) m. 2005 S. ZL: 5.45).<sup>198</sup>*

*El miedo de hablar errado portuguê[s] [...] dificulta yo nó falo muito bem [...] então em uma conversação com um brasileiro se tem que falar bien. Se você não fala bem ele não intende você (B5 (19) m. 2008 S. ZL: 5.35).<sup>199</sup>*

Even though many consider the lack of Portuguese skills as problematic, they do not regard it as too important: Because of the similarity of the languages the basic aims of communication, at least passive knowledge, can be reached easily without much language proficiency.

Thus, in their urban space and transnational network strongly dominated by the Bolivian presence, Portuguese plays just a subordinated role, as this citation affirms:

*pero meus filhos [...] eles tem que aprender meu idioma espanhol primeiro é segundo que viene portuguê[s] não pode aprender diretamente a portuguê[s], seria mal né não saber sua idioma legítima e saber outra idioma<sup>200</sup> (B4 (27) m. 2003 S. Ce: 9.41).*

However, they participated in this study and I talked to them in Portuguese. Therefore, what is the Portuguese of the NAW like? Which linguistic processes can explain their speech?

### 8.1.2 Linguistic characterization

---

---

<sup>197</sup> Paraphrasing in English: “I know that, as we speak only with colleagues here, at home with colleagues, with people from our home country, we do not learn to speak well” (own translation, sn).

<sup>198</sup> Paraphrasing in English: “Additionally, there are some Brazilians who speak quickly; it is not possible to understand them” (own translation, sn).

<sup>199</sup> Paraphrasing in English: “The fear of speaking Portuguese not correctly makes it difficult. I do not speak Portuguese well. But, in a conversation with a Brazilian you have to speak well. If you do not speak well, he will not understand you” (own translation, sn).

<sup>200</sup> Paraphrasing in English: “But my sons [...], they have to speak my language, Spanish, first. Portuguese comes second. They cannot learn Portuguese directly. This would be bad, right, not to know your own legitimate language and to know another language” (own translation, sn).

In the theoretical approach I hypothesized that for speakers in the transnational networks the first decision is if they actually move in the direction of the new country and subsequently of the new language or if they maintain their original language as far as possible. Considering figure 8.1., the network of the NAW is not necessarily translingual, because the necessity and the opportunities to speak Portuguese at all are limited.

Correspondingly, I hypothesized that their speech would primarily be characterized by L1 transfer, code-mixing and SLM rather than by accommodation towards the Brazilian varieties. Having a close look at the results of my qualitative analysis, this hypothesis is confirmed although simplification plays a further important role.

#### 8.1.2.1 Accommodation

---

The NAW do not accommodate towards any of the varieties of Brazilian Portuguese. According to my theoretical approach, this is the first dividing line between the immigrants. In contrast to the other Bolivians, the NAW maintain Castilian as main language, noticeable in the many cases of code-mixing in particular, and the few cases of actual accommodation to the Brazilian varieties. Features that were classified as accommodation like the palatization of *-ti/di* and the aspiration of /r/ at the syllable onset are not regularly realized by the NAW, instead the realization according to Bolivian Castilian is found as in (96) and (101). An exception constitutes the diphthongization of final /s/ because in (135) there is even a case of hypercorrection of La Paz as [la'pais] by A28.

They also adopt Brazilians' pronominal system of *você* and *a gente* although *a gente* is often used in its original semantics of Castilian *la gente*, but most of the time with PP phonology, i.e. including palatization. On the other hand, the use of *nosotros* (233) and of *nós* (238) demonstrates once more that the dominant process is L1-transfer and not accommodation. In addition, the absence of morphological agreement (number, gender and verbal agreement) is frequently observable in the speech of the NAW. With regard to number and verbal agreement, this could be counted as accommodation towards PP, but the further absence of gender agreement points rather in the direction of various factors like CP-transfer as well as morphological simplification as a universal process in oral languages that operate in the same direction in terms of the weakening of agreement rules.

Additionally, the NAW are the only ones who do not employ Portuguese discourse markers, although they clearly belong to the most salient features and are therefore highly socio-indexical. This underlines once more that the NAW do not accommodate to Portuguese; on the contrary: The use of

Bolivian discourse markers is a clear sign of non-accommodation. With regard to the network of the NAW as illustrated in figure 8.1 though, this non-accommodation towards Portuguese is only logical and in accordance with my theory: Seeking social approval by the members of their social network they perceive as socially successful, the Bolivian NAW accommodate towards these contacts and they regard themselves as Bolivians not as Brazilians.

Thus, from the point of view of the Bolivian community, the process now discussed as L1-transfer would constitute accommodation, but from my point of view it constitutes the most obvious sign of non-accommodation.

#### 8.1.2.2 L1-transfer

---

It was repeatedly emphasized that L1-transfer must not be understood as one-to-one imposition but in the terms of Siegel (2008) as the transfer of individual elements or aspects from the first language. In this case, different types of this transfer can be distinguished that relate to different parts of the speaker's language: Referring to Krefeld's model of *Migrationslinguistik* (2004), the first type concerns the speech of the Bolivians, while the second concerns their speaking:

With regard to speech, there is transfer of structural aspects that becomes manifest in the realization of particular linguistic features, for example of /z/ as [s] or gender attribution and gender agreement. With regard to speaking, though, the code-mixing between Castilian and Portuguese is the dominant type of transfer.

Beginning with the transfer influencing the speech, again different processes of how transfer actually influences the speech of the Bolivians can be distinguished. There is structural transfer that can be explained by Flege's SLM (1995; cf. also Bohn 1998) with regard to phonological features. The similarity between the sounds in Castilian and Portuguese impedes the perception and therefore the production of the phonemic distinction. This is clearly the case with regard to the sibilants, but also the variation between open and closed /ɛ/ and /e/ and between /ɔ/ and /o/ and the free variation between /e/ and /i/, and /o/ and /u/. Naturally, the latter could also be influenced by the reduction of the vowels in atonal position in PP and therefore also signify some case of accommodation to PP.

Akin to these phenomena, the variation regarding nasal diphthongs, mostly in analogy with the corresponding Castilian vocabulary, indicates L1-transfer that also could be due to difficulties in the perception and following production of the sound because of similar sounds in Castilian like [an] [on] or even slightly nasalized forms like [ã<sup>n</sup>]. The fact that this form has not yet been described for

Bolivian Castilian might be due to the relatively few extensive phonological studies about the varieties in Bolivia and not to its absence in the Bolivians' speech. In this case, though, the process could also be described as congruent lexicalization (cf. Muysken 2000), i.e. the use of Portuguese vocabulary in combination with Castilian phonology even if SLM is the cognitive reason underlying it.

Until here, there is no difference according to social factors. Thus, these processes are obviously valid for all speakers and must be seen as linguistic factors that are probably universal in language contact situations.

A slightly different type of L1-transfer with regard to the NAW's speech that mostly occurs together with further factors like simplification conditions the speakers' realization of morphological structures. Examples in this regard are cases of absence of nominal agreement (number and gender), *incorrect* gender attribution, and the omission or confusion of prepositions. Although there are examples of direct L1-transfer like false gender cognates (cf. (188) and (189)), the other examples from (197) to (200) are not explainable by simple transfer of the Castilian structure to Portuguese, because in these cases both languages would require the same gender and gender agreement. Even though the absence of gender agreement is discussed for CP, this does not explain the fact that it occurs for almost all speakers, even monolingual academics<sup>201</sup> who have come to Brazil 20 years ago. Thus, it seems plausible that the category of gender is already weakened in the Castilian of the Bolivians. In the situation of language contact, this element is transferred to Portuguese and additionally reinforced by the generalizing process of morphological simplification, probably in analogy to the absence of number agreement frequent in PP.

Correspondingly, prepositions in CP are obviously no obligatory category; in the case of contact with Portuguese, this structural information is transferred and the universal internal tendency of languages for simplification further strengthens the process in the direction of omission of prepositions. The confusion of the prepositions might hence stem from the attempt to meet language norms of PC without actually having the information which preposition to choose, even if they were equal in CC. Thus, these examples are proof against structural one-to-one imposition, particularly one which considers only the standard varieties.

The case of SOV word order is similar: Although not all speakers display this feature, I cannot confirm that only bilingual speakers from rural areas realize it, quite the contrary. Thus, I suggest that

---

<sup>201</sup> *Monolingual* refers here only to the languages that were learnt in Bolivia, i.e. Castilian and indigenous languages.

the element that is transferred from L1 is a freer word order than in Portuguese and CC: SVO is not obligatory anymore but nonetheless preferred.

The examples discussed so far are common for speakers from all social backgrounds and subgroups. But the L1-transfer with regard to the Bolivians' speech is the unique feature of the NAW: Code-mixing as insertion of PP features in the Castilian matrix language, as in the examples (376) to (378), and congruent lexicalization where single words are actually translated into Portuguese, while the remains are maintained in Castilian. This technique underlines that the Bolivian part of the NAW's transnational network is clearly more important than the other part. Because of the typological closeness of the languages, the communication with cooperating Brazilians is achieved by little linguistic concessions, in particular regarding vocabulary and phonology, while morphosyntax is maintained in Castilian or simplified as described further on.

#### 8.1.2.3 Morphological simplification

---

The role of the universal tendency of morphological simplification as a kind of catalyzer of structural CP-transfer in the speech of the Bolivians was already discussed, particularly in the case of prepositions and gender agreement / gender attribution. Naturally, this process is strongly connected to orality and the lack of educational institutions that impede the radicalization of the process. Correspondingly, morphological simplification is widely spread among the NAW because structurally they are not assimilated, i.e. they have no access to institutions or written linguistic norms.

#### 8.1.3 Evaluation by *Paulistanos*

---

For the tests of language regard, S I was selected as most representative of this group because he showed most insertion of Castilian vocabulary and the most prominent transfer from Castilian. In comparison to the NAW though, who actually insert PP vocabulary into the Castilian matrix language, his example seems rather distant from their speech. Thus, his rather positive evaluation might not be too revealing with regard to the estimation of NAW by *Paulistanos* but rather confirms that speech influenced by L1-transfer from Castilian alone provokes no negative stereotype. However, the fact that examples of the speech of actual NAW with stronger Castilian influence were declined already by the piloting group as "too difficult" and "not understandable", might reveal the actual evaluation of the NAW's speech.

Additionally, some informants actually talked about the Bolivian immigration, even though this could not be further reflected in the statistical analysis. Below, two examples from the questionnaires:

*[...] para trabalhar em uma vaga onde fosse necessário um bom português brasileiros ele provavelmente teria dificuldades e com certeza sofreria muito preconceito. Como desconfio que seja um boliviano, somente a sua aparência étnica já seria o suficiente para decrescer o valor da sua mão de obra entre os paulistanos, sua imagem estereotipada está muito associada ao trabalho escravo, a ilegalidade e ao subemprego. [...] [A imigração recente] acho excessiva por ser descontrolada, e isto acaba inflando a informalidade e contribuindo para o não cumprimento das leis trabalhistas<sup>202</sup> (Paulistano, 26, 13.5.2010).  
[...] Mas sobre a atual imigração [...] bolivianos descobertos como escravos de confecções textil em São Paulo [...] Mafia chinesa! [...] Sou contra esse tipo de imigração que suja o país<sup>203</sup> (Paulistana, 26, 28.6.2010).*

Thus, the NAW are seen rather negatively, even if these examples are not representative.

#### 8.1.4 Resume

---

Summarizing the linguistic and social characteristics of the NAW, they are neither structurally, nor identificationally, nor socially assimilated. Their contact with Brazilians is strongly limited, because they are oriented towards the Bolivian side of their transnational network. Correspondingly, they do not assimilate cognitively, they do not learn Portuguese. Because of the closeness between Castilian and Portuguese, only small adjustments seem to be enough to ensure that they can fulfill the little communicational needs when speaking with Portuguese-speakers. These adjustments concern mainly vocabulary and phonology, e.g. the pronouns. But as they clearly non-accommodate to their Brazilian contacts (if existing), social indexicality is not influential in direction of Portuguese. On the contrary, their code-mixing and e.g. use of Castilian discourse markers emphasizes their identity as Bolivians. As the interviews were held in public places where Bolivian bystanders might have listened, this linguistic behavior can be interpreted as accommodation to them and not to the interviewer, or to

---

<sup>202</sup> Paraphrasing in English: “He would probably have difficulties if he had to work in a job where a good Brazilian Portuguese was necessary, and would certainly suffer of many prejudices against him. As I suspect that he is Bolivian, his ethnic appearance alone would be enough to diminish the worth of his manpower among the *Paulistanos*, his stereotype is strongly associated with slave work, illegality and the lack of employment. [...] I think, the [recent immigration] is excessive, because it is uncontrolled; and this ends up augmenting the informality and not observing the rights of the workers” (own translation, sn).

<sup>203</sup> Paraphrasing in English: “But [talking] about the current immigration [...] Bolivians that are discovered as slaves of the textile industry in São Paulo. [...] Chinese mafia! I am against this type of immigration that pollutes the country” (own translation, sn).

non-present third persons still influencing the speech of the NAW at the time of the interview.<sup>204</sup> For real conversations, to make friends with Brazilians, though, these little adjustments are not enough, but as stated above, the *cantinho*, the corner of the NAW in São Paulo itself is big and heterogeneous enough to compensate for the little assimilation, not to speak of the strong relations to Bolivia. From the outside, though, the NAW seem either lamentable or are criticized and discriminated for their lack of assimilation, not only by the *Paulistanos* but often by their fellow countrymen, too.

## 8.2 The socially assimilated workers (SAW)

---

*Os meus melhores amigos [...] são os Brasileiros (A13 (33) m. 2000 D. ZL: 2.42).<sup>205</sup>*

As this quote expresses, these speakers differ from the NAW, because at least they are socially assimilated. This means that they are rather oriented towards the Brazilian speaking side of their transnational network, that they have friendships or even relationships with *Paulistanos*, even if their life circumstances are similar to those described above.

### 8.2.1 The social network

---

As the SAW maintain their contacts in Bolivia, the network of the SAW is also transnational, but the workers do not focus as much on their home country and going back to Bolivia as the NAW. Many of them explicitly say that they do not want to return to Bolivia because they are socially more integrated than the NAW:

*[...] nó dá pra morar lá porque a gente já está acostumado aqui às comunidades que a gente tem [...] decidiu nó mais voltar pra lá, só pra ir visitar a família<sup>206</sup> (A9 (42) m. 1994 D. ZL: 15.13).*

A6 also claims that he had problems in São Paulo in the beginning because he knew no one. Hence he returned to Bolivia, but came back once again, and started to like it when he made friends: “Com conhecer pessoas aí ya realmente ya comecei a gostar” (A6 (34) m. 1999 S. ZN: 1.06). For A5, the most important reason for staying is that in São Paulo she has work, and therefore as long as she can she will stay in Brazil:

---

<sup>204</sup> Bell (2001: 147) applies his concept of audience design and initiative style-shifting to “referee design”, i.e. styles that “derive their force and their direction of shift from their underlying association with [...] often-absent reference group [...] rather than the present addressee. Referees are third persons not usually present at an interaction but possessing such salience for a speaker that they influence style even in their absence”.

<sup>205</sup> Paraphrasing in English: “My best friends [...] are the Brazilians” (own translation, sn).

<sup>206</sup> Paraphrasing in English: “It is not possible to live there [in Bolivia, sn], because we are already used to being here, to the communities we have. [...] We decided to not return there, only to visit the family” (own translation, sn).

*Se é oportunidade para ficar eu fico. [...] Se nó nó é necessário assim [...] Se donde onde tem trabalho eu vou pra atrás<sup>207</sup> (A5 (28) f. 2000 D. ZL: 10.52).*

Not one of them has the goal of going back to Bolivia in the near future and even if they have their family there, the connections in Brazil are stronger. Akin to the NAW, many of them work and live together with family members or are married to Bolivians with whom they also work together in the textile industry. If they have children they grow up learning Portuguese, at least at school. As the children correct the Bolivian adults in their speech, this is an important factor influencing the SAW's Portuguese:

*A verdade eles [os filhos, sn] que me estão ensinando mais né porque se alfabetizaram aqui<sup>208</sup> (A13 (33) m. 2000 D. ZL: 8.38)*

A2, who worked as nanny for children growing up in Brazil, also states that she learned Portuguese with the children:

*As crianças me falaram que não se fala desse jeito, se fala do outro jeito<sup>209</sup> (A2 (18) f. 2005 SA ZL: 2.19).*

Most of them are single, but one of the *donos* is married to a Brazilian woman and one of the seamstresses has a child with a Brazilian. These examples emphasize the different quality of the social network of the SAW and the NAW, although not all SAW are as well connected as these two particular persons.

B3, for example, only has two friends in São Paulo:

*Tem dois brasileiros que eu conheço dois somos [...] de conhecer na rua eles vivem perto da sua casa do meu tio<sup>210</sup> (B3 (21) m. 2007 S. ZL: 4.19).*

Others also confirm that they are friends with Brazilians, interestingly in both cases with people from the northeast of Brazil because they live in the same geographic region as the Bolivians do:

*Nessa favela era a gente brasileira era como uma família todo mundo hablava saludava ninguém olhava. Todo mundo associable [...] <sup>211</sup> (A8 (27) f. 2005 R. Ce: 0.54).*

*Eu morei com uma brasileira cinco anos. Ela era de acho de Ceará então eu fiz amizade com elas. [...] Eu morava lá e final de semana a gente ia para chácara divertia muito os*

---

<sup>207</sup> Paraphrasing in English: "If there is the opportunity to stay I will stay. [...] If not, it is not that necessary. [...] Wherever there is work, I will go along (own translation, sn).

<sup>208</sup> Paraphrasing in English: "Actually, the children are rather teaching me, because they were alphabetized here" (own translation, sn).

<sup>209</sup> Paraphrasing in English: "The children told me that one does not speak this way but the other way" (own translation, sn).

<sup>210</sup> Paraphrasing in English: "There are two Brazilians I know, we are two [...] by meeting them on the street, they live near my uncle's house" (own translation, sn).

<sup>211</sup> Paraphrasing in English: "In this favela, there were the Brazilians, they were like a family; all talked to each other, greeted, no one watched. Everyone very sociable" (own translation, sn).

*cinco anos eu morei com elas a mãe era muito amiga minha*<sup>212</sup> (A7 (32) f. 1994 S. Ce: 2.27).

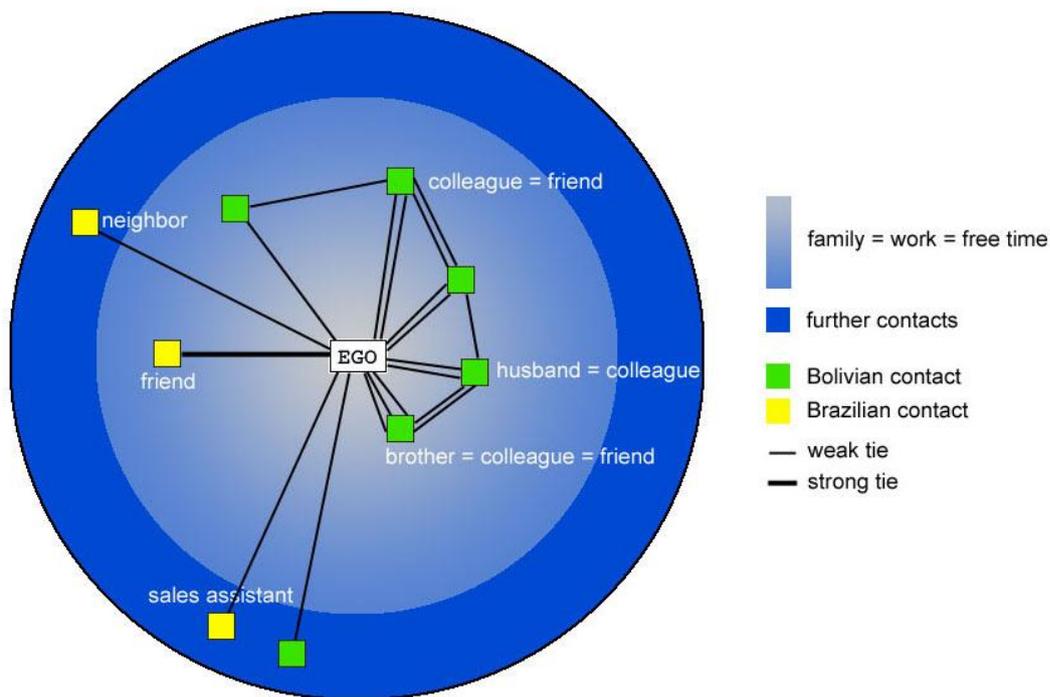
The SAW are more mobile than the NAW, some even travelled outside São Paulo, even if mostly because of work: “I was living there near the city Americana, in the interior of São Paulo (“tava morando lá na Americana que é o interior de São Paulo” (A4 (21) m. 2006 S. ZL: 0.46), own translation, sn). Due to the quite common regular change of working places, they have been to various districts in São Paulo, all of them located either in the typical Bolivian districts *Bom Retiro* or *Brás*, or in the poorer areas of the *Zona Leste*, *Norte*, and the *Periferia*. Since they are in the same economic situation as the NAW, they also have little leisure time and no money to use public transport. But still, they frequent the Bolivian places like the markets at *Praça Kantuta* and *Rua Coimbra*, and also know the city center and the parks. Thus, a little more cognitive assimilation with regard to mobility can be observed. Additionally, the following citation underlines that they are rather oriented towards the Brazilian side of their network and not only the Bolivian community: “Nó estamos muito comunicados com ((0,9s)) com aí Kantutas”<sup>213</sup> (A10 (22) m. 2006 S. Pe: 7.45). Particularly for those living in the periphery, the Bolivian infrastructure is not as easily reachable as for the people living in *Bom Retiro* or other central districts. Furthermore the neighborhood in the periphery is not dominated by Bolivians, a fact that also leads to a slightly different social network of the SAW:

---

<sup>212</sup> Paraphrasing in English: “I lived with a Brazilian woman for five years. She was, I think, from the *Ceará*. So, I became friends with them. [...] I lived there and at the weekends we went to the country house; we had a lot of fun during the five years I lived with them; the mother was a very good friend of mine” (own translation, sn).

<sup>213</sup> Paraphrasing in English: We do not communicate much with [the Bolivians] at *Praça Kantuta* (own translation, sn).

Figure 8.2: Prototypical network of one of the SAW (EGO-centered).



In figure 8.2, there are obvious similarities with the social network of the NAW because of the same economic situation of the two groups that strongly determines their way of living. The fluent transition among the various communities of practice family, work, and leisure time is again due to the low degree of structural assimilation of the speakers, i.e. the illegality, the amount of working hours, and the bad financial situation. Therefore, one half of the network is multiplex and very close-knit. The other half, though, and this makes the decisive difference between the NAW and the SAW, is rather loosely knit and various Brazilian contacts are present here. But these contacts have not only increased quantitatively but particularly qualitatively: The SAW have at least one strong-tie contact to a Brazilian person, reflecting the social assimilation of those speakers.

Naturally, some of the *donas* or the seamstresses who have spent many years in São Paulo might have more contacts with the Brazilian community. Since they send their children to school, they participate in the educational system of Brazil. One of the seamstresses participates in an NGO, thus for some of them, a higher degree of structural assimilation can be observed, too. Even the identificational assimilation may be high, particularly of those who never want to return to Bolivia: “I love this country [Brazil, sn]” (“Amo esta terra” (A9 (42) m. 1994 D. ZL), own translation, sn).

For the linguistic development, only this one strong-tie contact has an enormous influence because it is the most prominent linguistic input the SAW get, since none of them has frequented any language course:

*Eu não estudei assim pra falar português nada. Só conversando com brasileiros assim, perguntando, só isso. [...] Tem amigos brasileiros ((0,7s)) então eu pergunto pra eles que quê [...] quer dizer essa palavra? [...] Ou::: ou às vezes [...] tô falando errado. Ele disse tá errado assim, não, é esse, tem que ser assim<sup>214</sup> (B4 (27) m. 2003 S. Ce: 5.38).*

Correspondingly, they are eager to get the social approval of this contact and thus accommodate to her speech as hypothesized.

### 8.2.2 Linguistic characterization

---

According to the theoretical approach, I expect that the SAW show rather less L1-transfer but more linguistic signs for the accommodation towards the Brazilian speaking part of their transnational network. As the strong-tie contacts of the SAW are often *Nordestinos* or at least people in the neighborhood of Bolivians, i.e. districts of rather low social status, the Brazilian contacts of the SAW speak rather PP than PC as A6 confirms:

*Não era português que eles falava né. Era totalmente informal e aí aprendia mas aprendia totalmente errado. [...] A maior parte que moram aqui na região Zona Leste diz que vem do Nordeste [...] eles não falam muito bem português falam solamente com gírias<sup>215</sup> (A6 (34) m. 1999 S. ZN: 27.53).*

Thus, the SAW are expected to accommodate towards PP by the features they perceive to be most socio- indexical.

#### 8.2.2.1 Accommodation

---

All of the SAW frequently use phonological features that are considered as typical for São Paulo's Portuguese, i.e. the aspiration of the vibrant at the syllable onset, the palatization of *-ti/di* and the diphthongization before final */s/*. For the latter two, there is even hypercorrection (cf. (91), (136) and (137)), and some of the seamstresses are those who aspirate [r] most regularly. This regularity already differentiates them from the NAW.

---

<sup>214</sup> Paraphrasing in English: "I did not do any language course to learn Portuguese, nothing. Only speaking to Brazilians like that, asking, just this. [...] I have Brazilian friends, so I ask: 'What does this word mean?' [...] Or sometimes, when I am saying something wrong, he says: 'It is wrong this way, ok, it's like that, it has to be like that'" (own translation, sn).

<sup>215</sup> Paraphrasing in English: "They did not speak Portuguese, right. It was totally informal, and this way, I learned, but I learned it totally wrong. [...] The majority of those who live here in the region *Zona Leste* say that they are from the Northeast. They do not speak Portuguese very well, they speak only with *gírias*" (own translation, sn. For a discussion of the term "gírias", c.f. 6.1.14).

As to phonology, though, I cannot speak of any certain indicator towards PP because the phonological markers are not socially perceived as typical for PP but mostly generalized for PC, too.

More interesting are thus morphology and vocabulary: The pronominal system of PP, including *você* and *a gente*, is adopted by all SAW, only the *dona* uses *nós*, the others regularly employ *a gente*, although they still use it in the generalizing semantics, too. The men who have spent several years in São Paulo also employ the cliticized form of *você*, *cê*. Additionally, the absence of number agreement and verbal agreement seem highly generalized for these speakers, but not the absence of gender agreement. Here, some of the seamstresses stick out: One of them even “corrects” a PC sentence with norm-following number agreement to become a PP sentence without number agreement (187). Thus, it seems that for them, in particular, the absence of number and verbal agreement is a socio-indexical feature of PP, the variety they accommodate to.

This finding is confirmed by the analysis of discourse markers and Common Expressions. All of the SAW regularly use *né*, and also *entendeu* or *viu*, but no Castilian markers. The speakers just mentioned, though, use *né* over-frequently after almost every word. Additionally, one of them is the only one who uses *meu*, the marker typically associated with São Paulo’s youngsters. Moreover, he and one other seamstress are the ones who use Common Expressions that are still socially rather tabooed, i.e. by definition they are socially indexical of a norm-distant variety. Thus, it seems that all speakers show a certain degree of accommodation towards PP, in particular by phonology, the pronouns, and discourse markers.

Among them, however, three young males can be distinguished who show a stronger orientation towards PP in every aspect: hypercorrection of PP phonology, generalization of PP morphology, and particularly the frequent use of discourse markers as well as tabooed regular expressions. Therefore, I postulate that A6 ((34) m. 1999 S. ZN), A11 ((20) m. 2002 S. Pe) and A14 ((28) m. 2006 S. ZN) are those who accommodate strongest to PP and who are therefore prototypical for the group of SAW. Interestingly, apart from their sex and their work as seamstresses, they differ quite strongly regarding their age and duration of stay in Brazil. Consequently, I deduce that the attitude of the three to participate in their urban environment, i.e. make contact with the Brazilians they meet in the poor districts of São Paulo, is the decisive factor for this enforced accommodation. This does not mean, however, that they also show least L1-transfer. On the contrary, A 14 clearly displays a high degree of L1-transfer; his speech sounds rather like Castilian structure congruently lexicalized by socio-indexical features of PP.

#### 8.2.2.2 L1-transfer

---

Looking at the L1-transfer, the same features as for the NAW are observed regarding SLM because this process obviously does not depend on social factors such as accommodation. The big difference in comparison with the NAW is, though, that the code-mixing by the SAW is rather presented by congruent lexicalization or sometimes alternation. The insertion of individual Portuguese terms into Castilian, i.e. the clear dominance of Castilian, is not observable for the SAW. Thus, my hypothesis is confirmed that their looser-knit network with more Brazilian contacts causes all of them to accommodate to Portuguese in the case of conversations with Brazilians. Naturally, this does not mean that they shift completely to PP: Their network is actually translingual. Hence, the strategy drawn for A 14, i.e. congruent lexicalization using CP structure, salient phonological features, and socially highly indexical vocabulary like pronouns, discourse markers and regular expressions, seems also to be employed by other SAW.

#### 8.2.2.3 Morphological simplification

---

Morphological simplification can be observed for the SAW the same way as for the NAW because they are also structurally little assimilated, particularly with regard to access to educational institutions and written language. Interestingly, the *dono* who claims that he loves reading and emphasizes that he has no problems regarding reading or speaking (“quanto na leitura na fala né”, A13 (33) m. 2000 D. ZL: 1.50) is the only SAW who shows not one occurrence of the absence of number agreement. This could indeed be an indication for the connection between morphological simplification and orality. Additionally, this process is even fostered by accommodation to PP because in this variety the absence of number and verbal agreement is already highly salient and socio-indexical.

#### 8.2.3 Evaluation by *Paulistanos*

---

The example representing the SAW is S II, obviously comparable to the three young men who accommodated most to PP. Stereotypically, he also uses the discourse markers *né* and *meu*, socially indexical Common Expressions like *pateta*, but little L1-transfer, although there are some recognizable features of this.

I hypothesized that the persons he accommodates to – obviously young native Brazilian speakers living in the socially unprivileged areas – would actually approve of him while others might consider his speech as rather negative. Interestingly, the vast majority of the *Paulistanos* did not realize that he

was a foreigner at all. Most of them thought he was a native speaker of Portuguese and of PP as spoken in the *ZL* or the *Periferia*. Nonetheless, he got the most negative evaluation by far regarding social as well as personal perception.

Apart from the rather astonishing finding that *Paulistanos* see their fellow countrymen much more negative than foreigners, at first sight, this result clearly contradicts the hypothesis of a positive evaluation by the persons he accommodates to. A closer look reveals, though, that the youngest age group (15-24 years) living in the socially unprivileged areas shows a significantly better result regarding likeability than the others. Given that this is exactly his peer group, my hypothesis is at least partly confirmed.<sup>216</sup> Additionally, the test shows that socio-indexical features like the discourse markers or Common Expressions can easily “confuse” the language regard of the hearers, as almost none of them heard the rather obvious L1-transfer of little salient phonological features and vocabulary. In addition, various informants even claim that the absence of agreement “*falta de concordância*” and the use of “*mano*”, another discourse marker, are the decisive linguistic factors for their estimations. Since both features are not even used by the informant, these comments show for once that the absence of verbal and number agreement is also a socio-indexical feature associated with this stereotype of speech; additionally, it becomes clear that social indexicality is not applied to one feature alone but highly context-driven.

#### 8.2.4 Resume

---

Summing up the social and linguistic characteristics of the SAW, their social assimilation makes quite a difference for the orientation inside their social network and correspondingly for the speech. My data confirm that social assimilation is coupled with cognitive assimilation and thus linguistic accommodation. Roughly said, the SAW accommodate towards features that are associated with *Português Popular*, because their contacts are mostly people who live in the socially unprivileged areas of São Paulo and are often born in the northeast of Brazil. The features employed for accommodation are mainly salient phonology and vocabulary like discourse markers and regular expressions. The generalization of the absence of verbal and number agreement further points in this direction, but these features could also be influenced by morphological simplification and CP-transfer.

---

<sup>216</sup> I have already pointed out that by using the method of an online-survey poorer people without internet access were systematically excluded. A higher grade of representativeness exactly for this group might also reveal positive attitudes with regard to their own speech.

I identified some of the SAW who accommodated more obviously than others by features from all linguistic areas just enumerated. The test of language regard reveals not only that the Bolivian accommodating by these socio-indexical features was taken for a Brazilian, but also the social stigmatization of this style in São Paulo. Thus, the accommodation strategy towards PP by socio-indexical features as employed by the SAW is obviously successful in the sense of “I want to talk like Brazilians; I want to be perceived as a *Paulistano*”.

However, the group is more heterogeneous than I anticipated: Some of them are also structurally assimilated showing hence similarities with the academics which will be discussed below, while others differ little from NAW. The target variety of most SAW can still be characterized as PP, even if the individual styles might be quite different, depending for example on gender and situation: For a woman speaking to another woman at her home with her little daughter coming into the room, it would simply be strange to use socially tabooed Common Expressions as they were used by the men above. Understanding PP as the pole of a continuum, the speech of the young men has to be regarded as rather close to the pole, i.e. displaying the highest degree of informality, orality and lack of education, while the others’ speech is more distant from this pole, but still shows the same tendencies.

### 8.3 The structurally assimilated academics (StAA)

---

*Sin documento nó são ninguém. Eles têm que fazer os documentos para se tornar um cidadão (A21 (62) m. 1964 E. Ce: 3.56).<sup>217</sup>*

The third group of Bolivians differs strongly from the others because they come to Brazil legally, frequent Brazilian educational institutions, and they have an income that gives them the opportunity to be part of São Paulo’s middle class. Thus, from the very outset, the structural assimilation that A21 considers as very important is already given to a high degree.

#### 8.3.1 Social network

---

The social network of the StAA is also transnational. Because of their better financial means they have the possibility to travel regularly between the two countries. A23 even has pharmacies in La Paz (in addition to a catering service in São Paulo) and her son studies at a Bolivian university after

---

<sup>217</sup> Paraphrasing in English: “Without document (i.e. visa that legalize the status of the Bolivians in Brazil) they are nobody. They have to do the documents to become a *cidadão*, an actual citizen” (own translation, sn).

having gone to school in Brazil. The others also confirm that they regularly visit their parents in Bolivia, but even the younger ones want to stay in Brazil, although not necessarily in São Paulo:

*Pretendo [ficar aqui] sim mas não aqui no São Paulo ((0,3s)) eu éh quero mudar para o [...] Sul [...] eu gosto do Su<sup>18</sup> (A18 (30) m. 2004 P. Ce: 12.48).*

Additionally, A18 tells about his journey to *Brasília*, A20 has been to *Rio Grande do Sul*, etc. Thus, the StAA are a lot more mobile in Brazil than the other groups. Most of them even have cars:

*Graças a deus a gente con a que yo moro tenha auto. [...] Entonces siempre salimos en auto siempre vamos pra outro lugar shoppings baladas tudo<sup>19</sup> (A22 (27) f. 2008 P. ZS: 18.25).*

This already indicates that the StAA are also assimilated with regard to mobility and that they participate in urban life like middle-class *Paulistanos*.

Nonetheless, the contact with Bolivians is strong because almost all of them are married to Bolivians, only the wives of A19 and A21 are Brazilian. The younger physicians who have come to São Paulo in the last two years are single but live with Bolivian kinship. Many of them have children who grow up as Brazilians. Interestingly, in comparison to the opinion of B4 that his child has to learn Castilian first, A18 states that he wants his daughter to first learn Portuguese (“quero que ela aprenda primeiro bem o o por...português”, A18 (30) m. 2004 P. Ce: 7.31, own translation, sn). All except one have someone in the family they talk Castilian to, either husband, siblings or further kinship.

*Eu falava espanhol com o meu esposo e o primeiro filho falava nem espanhol nem português. [...] Eu fui na psicóloga de tratamento e ela falou assim: Tem que falar só português com ele (B8 (48) f. 1989 P. Os: 14.32).*

Apart from the family, their further environment is clearly not Bolivian, not one of them lives in *Bom Retiro*, *Brás* or other Bolivian shaped neighborhoods, on the contrary:

*Lá em casa todos os meus amigos são Brasileiros, os vizinhos (A16 (32) f. 2004 P. ZL: 3.31).*

*Donde yo moro allá en Santo Amaro não tem muito boliviano entonces eu, a minha família, [...] yo acho que es en essa setor os únicos bolivianos que moram lá (A22 (27) f. 2008 P. ZS: 19.46).*

At work, almost all of them are in contact with the Bolivian community. A19 ((42) m. 1988 P. Ce) is the only one who claims that he has no further contact with Bolivians apart from random visits to *Praça Kantuta*.

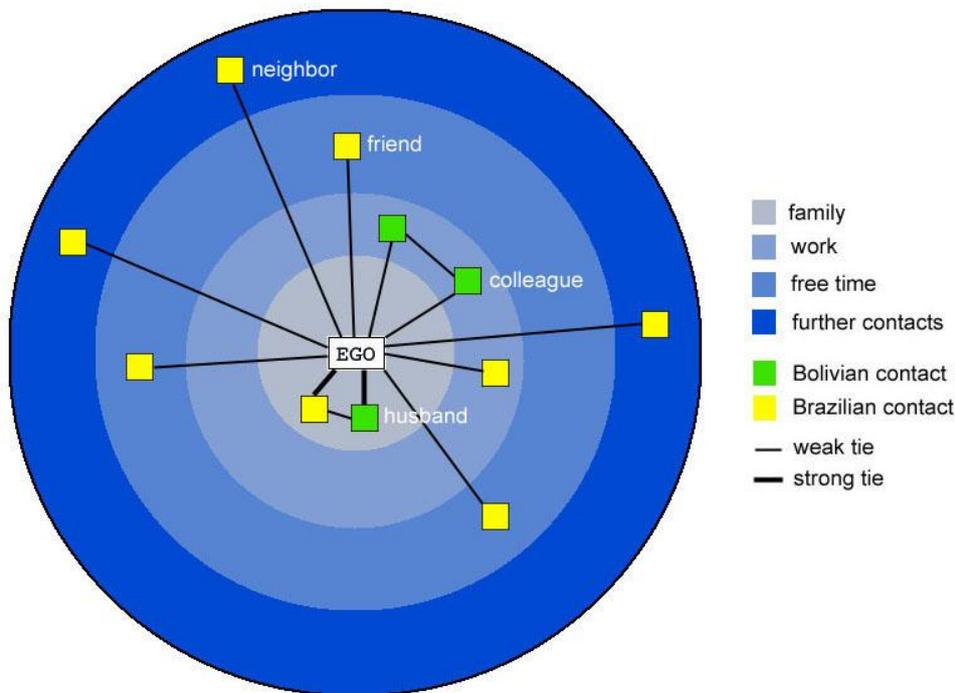
---

<sup>218</sup> Paraphrasing in English: “I want to stay here, yes, but not here in São Paulo. I want to move to the South, I like the South” (of Brazil, own translation, sn).

<sup>219</sup> Paraphrasing in English: “Fortunately, the people I live with have a car. [...] This way, we always go out by car, to shopping centers, to parties, everything (own translation, sn).

Leisure time is spent either with the family or with friends who mostly are Brazilians; they form their *turma* as A20 ((27) f. 2006 P. Gu) calls it. Thus, the StAA are not only structurally but also socially assimilated as figure 8.3 shows:

Figure 8.3 Prototypical network of one of the StAA (EGO-centered).



In contrast to the other networks, a clear division of the different CoP can be seen which also constitute different communicative areas: family, work, leisure time, and further contacts like (not better known) neighbours or the sales assistant at the supermarket. The transnational network is loose-knit – the connections between the ego and its contacts are simplex. In contrast to the other groups, here, a colleague is only a colleague and not also a family member. In addition, there are many Brazilian contacts in their network: Bolivian contacts are only those at work and the family member mentioned above.

Correspondingly, the communities of practice where they actually speak Castilian are restricted to work and to the family member, although even here, Portuguese plays a strong role because of the education of the children as Brazilians. In analogy, their attitude towards Portuguese learning is very positive; obviously they think that cognitive assimilation is a precondition for social assimilation as the following quotations show:

*Estamos num país onde se fala português e temos que aprender a língua [...] eu acho que precisa sim [aprender português] porque assim o pessoal se relaciona mais e está mais integrado à sociedade<sup>220</sup> (A20 (27) f. 2006 P. Gu: 4.50).*

*Eu vi a discriminação com os bolivianos por fato de que eles não sabem falar português<sup>221</sup> (A16 (32) f. 2004 P. ZL: 3.55).*

*Problema é comunicação. Eu acho que é o mais importante<sup>222</sup> (A18 (30) m. 2004 P. Ce: 8.57).*

Regarding their own process of language learning, some of them did a language course – A16 ((32) f. 2004 P. ZL) even before planning to go to Brazil – while others watch television or simply talk to people:

*Mas ao princípio o idioma era um pouco complicado mas [aprendí português] escutando falando. [...] O que eu gosto de fazer é falar com as pessoas. Às vezes em um ponto de ônibus eu falo com o pessoal [...] para escutar e falar<sup>223</sup> (A17 (31) m. 2007 P. Ce: 2.06).*

Even more importantly, they also have access to educational institutions. Even if they do not mention it explicitly, they all frequented courses at Brazilian universities, although none of them took an actual Portuguese course in Brazil.

### 8.3.2 Linguistic characterization

---

The StAA have loose-knit networks, positive attitudes towards Brazil and Portuguese, contacts with Brazilian middle-class, and access to educational institutions. Thus, I hypothesize that L1-transfer should be rather limited, because of the orientation towards speakers of Brazilian Portuguese in their network. The StAA rather accommodate towards PC because the contacts in their network are rather middle-class and they learned PC as standard at the universities. This and their professions imply that they are familiar with the written language. Therefore, the process of morphological simplification associated with orality should be reduced.

#### 8.3.2.1 Accommodation

---

With regard to the phonological features attributed to accommodation towards Brazilian Portuguese, diphthongization, aspiration of the vibrant, and palatization, most of the StAA realize them highly

---

<sup>220</sup> Paraphrasing in English: “We are in a country where you speak Portuguese; and we have to learn the language. [...] I think learning Portuguese is necessary, yes, because this way, people connect more and are more integrated in society” (own translation, sn).

<sup>221</sup> Paraphrasing in English: “I saw the discrimination of the Bolivians, because they do not know how to speak Portuguese” (own translation, sn).

<sup>222</sup> Paraphrasing in English: “The problem is communication. I think that it is most important” (own translation, sn).

<sup>223</sup> Paraphrasing in English: “But in the beginning, it was a little difficult, but [I learned Portuguese] listening and talking. [...] What I like is talking to people. Sometimes, at a bus-stop, I talk to the people [...] to listen and to talk” (own translation, sn).

regularly, for the latter even hypercorrection by A20 ((27) f. 2006 P. Gu). The men who have been in São Paulo the longest, though, show few occurrences of palatization and aspiration. Similarly, *çê* and *nós* are used only by the elder male StAA, while the others use the pronouns *você* and *a gente*, the latter only with the semantics of “we”; the generalized meaning is expressed by vocabulary like *o pessoal*.

This division according to gender could indicate that *çê* is too colloquial for the women in the context of an interview. Interestingly, two of the four women did not address me directly at all, the third only once. The use of *nós*, the non-palatization and non-aspiration by the eldest speaker who has spent most time in São Paulo, though, could signify actual non-accommodation to Portuguese while maintaining his identity as an immigrant. I observed similar linguistic behavior for *Paulistanos* with Italian ancestries, who confirm their identity as *sons of Italians* by avoiding palatization and aspiration (cf. Niehoff 2006). Thus, these features might present a kind of a hidden prestige form: This is possible because no one will evaluate him as a poor immigrant anyway due to his speech conforming highly to the grammatical standard PC, e.g. *nós* with corresponding verbal agreement. Thus, he uses the social indexicality of the phonological features to construct his identity as a successful immigrant. This interpretation is supported by the fact that the informant proudly tells me about the merits of Bolivians in Brazil and also his marriage to an Italian woman (cf. A21 (62) m. 1964 E. Ce: 15.13).

In addition, the frequent use of the *r-caipira* by the young academics must be mentioned. It actually seems that this feature has undergone a process towards more prestige that is particularly fostered by young women (cf. Beline Mendes 2009).<sup>224</sup>

In addition, they also use *né* as discourse marker as well as regular expressions of religious origin, even the one who has only come to São Paulo three months before the interview.

Thus, the empirical data confirm that most StAA use phonological features and vocabulary by which they accommodate to Brazilian Portuguese but not explicitly to PP styles like the SAW before. However, some of them choose to non-accommodate with regard to phonology to confirm their identity as immigrant.

#### 8.3.2.2 L1-transfer

---

<sup>224</sup> Interestingly, the only SAW realizing the *r-caipira* is one of those who accommodate most to PP, using highly socio-indexical markers. This could also affirm my interpretation of a certain prestige of this feature.

Regarding L1-transfer I would expect that the academics show few CP-features because they are not bilingual<sup>225</sup> and ethnically they do not belong to the indigenous of Bolivia. On the contrary, they always speak of the other Bolivian immigrants with their indigenous faces as clearly another social group. Indeed, there is no occurrence of a word order differing from SVO – therefore, this particular might actually be due to direct CP-transfer. The other features, however, that could also be attributed to CP-transfer can all be found for the StAA, too, as will be elaborated in the next section.

The StAA are quite reflective about their own linguistic behavior, in particular their Portuguese. The typological closeness between Castilian and Portuguese is not seen as an advantage:

*[...] A língua só aqui que aprendí. Ficou um pouco difícil, por quê? Porque é muito parecido ao espanhol então no final quando você começa você não sabe se está usando o término corretamente o se já está misturando com a língua materna, né?<sup>226</sup> (A18 (30) m. 2004 P. Ce: 1.13).*

The process that A18 describes here is obviously *congruent lexicalization* as described by Muysken (2000), where it is unclear for the speaker which language he is using while he speaks. Since most of the speakers show very little L1-transfer, four speakers must be considered in particular: On the one hand, A21, the eldest speaker mentioned above, and A20, the young physician with the r-*caipira*, are the only ones who show no signs of code-mixing at all. The transfer that occurs in their speech is restricted to the voiceless realization of the sibilants, the denasalized nasal diphthongs, and the variation between open and closed vowels, all features attributable to SLM. This clearly tags them as the speakers whose accommodation to PC has progressed furthest.

On the other hand, there are two women for whom alternation and congruent lexicalization is quite common: A23 ((45) f. 1990 Ph ZO) and A22 ((27) f. 2008 P. ZS). For the latter, this is clearly due to her short stay of only three months in São Paulo. She is the extreme case, included to test for occurrences owing to little proficiency. Interestingly, even she accommodates phonologically and uses the socio-indexical discourse marker and pronouns, lexicalizing hence her mostly Castilian grammar. A23 is the woman who is also professionally active in São Paulo. Out of the academics her network is clearly the most transnational one with most influence from the Bolivian side. Thus, the higher degree of code-mixing could also be seen as an emphasis of her transnational personality. As she

---

<sup>225</sup> Bilingual here refers only to the situation in Bolivia, i.e. growing up with one of the indigenous languages in addition to Castilian.

<sup>226</sup> Paraphrasing in English: “I only learned the language here. It was a little difficult, why? Because it is very similar to Spanish; so, in the end, when you start you do not know if you are using the term correctly or if you are already mixing with your mother tongue, right?” (own translation, sn).

clearly belongs to Bolivia's as well as Brazil's upper-class and was never discriminated, it does her no harm to strengthen her Bolivian identity in the middle of a Bolivian restaurant in São Paulo.

### 8.3.2.3 Morphological simplification

---

As was illustrated in 8.1.2.2 and 8.1.2.3, the morphological simplification with regard to prepositions actually concerns all speakers. With regard to verbal agreement it concerns all StAA but A20, the female physician just mentioned. The absence of gender and number agreement are also found for almost all StAA but A21, the eldest speaker in the case of number agreement, and another young female physician in the case of gender agreement. This divulgation even among academics, who speak no indigenous language, again constitutes an argument against the attempt to explain these phenomena by one-to-one transfer of CP structure.

At the same time, this means that despite their access to educational institutions and written Portuguese, the morphology particularly of the younger men would be rather associated with *Português Popular*. Since they indicate that they have learned Portuguese on the streets and by conversation, it underlines my understanding of PP as rather oral and informal variety that is also spoken by middle-class *Paulistanos*. Hence, the Bolivians perceive exactly those features as part of oral Brazilian Portuguese without realizing the social stereotype associated with them.

In addition, it must be considered that the three linguistic processes – accommodation to PP, transfer from PC, and the universal process of simplification – point in the same direction. These forces together are simply stronger than any language norm, particularly in the case of natural language learning.

### 8.3.3 Evaluation by *Paulistanos*

---

The group of StAA is represented by S III, whose speech is marked by some L1-transfer, particularly regarding phonology, but also by the highest degree of features associated with PC (cf. table 5.1), in particular with regard to morphology. Accordingly, he was estimated the wealthiest, most hard-working and knowing speaker of all. This fact clearly shows the social indexicality of those linguistic features, particularly since many people mentioned that he spoke like a professor.

At the same time he was evaluated as a foreigner, first and foremost due to cases of the insertion of Castilian vocabulary and phonological features like for example the realization of /v/ as [b].<sup>227</sup> Inter-

---

<sup>227</sup> This linguistic feature did not occur in the corpus and was therefore left out.

estingly, despite the identification of those linguistic features, almost no one thought that he came from Bolivia, despite the unambiguous Castilian-transfer. Many informants estimated that he came from a rather wealthy country in Europe. Correspondingly, they thought that he lived in one of the socially privileged areas of São Paulo, like *ZS* or *ZO*, confirming hence my theses of social and linguistic separation of São Paulo, where PC is clearly correlated with the higher socioeconomic classes.

There was no significant correlation between the personal evaluation and the status or educational level of the informants; although I had expected that the more educated informants would evaluate him more positively than the others. However, this might be due to a too strong perception of him being foreign to build a common identity. It revealed, though, that the better the social status of the speaker is, the more L1-transfer the speech can display without social consequences for the speaker. This confirms my explanation of the linguistic behavior of A21 and A23 as given above.

#### 8.3.4 Resume

---

The StAA are strongly assimilated in São Paulo, not only structurally but also socially and cognitively. Their transnational network is by far the loosest-knit as compared to the NAW and SAW; they participate most in all aspects of urban life and live the life of a middle/upper class *Paulistano*. Despite the huge social differences between the groups, the linguistic differences seem to be rather gradual, particularly if compared with some of the *donas* of the SAW. SLM, L1-transfer and morphological simplification can be observed in the speech of the StAA as well, although there are some speakers who show neither code-mixing nor any absence of number agreement. The far spread absence of gender agreement and particularly the omission and confusion of prepositions even among the StAA owes to various factors that all operate in the same direction. The fact that they show no explicit accommodation towards PC is easily explained by the fluent limits between PC and PP and the oral medium of the interview. As PC constitutes one pole of the continuum representing the highest degree of communicative distance associated with the written medium, a conversation in this style would be rather inappropriate. The fact that various informants thought that S III was reading confirms this view. Nonetheless, the two speakers that objectively speak the most elevated Portuguese are found in this group.

#### 8.4 Summing up

---

In this chapter the heterogeneity of the Bolivian's community was shown. Its different groups not only differ greatly by their profession and educational background but subsequently by their grade of assimilation and their linguistic characteristics. The distinction between workers in the textile industry and academics resulted in the formation of two groups because the profession determines greatly the social networks of the Bolivians in a segregated city like São Paulo: As long as the Bolivians work as seamstresses in the textile industry, they belong to a low socioeconomic class, have little geographic mobility and rather few opportunities to make contact with Brazilians. Thus, their network is strongly conditioned by the environment. The personal attitude of whether to make friends with, or at least talk to *Paulistanos* is nonetheless the decisive factor for an almost completely Bolivian versus an actual transnational network. The Bolivian academics, though, have all the possibilities to form their network in São Paulo's middle class, moving inside the city, attending different communities of practice and even educational institutions that transport PC.

Investigating the social networks of the various groups, they differed greatly in terms of their assimilation to São Paulo's society, taking into account the four dimensions of assimilation established by Esser (2001): cognitive, social, structural, and identificational assimilation. Because of the institutional advantages of the academics with regard to e.g. education and income, they were called structurally assimilated academics.

Among the workers, though, the not assimilated were distinguished from the assimilated because the linguistic differences between the groups revealed clear disparities. The network of the socially assimilated workers contains at least one strong-tie contact with a Brazilian speaking person.

Considering all three groups, the NAW and the StAA constitute two opposite poles regarding their assimilation to São Paulo. While the NAW live their lives isolated from the Portuguese-speaking society, almost without being influenced by the characteristics of their new environment, the StAA are highly mobile, they take part in associations, in public institutions like university or school and also have private contacts to Brazilians. The SAW are located between these two poles: some of them lean more to the NAW, others more to the StAA.

Obviously, these social differences are also reflected by the linguistic characterization:

- The lack of assimilation of the NAW results in the non-accommodation towards the Brazilian society and their speech, making only as few adjustments as necessary to be able to communicate with the investigator. The dominant linguistic process for this group is thus

clearly L1-transfer, particularly in terms of the insertion of Portuguese vocabulary in Castilian as matrix language. Their non-assimilation and non-accommodation is little tolerated by the *Paulistanos*, but by definition, the NAW do not care about Brazilian society.

- By their social assimilation, the SAW are clearly oriented towards the Brazilian side of their transnational networks. Correspondingly, the dominant linguistic process is accommodation towards the speech of the Brazilian contacts they have in the socially unprivileged areas of São Paulo. As those Brazilian contacts speak styles that are rather associated with PP, this is also the target variety of the SAW. The employment of features like certain discourse markers and Common Expressions, indicated to the style of youngsters living in the periphery, as the examples of three young men underlined. The success of this type of accommodation was confirmed by the evaluation of the *Paulistanos*, at the same time revealing a strong social stigmatization of this style of PP.
- The assimilation of the StAA can be seen for all dimensions. Correspondingly, the speakers accommodate towards Portuguese and those whose speech is nearest to PC are found in this group. Among them, there is particularly one whose social status is high enough so that he non-accommodates by phonological markers to confirm his identity as a successful immigrant. L1-transfer is very low in this group. Only one of them emphasizes her transnational network by congruent lexicalization and alternation. The perception test confirmed that L1-transfer is accepted if particularly morphology shows PC conformity.

The linguistic factors that were relevant for all groups were mainly phonological features explainable by SLM and also morphological simplification regarding verbal, number, and gender agreement and the collocation of prepositions. The frequency of these processes for almost all speakers, even the StAA, are quite surprising and only explicable if it is assumed that various factors like accommodation, L1-transfer and simplification operate together in a situation where language norms are little prominent.

This means that my hypotheses about the speech of Bolivian immigrants in São Paulo are correct: Accommodation theory based on assimilation to social networks and by social indexicality has high explanatory power for language contact in the city. In addition, L1-transfer in terms of congruent

lexicalization as well as morphological simplification are further linguistic processes which strongly influence the speech of the immigrants.

## 9. Conclusion and outlook

---

In this chapter the study is concluded by shortly answering the questions that had been asked in the first chapter: Who speaks Portuguese? In which way? Why? Even though not every aspect could be answered extensively, the theoretical approach explaining the Portuguese of the Bolivian immigrants as product of the linguistic processes – L1-transfer, morphological simplification and particularly accommodation – was confirmed by the empirical data.

The migrants linguistically assimilate and accommodate to the receiving society as their audience, and this, in particular, must be considered for studies focusing on such complex social phenomena like migration. Even though the perception test only hinted at the potential of this method, it is an important step to include the too long neglected aspect of the hearer and thus verify sociolinguistic hypotheses about structural and especially stylistic variation. I therefore postulate that for the investigation of social indexicality, clearly one of the central concepts of today's sociolinguistics, scholars have to resort to mixed methods research including both speaker and hearer.

---

In the first chapter of this study various questions were asked that arise when putting together migration theory, sociolinguistic as well as language contact theories, and the specific qualities of São Paulo as the social environment. They were subsumed by the following questions: Which of the immigrants speak Portuguese, in which way and why?

The first question is already tricky, because, in order to define which of the immigrants speak Portuguese and who does not, I need an objective definition of where Castilian stops and Portuguese starts. This enterprise is actually impossible since the two languages are so narrowly related that the congruent use of both, e.g. congruent lexicalization as suggested by Muysken (2000), is one of the most important linguistic processes for this language contact.

While the Bolivian immigrants themselves regard this sort of interlanguage as sufficiently Portuguese to agree to be interviewed by a Portuguese-speaking researcher, the *Paulistanos* as the receiving society do not accept this interlanguage as Portuguese. Those immigrants who show no signs of assimilation, come to work in the textile industry, mostly because of a social network that already contains contact persons working there and that reaches into São Paulo. They maintain a transnational network that does not contain any Brazilian strong-tie contact, but strong multiplex ties with other members of the Bolivian community. Many of them could be characterized as transmigrants, as they stay some time in São Paulo and then return to Bolivia, to then come back again. Correspondingly, even those who have been in Brazil for 20 years and have this kind of transnational network speak very little Portuguese. The study also shows that only one strong-tie contact in the network is enough for the speaker to become socially and cognitively assimilated and hence will learn Portuguese.

Therefore, the definition of the cognitive dimension as precondition for the further dimensions of assimilation, as hypothesized by one of the Bolivian academics, cannot be confirmed:

*A principal dificuldade é o idioma né eu acho que: por causa disso que os bolivianos não se integram à sociedade [...] brasileira (A20 (27) f. 2006 P. Gu: 6.44).*

The contrary is the case: social assimilation preconditions cognitive assimilation. Because of the typological relatedness of the two languages, the Bolivians quickly have sufficient linguistic means to make contact with Brazilians. However, if the personal environment presents too many obstacles for this contact, i.e. by illegality or fear of being discriminated and robbed, this first contact does not happen and any further assimilation is also impeded. It was shown, though, that many Bolivians manage to make contact and assimilate socially, even if the social conditions of their environment are equally hostile as those of the other workers in the textile industry.

Thus, positive attitude and individual effort play an important role for the quality of the social network. Additionally, categories like educational level and particularly socio-economic class are also highly relevant for the networks: The academics, for example, have simply more opportunity to communicate with Brazilians, because they work outside the textile industry, belong to São Paulo's middle class and live in completely different conditions.

Because of São Paulo's size and heterogeneity, varying lifestyles and networks are possible. These factors allow the Bolivian community to develop a highly institutionalized Bolivian environment inside the city, where people can continue their transnational networks, even without actually going back to Bolivia. Additionally, São Paulo's heterogeneous society is highly segregated with regard to socio-economic classes, geographic distribution – and even linguistically as the questionnaires have shown. Consequently, the Bolivians assimilate only to one part of this differentiated society and correspondingly accommodate linguistically to the variety or style that is socially most important for the single speaker, i.e. the variety that is spoken where they live, by the contacts of their social networks. Naturally, this leads to a high degree of variation in the Portuguese of different Bolivian immigrants. Looking closer at the way of how Bolivians speak Portuguese and how they accommodate, it was found that accommodation functions mainly by socio-indexical features, i.e. linguistic features that are highly salient and indicative of the social group the immigrant wants to belong to or seeks approval of. My data confirms that social indexicality is possible for all linguistic areas; vocabulary like discourse markers and Common Expressions and phonology are most prevalent. Asking the *Paulistanos* for their opinion, it seems that the poorer and socially unprivileged Bolivians experience

more pressure to accommodate to their Brazilian environment than the wealthier ones. This also means that, although their preconditions are clearly worse, the social pressure on the workers to assimilate in every dimension is higher than the pressure on the academics. Thus, the estimation of B2 seems to be just right for the unprivileged immigrants who come to São Paulo to find their luck there:

*Quando você chega aqui em São Paulo você não sabe falar português você é meio isolado [...] você fica no zero até você aprender. Se você não aprender falar direitinho você não vai se dar bem com São Paulo. Agora se você falar bem você se dá bem com São Paulo*<sup>228</sup>  
(B2 (30) m. 2004 S. ZL: 13.58).

Altogether, my theoretical approach explaining linguistic assimilation by the process of accommodation and non-accommodation based on socio-indexical features and translingual social networks seems applicable to any migration situation. Interestingly, Bommers (2005) found for sociological migration studies that the basic processes described in the literature are still valid, but the social categories used could just not explain sociological reality. By analogy, many of the theoretical approaches presented in Chapter 3 are highly informative and intelligent, but their categories are mostly too broad and too imprecise to actually constitute a valid instrument for the linguistic explanation of situations of modern-day migration.

Network-oriented accommodation, though, constitutes a concept that actually clarifies the process that underlies the influence of social factors on the outcome of language contact, no matter if the languages are typologically related or not. Complemented by the process of L1-transfer and – in the case of natural SLA where normative language lessons are rather absent – by the universal tendency of simplification due to the orality of the contact situation, every linguistic behavior of the speakers could be explained, on the individual as well as on the macro-level. Additionally, these results were confirmed by the *Paulistanos* themselves.

This validates also the choice of my methodology: As assimilation, accommodation and social indexicality are two-sided and thus dependent on negotiations between migrant and receiving society, learner and contact person, speaker and audience, both sides have to be taken into account if reliable data about complex sociolinguistic relations, such as provoked by migration, is wanted.

---

<sup>228</sup> Paraphrasing in English: “When you come to São Paulo and cannot speak Portuguese, you are rather isolated. You stay at zero until you learn. If you do not learn to speak right you will not go along well with São Paulo. But, if you speak well you go along well with São Paulo” (own translation, sn).

The hearer's side has been neglected far too long by sociolinguistics. Similar to the concept of stylistic variation that Labov started in the 1960s and then reduced to mere attention to speech, he also used perception tests and therefore mixed methods for his research about language in NYC. Yet, he never integrated the hearer's perspective into the theoretical variationist model. On the contrary, the investigator – the audience at the moment of the interview – should be reduced to invisibility as not to interfere with the speaker's vernacular. It was ignored that the gender or age of the interviewer could strongly influence the characteristics of the speakers talk, although these factors most certainly led to major biases, for example in the case of research projects where interviewers of different gender, age, ethnicity etc. were involved to collect the data for one data base. Although I was unable to investigate this part, it became obvious that my own speaking style during the different interviews differed depending on social factors of the communicating people.

Thus, the concepts to integrate the hearer into sociolinguistic research must be elaborated together with the progress on stylistic variation. As Bell (1984 and 2009) started to show, audience design is most promising for the actual understanding of the social dimension in communicative practice. Correspondingly, the methodology to integrate the hearer's perspective must be expanded and refined. This study only offers limited representativeness because of the flaws of an online survey that biases the sample selection by the necessary tool of self selection. But nonetheless, I am convinced that mixed-methods research has the right amount of methodic advantages to equip sociolinguistics with a tool to grasp the current challenges that wait particularly in the cities.

Concluding this study, I turn my view back to São Paulo's history of immigration: My data revealed that despite the close relatedness between the two languages, Castilian and Portuguese, the effects of language contact in a natural situation are enormous, particularly regarding morphological simplification in the cases of the omission of prepositions and gender attribution, even for cases where the prescriptive norms of the two languages completely coincide. The fact that these phenomena are found for almost every speaker, irrespective of social status, of their education or their Portuguese proficiency might shed light on the development of the morphological system of current Brazilian Portuguese. Thus, I hope that my study motivates more scholars to investigate the current immigration to the melting pot São Paulo to finally understand its linguistic reality, history and maybe even future developments: Let's study the present to explain the future, the present and the past!

## 10. References

---

- Abaurre, M. B. M. (Ed.) (2002). *Gramática do português falado: Volume VIII: Novos estudos descritivos*. Campinas: Editora Unicamp.
- Abaurre, M. B. M. (2002). Nasalização no português do Brasil. In I. G. V. Koch (Ed.), *Gramática do Português Falado: Volume VI: Desenvolvimentos* (2nd ed., pp. 491–519). Campinas: Editora Unicamp.
- Abaurre, M. B. M. and Gozze Pagotto, E. (2002). Palatalização das oclusivas dentais no português do Brasil. In M. B. M. Abaurre (Ed.), *Gramática do português falado: Volume VIII: Novos estudos descritivos* (pp. 557–602). Campinas: Editora Unicamp.
- Akerberg, M. (2002). *Aprendizagem de uma língua próxima: Português para falantes de espanhol*. Stockholm: Stockholms univ. Institutionen för spanska portugisiska och latinamerikastudier.
- Akerberg, M. (2004). A importância da palavra escrita para a pronúncia. In A. R. Simões; A. M. Carvalho and L. Wiedemann (Eds.), *Português para falantes de espanhol: Artigos selecionados escritos em português e inglês* (pp. 115–124). Campinas, Spain: Pontes.
- Alves, U. S. (2011). *Imigrantes Bolivianos em São Paulo: A Praça Kantuat e o Futebol*. Tese de Doutorado, Universidade Estadual de Campinas, Campinas.
- Amaral, M. (22.2.2011). A panela de pressão que é o Metrô de SP. *Revista do Brasil*. Retrieved July 09, 2011, from [http://fenametro.org.br/site/index.php?option=com\\_content&Itemid=&task=view&id=557](http://fenametro.org.br/site/index.php?option=com_content&Itemid=&task=view&id=557).
- Auer, P. (1989). Natürlichkeit und Stil. In V. Hinnenkamp et al. (Eds.), *Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik* (pp. 27–59). Tübingen: Niemeyer.
- Auer, P., Arnhold, J. and Bueno-Aniola, C. (2005). Being a "colono" and being "daitsch" in Rio Grande do Sul: Language choice and linguistic heterogeneity as a resource for social categorization. *Caleidoscópio*, 3(3), 170–183.
- Auer, P., Barden, B. and Grosskopf, B. (1998). Subjective and objective parameters determining 'salience' in long-term dialect accommodation. *Journal of Sociolinguistics*, 2(2), 163–187.
- Badilho, M. (Ed.) (2007). 50 anos Imigração Boliviana em São Paulo 1957-2007. *CPM além fronteiras, Edição Histórica*. São Paulo.
- Bahé, M. and Lucena, P. (2011) *Com vocês, "Por uma vida melhor", o livro mais popular do Brasil*. Retrieved July 09, 2011, from <http://acertodecontas.blog.br/economia/com-vocs-por-uma-vida-melhor-o-livro-mais-popular-do-brasil/>.
- Baptista, D. M. (1998). *Nas terras do "Deus Dará": Nordestinos e suas redes sociais em São Paulo*. Tese de doutorado, PUC, São Paulo.
- Barros, C. J. (2000). *Senzalas Bolivianas*, from [www.tupinitango.hpg.ig.com.br/senzala.pdf](http://www.tupinitango.hpg.ig.com.br/senzala.pdf).
- Barros, M. (28.9.2008). "Capital cultural" é para poucos. *Folha de São Paulo especial*, 19–20. Retrieved July 10, 2011, from [www.nossasaopaulo.org.br/observatorio/biblioteca/DnaPaulistanoSPTtotal.pdf](http://www.nossasaopaulo.org.br/observatorio/biblioteca/DnaPaulistanoSPTtotal.pdf).
- Battisti, E., Dornelles Filho, A. A., Pires Lucas, J. I. and Bovo, N. M. P. (2007). Palatalização das oclusivas alveolares e a rede social dos informantes. *Revista Virtual de Estudos da Linguagem - ReVEL*, 5(9), 1–28.
- Bedinelli, T. (28.9.2008). Morador se sente seguro em bairro violento. *Folha de São Paulo especial*, 12–14. Retrieved July 10, 2011, from [www.nossasaopaulo.org.br/observatorio/biblioteca/DnaPaulistanoSPTtotal.pdf](http://www.nossasaopaulo.org.br/observatorio/biblioteca/DnaPaulistanoSPTtotal.pdf).
- Beebe, L. M. and Giles, H. (1984). Speech accommodation theories: A discussion in terms of second-language acquisition. *International Journal of the Sociology of Language*, 46, 5–32.

- Beline Mendes, R. (9.11.2006). *What is 'gay speech' in São Paulo, Brazil?* Paper presented at the IGALA 4. Valencia.
- Beline Mendes, R. (22.10.2009). *Who sounds /r/-ful? The pronunciation of /r/ in São Paulo Portuguese.* New ways of Analyzing variation. Ottawa.
- Bell, A. (1984). Language style as audience design. *Language in Society*, 13, 145–204.
- Bell, A. (1991). Accommodation in the mass media. In H. Giles, J. Coupland and N. Coupland (Eds.), *Studies in emotion and social interaction: Contexts of accommodation. Developments in applied sociolinguistics* (pp. 69–102). Cambridge: Cambridge Univ. Press [et al.].
- Bell, A. (2001). Back in style: reworking audience design. In P. Eckert and J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 139–169). Cambridge [et al.]: Cambridge Univ. Press.
- Bell, A. (2009). Language Style as Audience Design. In N. Coupland (Ed.), *The new sociolinguistics reader* (pp. 265–275). Basingstoke: Palgrave Macmillan.
- Benites, A. (30.4.2010). Mais violento, Brás tem até código antifurto. *Folha de São Paulo*, C4.
- Blaser, J. (1995). *Das Spanische in Südbrasilien: Die Zerstörung einer Legende durch mikrodialektologische Feldforschung.* Univ. Diss. Mainz, 1995. *Europäische Hochschulschriften Reihe 24, Ibero-romanische Sprachen und Literatur.* Frankfurt am Main, Berlin, Bern: Lang.
- Bógus, L. M. M. and Taschner, S. P. (Eds.) (2004). Como anda São Paulo. *Cadernos MetrÓpole*, (12-5). Rio de Janeiro.
- Bógus, L. M. M. and VÉras, M. P. B. (2000). A reorganização metropolitana de São Paulo: espaços sociais no contexto da globalização. *Cadernos MetrÓpole*, (3), 91–98, from [http://web.observatoriodasmetrosoles.net/download/cm\\_artigos/cm3\\_10.pdf](http://web.observatoriodasmetrosoles.net/download/cm_artigos/cm3_10.pdf).
- Bohn, O.-S. (1998). Wahrnehmung fremdsprachlicher Laute: Wo ist das Problem? In H. Wegener (Ed.), *Tübinger Beiträge zur Linguistik: Series A, Language development: Vol. 24. Eine zweite Sprache lernen. Empirische Untersuchungen zum Zweitspracherwerb* (pp. 1–20). Tübingen: Narr.
- Bollée, A. and Neumann-Holzschuh, I. (2008). *Spanische Sprachgeschichte* (1st edition, [4. print]). *Uni-WissenSpanisch.* Stuttgart: Klett.
- Boller, F. (2002). Sonorización en castellano vs. ensordecimiento en portugués en situaciones de contacto luso-español. In N. Díaz, R. Ludwig and S. Pfänder (Eds.), *Lengua y sociedad en el mundo hispánico: Vol. 9. La Romania americana. Procesos lingüísticos en situaciones de contacto* (pp. 197–253). Frankfurt am Main: Iberoamericana; Vervuert.
- Bommes, M. (2005). Transnationalism or Assimilation? *Journal of Social Science Education*, 4(1), 15-30.
- Bortoni-Ricardo, S. M. (1985). *The urbanization of rural dialect speakers: A sociolinguistic study in Brazil.* London New York, Cambridge: Cambridge Univ. Pr.
- Bortoni-Ricardo, S. M. (1998). A análise do português brasileiro em três continua: o continuum rural-urbano, o continuum de oralidade-letramento e o continuum de monitoração estilística. In S. Große and K. Zimmermann (Eds.), *Biblioteca luso-brasileira: Vol. 6. "Substandard" e mudança no português do Brasil. [colóquio internacional ..., Berlin outubro de 1997]* (pp. 101–118). Frankfurt am Main: TFM - Teo Ferrer de Mesquita.
- Brazilian government (2010). *Geografia*, from <http://www.brasil.gov.br/sobre/geografia/populacao/numeros-gerais>.
- Briggs, C. L. (2005). Sociolinguistic Interviews/Soziolinguistisches Interview. In U. Ammon, N. Dittmar and K. J. Mattheier (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft: Vol. 2. Sociolinguistics.*

- An international handbook of the science of language and society = Soziolinguistik* (2nd ed., pp. 1052–1062). Berlin: de Gruyter.
- Busse, W. (1999). Die Sprache(n) der Sepharden: Ladino. In N. Rehrmann and A. Koechert (Eds.), *Romania Judaica: Vol. 3. Spanien und die Sepharden. Geschichte, Kultur, Literatur* (pp. 133–143). Tübingen: Niemeyer.
- Cacciamali, M. C. and Gomes Azevedo, F. A. de (2006). Entre o tráfico humano e a opção da mobilidade social: A situação dos imigrantes bolivianos na cidade de São Paulo. *Cadernos PROLAM/USP*, 5(1), 129–143. Retrieved July 05, 2011, from [http://econ.fea.usp.br/cacciamali/ENTRE\\_TRAFICO\\_HUMANO\\_OPCAO\\_MOBILIDADE.pdf](http://econ.fea.usp.br/cacciamali/ENTRE_TRAFICO_HUMANO_OPCAO_MOBILIDADE.pdf).
- Caldeira, T. (2000). *Cidade de muros: crime, segregação e cidadania em São Paulo*. São Paulo: Edusp/Editora 34.
- Callou, D. (Ed.) (1992). *A Linguagem Falada Culta na Cidade do Rio de Janeiro: Materiais para o seu estudo*. Rio de Janeiro: UFRJ/CAPES.
- Callou, D. and Lopes, C. R. de S. (Eds.) (1993). *A Linguagem Falada Culta na Cidade do Rio de Janeiro: Materiais para seu estudo*. Rio de Janeiro: UFRJ/CAPES.
- Camacho Kadluba, R. Myrian (2007). Partilhando sorrisos. In M. Badilho (Ed.), *Vol. Edição Histórica. 50 anos Imigração Boliviana em São Paulo 1957-2007* (p. 30). São Paulo.
- Campbell-Kibler, K. (2008). I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society*, 37(4), 637–659.
- Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change*, 21, 135–156.
- Campos Castro, P. de (2009). *Ainda há espaço para alfaiates no mundo do 'pronto para vestir'?* Monografia de conclusão de curso, Universidade Federal de Juiz de Fora, Juiz de Fora, from Universidade Federal de Juiz de Fora: <http://ufjf.br/posmoda/files/2010/09/Monografia-Alfaiataria2.pdf>.
- Cancino, R. (2008). El Mosaico de las lenguas de Bolivia: Las lenguas indígenas de Bolivia Obstáculo o herramienta en la creación de la nación de Bolivia? *Diálogos Latinoamericanos*, (013), 1–20.
- Cargile, A. C., Giles, H., Ryan, E. B. and Bradac, J. J. (1994). Language attitudes as a social process: A conceptual model and new directions. *Language and Communication*, 14(3), 211–236.
- Carita, I. (1998). Português e espanhol: "falsos amigos". In P. F. Pinto and N. Júdice (Eds.), *Para acabar de vez com Tordesilhas* (pp. 31–40). Lisboa: Ed. Colibri.
- Carvalho, A. M. (2002). Português para falantes de espanhol: Perspectivas de um campo de pesquisa. *Hispania*, (85), 598–608.
- Carvalho, A. M. (2003a). Rumo a uma definição do português uruguaio. *Revista internacional de lingüística iberoamericana*, 2, 125–150.
- Carvalho, A. M. (2003b). The sociolinguistic distribution of (lh) in Uruguayan Portuguese: a case of dialectal diffusion. In S. Montrul and F. Ordonez (Eds.), *Linguistic Theory and Language Development in Hispanic Languages. Papers from the 5th Hispanic Linguistics Symposium and the 4th Conference on the Acquisition of Spanish and Portuguese* (pp. 30–43). Somerville, MA: Cascadilla Press.
- Carvalho, A. M. (2003c). Variation and diffusion of Uruguayan Portuguese in a bilingual border town. In C. Cabeza, X. P. Rodriguez Yañez and A. Lorenzo Suárez (Eds.), *Comunidades e individuos bilingües. Actas do I Simposio Internacional sobre o Bilingüismo* (pp. 642–651). Vigo: Universidade.

- Carvalho, A. M. (2004). I speak like the guys on TV: Palatalization and the urbanization of Uruguayan Portuguese. *Language Variation and Change*, 16, 127-151.
- Carvalho, A. M. (2006). Nominal Number Marking in a Variety of Spanish in Contact with Portuguese. In T. L. Face and C. Klee (Eds.), *Selected Proceedings of the 8th Hispanic Linguistics Symposium* (pp. 154–166). Somerville, MA: Cascadilla Proceedings Project.
- Carvalho, A. M. (2008). Estudo de atitudes lingüísticas sobre o português de hispano-falantes: Até que ponto oportunhol é aceitável? In L. Wiedemann (Ed.), *Português para Falantes de Espanhol: Ensino e Aquisição* Artigos selecionados escritos em Português e Inglês. (pp. 157–173). Campinas: Pontes.
- Castilho, A. T. de (Ed.) (1989). *Português culto falado no Brasil*. Campinas: Editora Unicamp.
- Castilho, A. T. de (Ed.) (2002). *Gramática do Português Falado: Volume IV: Estudos descritivos* (2. ed. rev.). Campinas: Editora Unicamp.
- Castilho, A. T. de (2010). *Nova gramática do português brasileiro*. São Paulo: Contexto.
- Castilho, A. T. de and Preti, D. (Eds.) (1987). *A Linguagem Falada Culta na Cidade de São Paulo: Materiais para seu estudo*. São Paulo: TAQ/Fapesp.
- Clements, J. C. (2009). *The linguistic legacy of Spanish and Portuguese: Colonial expansion and language change*. Cambridge [et al.]: Cambridge Univ. Press.
- Coello Vila, C. (1996). Aproximación al castellano boliviano. *Hispanorama*, 72, 26–34.
- Coetsem, F. v. (1988). *Loan phonology and the two transfer types in language contact. Publications in language sciences: Vol. 27*. Dordrecht: Foris Publ.
- Coetsem, F. v. (2000). *A general and unified theory of the transmission process in language contact. Monographien zur Sprachwissenschaft: Vol. 19*. Heidelberg: Winter.
- Cortés Conde, R. (2003). La Crisis Argentina de 2001-2002. *Cuadernos de economía*, 40(121), 762–767.
- Coulmas, F. (2005). *Sociolinguistics: The study of speakers' choices*. Cambridge: Cambridge Univ. Press.
- Coupland, N. (2001). Language, situation and the relational self: theorizing dialect-style in sociolinguistics. In P. Eckert and J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 185–210). Cambridge [et al.]: Cambridge Univ. Press.
- Coupland, N. (2007). *Style: Language variation and identity. Key topics in sociolinguistics*. Cambridge [et al.]: Cambridge Univ. Press.
- Coupland, N. and Jaworski, A. (1997). Relevance, accommodation and conversation: Modeling the social dimension of communication. *Multilingua*, 16(2/3), 233–258.
- Cowie, N. (2009). Observation. In J. Heigham and R. A. Croker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 165–181). Houndmills Basingstoke Hampshire England New York: Palgrave Macmillan.
- Cowles, M. A. and Pires, A. M. (2008). The voice of the learner. In L. Wiedemann (Ed.), *Português para Falantes de Espanhol: Ensino e Aquisição. Artigos selecionados escritos em Português e Inglês*. (pp. 237–256). Campinas: Pontes.
- Crespim Pereira, D. (2007). *Variação e mudança no uso dos pronomes reflexivos no português popular da capital paulista: Uma abordagem funcionalista e cognitivista*, Universidade de São Paulo - Faculdade de Filosofia, Letras e Ciências Humanas, São Paulo.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches* (2. ed.,). Thousand Oaks, Calif.: Sage Publ.

- Creswell, J. W. (2008). *Educational research: Planning, Conducting and evaluating quantitative and qualitative approaches to research* (3rd ed.). Upper Saddle River: Merrill/Pearson Education.
- Creswell, J. W. et al. (2003). Advanced Mixed Methods Research Designs. In A. Tashakkori and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240). Thousand Oaks, Calif.: SAGE.
- Crevels, M. and Muysken, P. (2009). *Lenguas de Bolivia: Volume 1 Ámbito andino* (1. ed.). La Paz: Plural Editores.
- Cukor-Avila, P. and Bailey, G. (2001). The effects of the race of the interviewer on sociolinguistic fieldwork. *Journal of Sociolinguistics*, 5(2), 254–270.
- Cymbalista, R. and Xavier, I. Rolnik (2007). A comunidade boliviana em São Paulo: definindo padrões de territorialidade. *Cadernos Metrôpole*, (17), 119–133, from [http://web.observatoriodasmetropoles.net/download/cm\\_artigos/cm17\\_96.pdf](http://web.observatoriodasmetropoles.net/download/cm_artigos/cm17_96.pdf).
- Dantas, T. (9.4.2011). Metrô de SP é o mais lotado do mundo. *Estado de S. Paulo*, from [http://www.estadao.com.br/estadaodehoje/20110409/not\\_imp703954,0.php](http://www.estadao.com.br/estadaodehoje/20110409/not_imp703954,0.php).
- Dicionário da língua portuguesa (Ed. rev. e atualizada.) (2002). Porto: Porto Ed.
- Dores Nicolau, E. Maria de (2000). Uma abordagem sociolinguística da ausência de concordância verbal no português do Brasil. In E. Gärtner, C. Hundt and A. Schönberger (Eds.), *Biblioteca luso-brasileira: Vol. 15. Estudos de sociolinguística brasileira e portuguesa* (pp. 121–139). Frankfurt am Main: TFM Teo Ferrer de Mesquita.
- Dörnyei, Z. (2009). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies* ([Reprinted]). *Oxford applied linguistics*. Oxford: Oxford Univ. Press.
- Dubois, S. and Horvath, B. M. (1998). Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change*, 10 (3), 245 – 261.
- Duller, C. (2006). *Einführung in die Statistik mit EXCEL und SPSS: Ein anwendungsorientiertes Lehr- und Arbeitsbuch. Physica-Lehrbuch*. Heidelberg: Physica-Verlag Heidelberg.
- Eckert, P. (in review). *Three waves of Variation Study: The emergence of meaning in the study of variation*. Retrieved May 26, 2011, from <http://www.stanford.edu/~eckert/PDF/ThreeWavesofVariation.pdf>.
- Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. New York: Teachers College Press.
- Eckert, P. (2000). *Linguistic variation as social practice: The linguistic construction of identity in Belten High. Language in society*, 27. Malden Mass. [et al.]: Blackwell.
- Eckert, P. (2001). Style and social meaning. In P. Eckert and J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 119–126). Cambridge [et al.]: Cambridge Univ. Press.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4), 453–476.
- Eckert, P. and McConnell-Ginet, S. (1999). New generalizations and explanations in language and gender research. *Language in Society*, 28, 185–201.
- Eckman, F. R. (1977). Markedness and the Contrastive Analysis Hypothesis. *Language Learning*, 27, 315–333
- Eckman, F. R. (2008). Typological Markedness and Second Language Phonology. In J. G. Hansen Edwards, & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 95 - 115). Philadelphia: John Benjamins Publishing Co.

- Edwards, J. (1999). Refining our Understanding of Language Attitudes. *Journal of Language and Social Psychology*, 18(1), 101–110.
- Edwards, J. (2009). *Language and Identity: An introduction. Key topics in sociolinguistics*. Cambridge: Cambridge Univ. Press.
- Elizaincín, A. (1984). Comparación de dos situaciones de contacto de dialectos. *Cuadernos del sur*, 17, 89–100.
- Elizaincín, A. (1992). *Dialectos en contacto: español y portugués en España y América*. Montevideo: Arca Ed.
- Elizaincín, A. (1996). Contacto entre lenguas genéticamente emparentadas. *Signo y Seña*, 6, 415–434.
- Elizaincín, A. (2002). Diacronía del contacto español-portugués. In N. Díaz, R. Ludwig and S. Pfänder (Eds.), *Lengua y sociedad en el mundo hispánico: Vol. 9. La Romanía americana. Procesos lingüísticos en situaciones de contacto* (pp. 255–261). Frankfurt am Main: Iberoamericana; Vervuert.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Elmentaler, M., Gessinger, J. and Wirrer, J. (2010). Qualitative and quantitative Verfahren in der Ethnodi-alektologie am Beispiel von Salienz. In C. A. Anders, M. Hundt and A. Lasch (Eds.), *Linguistik - Impulse and Tendenzen: Vol. 38. Perceptual dialectology. Neue Wege der Dialektologie* (pp. 67–88). Berlin: de Gruyter.
- EMPLASA. *Região Metropolitana de São Paulo: Dados*. Retrieved June 05, 2010, from [www.emplasa.sp.gov.br/portalemplasa/infometropolitana/rmsp/rmsp\\_dados.asp](http://www.emplasa.sp.gov.br/portalemplasa/infometropolitana/rmsp/rmsp_dados.asp).
- Ernst, S. (2007). *São Paulo: Motor der brasilianischen Wirtschaft*. Retrieved June 11, 2011, from Bundeszentrale für politische Bildung: [http://www.bwb.de/themen/SDOOH0,0S%E3o\\_Paulo.html](http://www.bwb.de/themen/SDOOH0,0S%E3o_Paulo.html).
- Escobar, A. María (2000). *Contacto social y lingüístico. El español en contacto con el quechua en el Perú*. Lima: PUCP - Fondo Editorial.
- Esser, H. (1980). Aspekte der Wanderungssoziologie: Assimilation und Integration von Wanderern, ethnischen Gruppen und Minderheiten. Neuwied: Luchterhand.
- Esser, H. (2001). Kulturelle Pluralisierung und strukturelle Assimilation. Das Problem der ethnischen Schichtung. In *Schweizerische Zeitschrift für Politikwissenschaft*, Vol. 7(2) (pp. 97-108).
- Estadão* (10.6.2009): São Paulo tem recorde histórico de congestionamento. From [www.estadao.com.br/noticias/cidade,sao-paulo-tem-recorde-historico-de-congestionamento,385598,0.htm](http://www.estadao.com.br/noticias/cidade,sao-paulo-tem-recorde-historico-de-congestionamento,385598,0.htm).
- Faingold, E. D. (1996). *Child language, creolization and historical change: Spanish in contact with Portuguese. Tübinger Beiträge zur Linguistik: Series A, Language development: Vol. 18*. Tübingen: Narr.
- Faingold, E. D. (2003). *The development of grammar in Spanish and the Romance languages*. Basingstoke: Palgrave Macmillan.
- Ferrari, L. V. (1997). Variação e Cognição: o caso das preposições "em" e "ni" no Português do Brasil. *Revista da ANPOLL*, 3, 121–134.
- Ferreira Coelho, R. (August 2006). *É nós na fita! Duas variáveis lingüísticas numa vizinhança da periferia paulistana.: O pronome de primeira pessoa do plural e a marcação do plural no verbo*. Dissertação de Mestrado, Universidade de São Paulo - Faculdade de Filosofia, Letras e Ciências Humanas, São Paulo. Retrieved March 06, 2011, from: [www.fflch.usp.br/dl/pos/Teses/COELHOrafaelferreira.pdf](http://www.fflch.usp.br/dl/pos/Teses/COELHOrafaelferreira.pdf).
- Flege, J. E. (1995). Second Language Speech Learning Theory, Findings and Problems. In W. Strange (Ed.), *Speech perception and linguistic experience. Issues in cross-language research; [experiences of participants in a conference entitled "Workshop on Cross-Language Speech Perception", held at the University of South Florida in 1992]* (pp. 233–277). Baltimore: York Press.

- Fontana, A. and Frey, J. H. (2009). The Interview: From Neutral Stance to Political Involvement. In N. K. Denzin and Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 695–727). Thousand Oaks, Calif.: SAGE.
- Fontanella de Weinberg, M. B. (2000). Sistemas pronominales de tratamiento usado en el mundo hispánico. In I. Bosque and V. Demonte (Eds.), *Colección Nebrija y Bello: Gramática descriptiva de la lengua española* (1st ed., pp. 1399–1425). Madrid: Espasa Calpe.
- Fromm, S. and Baur, N. (2008). *Datenanalyse mit SPSS für Fortgeschrittene: Ein Arbeitsbuch* (2., überarbeitete und erweiterte Auflage.). Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden.
- Frank Kersch, D. (2006). *A construção relativa na língua falada: Enfoque na fronteira bilíngüe do Brasil com o Uruguai, comparado ao espanhol e ao português riopratense e europeu*. Univ., Diss.-Kiel, 2005. *Dialectologia pluridimensionalis Romanica: Vol. 6*. Kiel: Westensee-Verlag.
- Freire da Silva, C. (2008). *Trabalho Informal e Redes de Subcontratação: Dinâmicas Urbanas da Indústria de Confecções em São Paulo*. Dissertação de Mestrado, Universidade de São Paulo - Faculdade de Filosofia, Letras e Ciências Humanas, São Paulo. Retrieved July 05, 2011, from: [http://fflch.usp.br/sociologia/pos-graduacao/sites/trajetorias/txts/Carlos\\_Freire.pdf](http://fflch.usp.br/sociologia/pos-graduacao/sites/trajetorias/txts/Carlos_Freire.pdf).
- Freitas, S. M. (1999). *Falam os Imigrantes... Memória e Diversidade Cultural em São Paulo*. São Paulo.
- Friedrichs, J. (1995). *Stadtsoziologie*. Opladen: Leske + Budrich.
- Gal, S. (1979). *Language Shift: Social Determinants of Linguistic Change in Bilingual Austria*. New York et al.: Academic Press.
- Gallois, C., Ogay, T. and Giles, H. (2005). Communication Accomodation Theory. In W. B. Gudykunst (Ed.), *Theorizing about intercultural communication* (pp. 121–148). Thousand Oaks: Sage Publ.
- Gambhir, S. K. (1981). *The East Indian speech community in Guyana: A sociolinguistic study with special reference to koine formation*. Ph.D. dissertation, University of Pennsylvania.
- Garrido Meirelles, V. A. (2009). O português da fronteira Uruguai-Brasil. In A. M. Carvalho (Ed.), *Linguística luso-brasileira. Português em contato* (pp. 257–275). Madrid, Frankfurt am Main: Iberoamericana; Vervuert.
- Giannini, D. (05/12/2009). *Eleição entusiasmo bolivianos em São Paulo*, from: [www.noticias.r7.com/internacional/noticias/eleicao-entusiasma-bolivianos-em-sao-paulo-20091205.html](http://www.noticias.r7.com/internacional/noticias/eleicao-entusiasma-bolivianos-em-sao-paulo-20091205.html).
- Giles, H., Coupland, N. and Coupland, J. (1991). Accommodation theory: Communication, context and consequence. In H. Giles, J. Coupland and N. Coupland (Eds.), *Studies in emotion and social interaction. Contexts of accommodation. Developments in applied sociolinguistics* (pp. 1–68). Cambridge: Cambridge Univ. Press [et al.].
- Giles, H. et al. (1987). Speech accommodation theory: The first decade and beyond. In M. L. McLaughlin (Ed.), *Communication yearbook 10* (pp. 13–48). Newbury Park, CA: SAGE.
- Godenzi, J.-C. (2007). El español de América y el español de los Andes: universalización, vernacularización y emergencia. In M. Schrader-Kniffki, L. Morgenthaler García and K. Zimmermann (Eds.), *Lengua y sociedad en el mundo hispánico: Vol. 18. La Romania en interacción. Entre historia, contacto y política ; ensayos en homenaje a Klaus Zimmermann* (pp. 29–50). Frankfurt am Main: Vervuert [et al.].
- Gomes Azevedo, F. A. de (2005). *A presença de trabalho forçado na cidade de São Paulo: Brasil/Bolívia*. Dissertação de Mestrado, Universidade de São Paulo, São Paulo.

- Gordon, A. (1980). Notas sobre la Fonetica del Castellano en Bolivia. In A. Gordon and E. Rugg (Eds.), *Actas del Sexto Congreso de la Asociación Internacional de Hispanistas* (pp. 349–352). Toronto.
- Governo de São Paulo (2008). Retrieved October 12, 2013, from <http://www.saopaulo.sp.gov.br/imigracaojaponesa/historia.php>
- Grannier, D. M. (2000). Uma proposta heterodoxa para o ensino de português a falantes de espanhol. In N. Júdice (Ed.), *Português para estrangeiros: perspectivas de quem ensina* (pp. 57–80). Niterói: Intertexto.
- Grannier, D. M. (2004). Grandes dificuldades de comunicação devidas a falhas de pronúncia. In A. R. Simões, A. M. Carvalho and L. Wiedemann (Eds.), *Português para falantes de espanhol: Artigos selecionados escritos em português e inglês* (pp. 175–191). Campinas, Spain: Pontes.
- Greene, J. C. and Caracelli, V. J. (2003). Making paradigmatic sense of mixed methods practice. In A. Tashakkori and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 91–110). Thousand Oaks, Calif.: SAGE.
- Gregg, R. J. (1964). Scotch – Irish urban speech in Ulster. In B. Adams (Ed.), *Ulster Dialects: An Introductory Symposium*. Holywood, Co. Down: Ulster Folk Museum.
- Gudykunst, W. B. (Ed.) (2005). *Theorizing about intercultural communication*. Thousand Oaks: Sage Publ.
- Gutiérrez Marrone, N. (1992). *El habla de la ciudad de La Paz: Materiales para su estudio* (1. ed.). La Paz: Ed. Signo.
- Haug, S. and Sauer, L. (2006). Bestimmungsfaktoren internationaler Migration: Ein Überblick über Theorien zur Erklärung von Wanderungen. *Sozialwissenschaftlicher Fachinformationsdienst soFid*, (1), 7–34.
- Heigham, J. and Sakui, K. (2009). Ethnography. In J. Heigham and R. A. Croker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 91–111). Houndmills [et al.]: Palgrave Macmillan.
- Hofmeister, B. (1999). *Stadtgeographie* (7th edition). *Das geographische Seminar*. Braunschweig: Westermann.
- Hora, A. F. da (2000). Comparación fonológica del español y del portugués de Brasil. *Anuario brasileño de estudios hispánicos*, 10, 15–29.
- Ianni, O. (1979). Aspectos Políticos e Econômicas da Imigração Italiana. In Instituto Superior Brasileiro-Italiano de Estudos (Ed.), *Imigração italiana: estudos* (pp. 11–20). Porto Alegre: Universidade de Caxias do Sul.
- IBGE - Instituto Brasileiro Geral de Estatística (2010). *Censo 2010*, from [http://www.ibge.gov.br/home/estatistica/populacao/censo2010/tabelas\\_pdf/total\\_populacao\\_sao\\_paulo.pdf](http://www.ibge.gov.br/home/estatistica/populacao/censo2010/tabelas_pdf/total_populacao_sao_paulo.pdf).
- Ilari, R. (Ed.) (2002). *Gramática do Português Falado: Volume II: Níveis de análise lingüística* (4. ed. rev.). Campinas: Editora Unicamp.
- Ilari, R. et al. (2002). Os pronomes pessoais do português falado: roteiro para a análise. In A. T. de Castilho (Ed.), *Gramática do Português Falado: Volume IV: Estudos descritivos* (2nd ed., pp. 73–159). Campinas: Editora Unicamp.
- Ilari, R. and Basso, R. (2006). *O português da gente: A língua que estudamos, a língua que falamos*. São Paulo: Ed. Contexto.
- Instituto Boliviano de Comercio Exterior (2008). Bolivia: Migración, Remesas y Desempleo. *Comercio Exterior*, 16(159).
- Instituto Brasileiro de Geografia e Estatística. *De 2005 para 2008, acesso à Internet aumenta 75,3%*. Retrieved July 10, 2011, from [http://www.ibge.gov.br/home/presidencia/noticias/noticia\\_visualiza.php?id\\_noticia = 1517](http://www.ibge.gov.br/home/presidencia/noticias/noticia_visualiza.php?id_noticia = 1517).

- Instituto Nacional de Estadística de Bolivia. *Actualidad Estadística*. Retrieved September 7, 2011, from [http://www.ine.gob.bo/pdf/Bo\\_Es\\_Na/BEN\\_2010\\_1.pdf](http://www.ine.gob.bo/pdf/Bo_Es_Na/BEN_2010_1.pdf).
- Ivankova, N. and Creswell, J. W. (2009). Mixed Methods. In J. Heigham and R. A. Croker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 135–161). Houndmills Basingstoke Hampshire England New York: Palgrave Macmillan.
- Ivankova, N. and Kawamura, Y. (2010). Emerging trends in the utilization of integrated designs in the social, behavioral and health sciences. In A. Tashakkori and C. Teddlie (Eds.), *Sage handbook of mixed methods in social and behavioral research* (2nd ed., pp. 581–611). Los Angeles, Calif.: SAGE.
- Janssen, J. and Laatz, W. (2007). *Statistische Datenanalyse mit SPSS für Windows* (6th ed.). Berlin: Springer (eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests.).
- Jenny, R. K. and Obaid, T. (2004). *Meeting the Challenges of Migration: Progress since the ICPD*. Retrieved July 14, 2011, from United Nations Population Fund (UNFPA): [www.unfpa.org/webdav/site/global/shared/documents/publications/2004/migration\\_icpd.pdf](http://www.unfpa.org/webdav/site/global/shared/documents/publications/2004/migration_icpd.pdf).
- Jensen, J. B. (1989). On the Mutual Intelligibility of Spanish and Portuguese. *Hispania*, 72(4), 848–852.
- Jensen, J. B. (2004). The relative influence of Spanish and English in the Portuguese writing of bilingual students. In A. R. Simões, A. M. Carvalho and L. Wiedemann (Eds.), *Português para falantes de espanhol: Artigos selecionados escritos em português e inglês* (pp. 67–82). Campinas, Spain: Pontes.
- Johnson, R. B., Onwuegbuzie, A. J. and Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, 2(1), 112–133.
- Kellman, S. G. (2003). *Switching languages: Translingual writers reflect on their craft*. Lincoln: Univ. of Nebraska Press.
- Kerswill, P. (2008). Koineization and Accommodation. In J. K. Chambers, P. Trudgill and N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change* (pp. 669–702). John Wiley and Sons Ltd.
- Kerswill, P. and Williams, A. (2005). New towns and koineization: linguistic and social correlates. *Linguistics*, 43(5), 1023–1048.
- Kewitz, V. (2009). Gramaticalização, semantização e discursivização das preposições a e para no português brasileiro dos séculos XIX e XX. In A. T. de Castilho (Ed.), *História do português paulista, Vol. 1* (pp. 603–633). Campinas: Instituto de Estudos da Linguagem/Unicamp.
- Klee, C. and Caravedo, R. (2009). Andean Spanish and the Spanish of Lima: Linguistic Variation and Change in a Contact Situation. In C. Mar-Molinero and M. Stewart (Eds.), *Language and globalization: Globalization and language in the Spanish speaking world. Macro and micro perspectives* (pp. 94–113). Basingstoke: Palgrave Macmillan.
- Koch, P. and Oesterreicher, W. (1990). *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch. Romanistische Arbeitshefte: Vol. 31*. Tübingen: Niemeyer.
- Koch, P. and Oesterreicher, W. (2011). *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch* (2nd rev. ed., Vol. 31). *Studium*, 31. Berlin: de Gruyter.
- Kohlhepp, G. (1997). Ersticken die Metropolen an sich selbst? São Paulo – Größter industrieller Ballungsraum Lateinamerikas. *Großstädte. Der Bürger im Staat*, 47, 137–143.
- Korff, R. (2007). *Einführung: Wie urban ist die Welt?: Megastädte und die zunehmende Urbanisierung*. Retrieved June 11, 2011, from Bundeszentrale für politische Bildung: [http://www.bpb.de/themen/3WXP60D,0,0,Einf%FCChrung%3A\\_Wie\\_urban\\_ist\\_die\\_Welt.html](http://www.bpb.de/themen/3WXP60D,0,0,Einf%FCChrung%3A_Wie_urban_ist_die_Welt.html).
- Krefeld, T. (2004). *Einführung in die Migrationslinguistik*. Tübingen: Gunter Narr Verlag.

- Krefeld, T. and Pustka, E. (2010). Für eine perzeptive Varietätenlinguistik. In T. Krefeld and E. Pustka (Eds.), *Spazi comunicativi: Vol. 8. Perzeptive Varietätenlinguistik* (pp. 9–28). Frankfurt am Main: Lang.
- Labov, W. (1963). The social motivation of a sound change. *Word*, 18, 1–42.
- Labov, W. (1966). *The social stratification of English in New York City. Urban language series*. Washington: Center for Applied Linguistics.
- Labov, W. (1972). *Language in the inner city: Studies in the black English vernacular*. Philadelphia: University of Pennsylvania.
- Labov, W. (1974). On the use of the present to explain the past. In L. Heilmann (Ed.), *Proceedings of the 11th International Congress of Linguists* (pp. 825–851). Bologna: Il Mulino.
- Labov, W. (1991). *Sociolinguistic patterns* (11. print.). *Conduct and communication series*. Philadelphia: Univ. of Pennsylvania Press.
- Labov, W. (2004). Ordinary Events. In C. Fought (Ed.), *Sociolinguistic Variation: Critical Reflections* Oxford Studies in Sociolinguistics (pp. 31–43). New York et al.: Oxford University Press.
- Labov, W. (2006). *The social stratification of English in New York City* (2. ed.). Cambridge: Cambridge Univ. Press.
- Lambert, W. E. et al. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, 3, 44–51.
- Leather, J. and James, A. (1991). The Acquisition of Second Language Speech. *Studies in Second Language Acquisition*, 13, 305–341.
- Lee, E. S. (1966). A Theory of Migration. *Demography*, 3, 47–57.
- Leiria, I. (1998). Falemos antes de "verdadeiros amigos". In P. F. Pinto and N. Júdice (Eds.), *Para acabar de vez com Tordesilhas* (pp. 11–29). Lisboa: Ed. Colibri.
- Lenz, A. N. (2010). Zum Salienzbegriff und zum Nachweis salienter Merkmale. In C. A. Anders, M. Hundt and A. Lasch (Eds.), *Linguistik - Impulse and Tendenzen: Vol. 38. Perceptual dialectology. Neue Wege der Dialektologie* (pp. 89–110). Berlin: de Gruyter.
- Loffredo D'Ottaviano, M. C. (2006). *Condomínios Fechados na Região Metropolitana de São Paulo: fim do modelo centro rico versus periferia pobre?* from [www.abep.nepo.unicamp.br/entontro2006/docspdf/ABEP2006\\_590.pdf](http://www.abep.nepo.unicamp.br/entontro2006/docspdf/ABEP2006_590.pdf).
- Lipski, J. M. (2010). Spanish and Portuguese in Contact. In R. Hickey (Ed.), *Blackwell handbooks in linguistics. The handbook of language contact* (1st ed., pp. 550–580). Malden, Mass: Wiley-Blackwell.
- Lope Blanch, J. M. (1986). *El estudio del español hablado culto. Historia de un proyecto*. México: UNAM.
- Lopes, C. R. dos S. (2003). *A inserção de "a gente" no quadro pronominal do português* (Vol. 18). *Linguística iberoamericana*, 18. Madrid: Iberoamericana [et al.].
- Macedo, L. (16.9.2010). Paulistano gasta, em média, 2 horas e 42 minutos no trânsito todos os dias. *Globo São Paulo*. Retrieved July 09, 2011, from <http://g1.globo.com/sao-paulo/noticia/2010/09/paulistano-gasta-em-media-2-horas-e-42-minutos-no-transito-todos-os-dias.html>.
- Machline Oliveira e Silva, G. de (1992). Coleta de Dados. In M. C. Mollica and M. L. Braga (Eds.), *Cadernos Didáticos UFRJ: . Introdução à Sociolinguística Variacionista. o tratamento da variação* (pp. 101–114). Rio de Janeiro: Contexto.
- Marcuschi, L. A. (1989). Marcadores conversacionais no português brasileiro: formas, posições e funções. In A. T. d. Castilho (Ed.), *Português culto falado no Brasil* (pp. 281–321). Campinas: Editora Unicamp.

- Mayerthaler, W. (1981). *Morphologische Natürlichkeit. Linguistische Forschungen: Vol. 28.* Wiesbaden: Athenaion.
- Memorial do Imigrante (10.6.2011). *Imigrantes Estrangeiros Entrados No Estado de São Paulo - 1885-1961*, from [www.memorialdoimigrante.org.br/portalmi/templates/historico/e6.htm](http://www.memorialdoimigrante.org.br/portalmi/templates/historico/e6.htm).
- Mendes, C. M. and Matte, A. (2008). Euforia e Disforia do /R/ retroflexo da Apresentadora Ana Maria Braga. *CASA - Cadernos de Semiótica Aplicada*, 6(2), 1–10, from <http://www.fclar.unesp.br/seer/index.php?journal=casa>.
- Mendes, E. A. M. de (1998). Ensino do léxico português para hispanofalantes. In P. F. Pinto and N. Júdice (Eds.), *Para acabar de vez com Tordesilhas* (pp. 91–99). Lisboa: Ed. Colibri.
- Mendoza, J. G. (1991). *El castellano hablado en La Paz: Sintaxis divergente.* La Paz: Talleres Gráficos de la Facultad de Humanidades y Ciencias de la Educación.
- Mendoza, J. G. (2008). Bolivia. In A. Palacios (Ed.), *El español en América: Contactos lingüísticos en Hispanoamérica.* Barcelona: Editorial Ariel.
- Meyerhoff, M. (2006). Communities of Practice. In J. K. Chambers; P. Trudgill and N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change* (3rd ed., pp. 526–548). Malden et al.: Blackwell Publishing.
- Milroy, L. (1980). *Language and social networks. Language in society.* Baltimore: University Park Press.
- Milroy, L. (2004). Language Ideologies and Linguistic Change. In C. Fought (Ed.), *Sociolinguistic Variation: Critical Reflections. Oxford Studies in Sociolinguistics* (pp. 161–177). New York et al.: Oxford University Press.
- Milroy, L. (2006). Social Networks. In J. K. Chambers; P. Trudgill and N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change* (3rd ed., pp. 549–572). Malden et al.: Blackwell Publishing.
- Milroy, L. and Gordon, M. (2006). *Sociolinguistics: Method and interpretation* ([Reprinted], Vol. 34). *Language in society*, 34. Malden, Mass.: Blackwell.
- Milroy, L. and Milroy, J. (1992). Social network and social class: Toward an integrated sociolinguistic model. *Language in Society*, 21, 1–26.
- Milroy, L. and Preston, D. R. (1999). Introduction. *Journal of Language and Social Psychology*, 18(1), 4–9.
- Mollica, M. C. (2006). Sobre processos sintáticos que migram da fala para a escrita. In T. L. Face and C. Klee (Eds.), *Selected Proceedings of the 8th Hispanic Linguistics Symposium* (pp. 161–171). Somerville, MA: Cascadilla Proceedings Project.
- Mollica, M. C. and Braga, M. L. (Eds.) (1992). *Introdução à Sóciolingüística Variacionista: o tratamento da variação. Cadernos Didáticos UFRJ.* Rio de Janeiro: Contexto.
- Moraes, J., Callou, D. and Leite, Y. (2002). O sistema vocálico do português do Brasil: Caracterização acústica. In M. A. Kato (Ed.), *Gramática do Português Falado: Volume V: Convergências* (2nd ed., pp. 33–51). Campinas: Editora Unicamp.
- Moraes, M. (20.12.2004). Invasão silenciosa. *Época*, pp. 51–54.
- Moya, O. (1999). Situación de tres fonemas en los medios de la comunicación de La Paz. *Revista Lengua*, (9), 123–137.
- Müller, S. (2003). Sprachkontakt in bilingualen Sprachgemeinschaften: Code-Switching an der Deutschen Schule Barcelona. Köln: ZSM-Studien, from <http://www.zsm.uni-koeln.de/files/mueller.pdf>.
- Mufwene, S. (1990). Transfer and the substrate hypothesis in creolistics. *Studies in Second Language Acquisition*, 12, 1–23.

- Município em Mapas: Panorama. Político Administrativo.* Retrieved July 06, 2011, from <http://sempla.prefeitura.sp.gov.br/mm/panorama/>.
- Município em Mapas: Expansão Urbana.* Retrieved July 06, 2011, from <http://sempla.prefeitura.sp.gov.br/mm/panorama/>.
- Município em Mapas: Economia Urbana. Distribuição Espacial da Atividade Econômica.* Retrieved July 08, 2011, from [http://www.9.prefeitura.sp.gov.br/sempla/mm3/index.php?texto=corpo&tema\\_cod=2](http://www.9.prefeitura.sp.gov.br/sempla/mm3/index.php?texto=corpo&tema_cod=2).
- Município em Mapas: Panorama. Economia Urbana.* Retrieved July 06, 2011, from <http://sempla.prefeitura.sp.gov.br/mm/panorama/>.
- Município em Mapas: Economia Urbana. Mercado de Trabalho.* Retrieved July 08, 2011, from [http://www.9.prefeitura.sp.gov.br/sempla/mm3/index.php?texto=corpo&tema\\_cod=4](http://www.9.prefeitura.sp.gov.br/sempla/mm3/index.php?texto=corpo&tema_cod=4).
- Município em Mapas: Índices sociais. Mapa da vulnerabilidade social.* Retrieved July 08, 2011, from [http://www.9.prefeitura.sp.gov.br/sempla/mm/index.php?texto=corpo&tema\\_cod=6](http://www.9.prefeitura.sp.gov.br/sempla/mm/index.php?texto=corpo&tema_cod=6).
- Município em Mapas: Índices sociais. Índice de Desenvolvimento Humano IDH.* Retrieved July 08, 2011, from [http://www.9.prefeitura.sp.gov.br/sempla/mm/index.php?texto=corpo&tema\\_cod=2](http://www.9.prefeitura.sp.gov.br/sempla/mm/index.php?texto=corpo&tema_cod=2).
- Município em Mapas: Analfabetismo. Analfabetismo no Município.* Retrieved July 08, 2011, from [http://www.9.prefeitura.sp.gov.br/sempla/mm2/index.php?texto=corpo&tema\\_cod=2](http://www.9.prefeitura.sp.gov.br/sempla/mm2/index.php?texto=corpo&tema_cod=2).
- Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge Univ. Press.
- Muysken, P. (2007). Mixed Codes. In P. Auer and L. Wei (Eds.), *Handbook of multilingualism and multilingual communication* (pp. 315-340). Berlin: de Gruyter.
- Muysken, P. (2010). Scenarios for Language Contact. In R. Hickey (Ed.), *Blackwell handbooks in linguistics. The handbook of language contact* (1st ed., pp. 265–281). Malden, Mass: Wiley-Blackwell.
- Myers-Scotton, C. M. (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, C. M. (2005). *Duelling languages: Grammatical structure in codeswitching* [Reprinted]. Oxford: Clarendon.
- Nastasi, B. K., Hitchcock, J. H. and Brown, L. M. (2010). An inclusive framework for conceptualizing mixed methods design typologies. In A. Tashakkori and C. Teddlie (Eds.), *Sage handbook of mixed methods in social and behavioral research* (2nd ed., pp. 305–338). Los Angeles, Calif.: SAGE.
- Niedzilski, N. (1999). The Effect of Social Information on the Perception of Sociolinguistic Variables. *Journal of Language and Social Psychology*, 18(1), 62–85.
- Niehoff, S. (24.8.2006). *Das Portugiesische bei italienischen Immigranten in São Paulo*. Diplomarbeit, Universität zu Köln, Köln. Retrieved March 09, 2011, from: [miradouro.de/niehoff\\_2006.pdf](http://miradouro.de/niehoff_2006.pdf).
- Noll, V. (2001). *Das amerikanische Spanisch: Ein regionaler und historischer Überblick* (Vol. 46). *Romanistische Arbeitshefte*, 46. Tübingen: Niemeyer.
- Noonan, M. (2010). Genetic Classification and Language Contact. In R. Hickey (Ed.), *Blackwell handbooks in linguistics. The handbook of language contact* (1st ed., pp. 48–65). Malden, Mass: Wiley-Blackwell.
- O Globo* (15.9.2010). Bolivianos eram mantidos como escravos em oficina de costura em SP. from <http://oglobo.globo.com/cidades/sp/mat/2010/09/15/bolivianos-eram-mantidos-como-escravos-em-oficina-de-costura-em-sp-917641165.asp>.
- Olhar São Paulo: Contrastes Urbanos. Território e Desigualdades Sociais.* Retrieved July 09, 2011, from <http://sempla.prefeitura.sp.gov.br/olhar>.

- Olhar São Paulo: Violência e Criminalidade*. Retrieved July 09, 2011, from <http://sempla.prefeitura.sp.gov.br/mm/criminalidade>.
- Onwuegbuzie, A. J. and Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, 13(1), 48–63.
- Padgett, D. K. (2009). Qualitative and mixed methods in social work knowledge development. *Social Work*, 54(2), 101–105.
- Paes Almeida Filho, J. C. de (2004). Questões da interlíngua de aprendizes de português a partir ou com a interposição do espanhol (língua muito próxima). In A. R. Simões, A. M. Carvalho and L. Wiedemann (Eds.), *Português para falantes de espanhol: Artigos selecionados escritos em português e inglês* (pp. 183–191). Campinas, Spain: Pontes.
- Paiva, M. d. C. and Duarte, M. E. L. (Eds.) (2003). *Mudança lingüística em tempo real*. Rio de Janeiro: Contra Capa Livraria.
- Paiva, M. da C. and Pereira Scherre, M. M. (1999). Retrospectiva Sociolingüística: Contribuições do Peul. *DELTA*, 15, 201–232.
- Pampuch, T. and Echalar Ascarrunz, A. (2009). *Bolivien* (Orig., 4th rev. ed. Vol. 813). *Beck'sche Reihe Länder, 813*. München: Beck.
- Pappas, P. A. (2008). Stereotypes, variation and change: Understanding the change of coronal sonorants in a rural variety of Modern Greek. *Language Variation and Change*, 20(3), 493–526.
- Pereira Scherre, M. M. (1991). A concordância de número nos predicativos e nos participios passivos. *Organon*, 5(18), 52–70.
- Pereira Scherre, M. M. (1998). Variação da concordância nominal do português so Brasil: influência das variáveis posição, classe gramatical e marcas precedentes. In S. Große and K. Zimmermann (Eds.), *Biblioteca luso-brasileira: Vol. 6. "Substandard" e mudança no português do Brasil. [colóquio internacional ..., Berlin outubro de 1997]* (pp. 153–188). Frankfurt am Main: TFM - Teo Ferrer de Mesquita.
- Pereira Scherre, M. M. (2001). Major Linguistics Patterns in Noun Phrase Agreement in Brazilian Portuguese. In R. B. Finazzi and P. Tornaghi (Eds.). *Cinquant'anni di ricerche linguistiche: problemi, risultati e prospettive per il terzo millennio. Atti del 9. Convegno Internazionale di Linguisti, Milano, 8-9-10 ottobre 1998* (pp. 461–473). Alessandria: Edizioni dell'Orso.
- Pereira Scherre, M. M. and Naro, A. J. (1991). Marking in Discourse: Birds of a Feather. *Language Variation and Change*, 3, 23–32.
- Pestana Allegro, F. R. and Madureira, S. (2008). Speech production and perception tasks: A case study of an intermediate Brazilian student learning Spanish as a second language. In L. Wiedemann (Ed.), *Português para Falantes de Espanhol. Ensino e Aquisição. Artigos selecionados escritos em Português e Inglês*. (pp. 191–218). Campinas: Pontes.
- Petendra, B. (2007). Integration von Migranten und Migrantinnen: Eine Analyse sozialraumorientierter Projekte. *Sic et Non. zeitschrift für philosophie und kultur. im netz.*, 8, 1–157. Retrieved July 15, 2011.
- Pfänder, S. (2002). Contacto y cambio lingüístico en Cochabamba (Bolivia). In N. Díaz, R. Ludwig and S. Pfänder (Eds.), *Lengua y sociedad en el mundo hispánico: Vol. 9. La Romania americana. Procesos lingüísticos en situaciones de contacto* (pp. 219–253). Frankfurt am Main: Iberoamericana; Vervuert.
- Pfänder, S. (2009). *Gramática mestiza*. La Paz: Academia Boliviana de la Lengua/Editorial Signo.
- Placencia, M. E. (2001). Inequality in Address Behavior at Public Institutions in La Paz, Bolivia. *Anthropological Linguistics*, 43(2), 198–217.

- Ploog, K. and Reich, U. (2005). Rasgos socio-indexicales en la dinámica urbana. *Lexis*, XXIX(1), 47–78.
- Ploog, K. and Reich, U. (2006). Urbane Prozess: Migration und Sprachdynamik. In T. Krefeld (Ed.), *Spazi comunicativi: Vol. 1. Modellando lo spazio in prospettiva linguistica. [... interventi tenuti ... in occasione del XXIX Deutscher Romanistentag (Kiel, 28/9 - 3/10/2003) ...]* (pp. 215–256). Frankfurt am Main: Lang.
- Portal Brasil (2010). *Infográfico: Numeros da população*. Retrieved August 20, 2013, from <http://www.brasil.gov.br/infograficos/numeros-da-populacao>
- Prada, J. J. (October 2001). *Marcadores del discurso en español: Análisis y representación*. Tesis de maestría, Universidad de la República, Uruguay. Retrieved March 25, 2011, from Universidad de la República: <http://www.fing.edu.uy/inco/pecediba/bibliote/tesis/tesis-prada.pdf>.
- Prefeitura de São Paulo*, from [http://www.prefeitura.sp.gov.br/cidade/secretarias/subprefeituras/subprefeituras/dados\\_demograficos/index.php?p=12758](http://www.prefeitura.sp.gov.br/cidade/secretarias/subprefeituras/subprefeituras/dados_demograficos/index.php?p=12758).
- Preston, D. R. (1989). *Perceptual Dialectology*. Berlin, New York: Mouton de Gruyter.
- Preston, D. R. (Ed.) (1999). *Handbook of perceptual dialectology: Vol. 1*. Amsterdam: John Benjamins Publishing Company.
- Preston, D. R. (2001). Style and the psycholinguistics of sociolinguistics: the logical problem of language variation. In P. Eckert and J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 279–304). Cambridge [et al.]: Cambridge Univ. Press.
- Preston, D. R. (2010). Perceptual Dialectology in the 21st Century. In C. A. Anders, M. Hundt and A. Lasch (Eds.), *Linguistik - Impulse und Tendenzen: Vol. 38. Perceptual dialectology. Neue Wege der Dialektologie* (pp. 1–29). Berlin: de Gruyter.
- Preti, D. (Ed.) (2003). *Análise de Textos Oraís* (6.ed.). São Paulo: Humanitas.
- Preti, D. (Ed.) (2004). *Estudos de língua oral e escrita. Série Dispersos*. Rio de Janeiro: Editora Lucerna.
- Pries, L. (1997). Neue Migration im transnationalen Raum. In L. Pries (Ed.), *Transnationale Migration. Soziale Welt: Sonderband 12* (pp. 15–46). Baden-Baden.
- Pustka, E. (2008). accent(s) parisien(s) - Auto- und Heterorepräsentationen stadtsprachlicher Merkmale. In T. Krefeld (Ed.), *Sprachen und Sprechen im städtischen Raum* (pp. 213–249). Frankfurt a. M. u.a.: Peter Lang.
- Quilis, A. (1979). Comparación de los sistemas fonológicos del español y del portugués. *Revista Española de Lingüística*, 1–29. Retrieved March 07, 2011, from <http://www.uned.es/sel/pdf/ene-jun-79/quilis79.pdf>.
- Rehrmann, N. and Koechert, A. (Eds.) (1999). *Spanien und die Sepharden: Geschichte, Kultur, Literatur* (Vol. 3). *Romania Judaica, 3*. Tübingen: Niemeyer.
- Rallis, S. F. and Rossmann, G. B. (2009). Ethics and Trustworthiness. In J. Heigham and R. A. Croker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 283–287). Houndmills Basingstoke Hampshire England New York: Palgrave Macmillan.
- Ramos, H. and et al. (Eds.) (2011). *Por uma vida melhor* (Vol. 6). *Viver, aprender, 6*. São Paulo: Global Editora e Distribuidora Ltda.
- Reich, U. (2002). *Freie Pronomina, Verbalklitika und Nullobjekte im Spielraum diskursiver Variation des Portugiesischen in São Paulo* (Bd. 62). *Romanica Monacensia, Bd. 62*. Tübingen: Narr (Univ., Diss--München, 2000.).
- Ribas Fialho, V. (2005). Proximidade entre línguas: algumas considerações sobre a aquisição do espanhol por falantes nativos de português brasileiro. *Espéculo. Revista de estudos literários.*, 31, from <http://www.ucm.es/info/especulo/numero31/falantes.html>.

- Ribeiro, L. Cesar Queiroz de (2008). O desafio das metrópoles. *Le Monde Diplomatique Brasil*, 12–13.
- Richards, K. (2009). Interviews. In J. Heigham and R. A. Croker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 182–199). Houndmills Basingstoke Hampshire England New York: Palgrave Macmillan.
- Rickford, J. R. (2001). Style and stylizing. In P. Eckert and J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 220–231). Cambridge [et al.]: Cambridge Univ. Press.
- Roberts, C. (1997). Transcribing Talk: Issues of Representation. *TESOL Quarterly*, 31(1), 167–172. Retrieved February 10, 2011, from <http://www.jstor.org/stable/3587983>.
- Rocha, D. J. de la (1987). *Apuntes sobre las lenguas nativas en el dialecto español de Bolivia*. La Paz: III.
- Rodrigues, Â. C. S. (1987). *A concordância verbal no português popular em São Paulo*. São Paulo: USP.
- Rodrigues, Â. C. S. (2000). Concordância verbal e saliência social no português popular em São Paulo. In E. Gärtner, C. Hundt and A. Schönberger (Eds.), *Biblioteca luso-brasileira: Vol. 15. Estudos de sociolinguística brasileira e portuguesa*. Frankfurt am Main: TFM Teo Ferrer de Mesquita.
- Rodrigues, Â. C. S. (2009). Fotografia Sociolinguística do Português do Brasil: o português popular em São Paulo. In A. T. d. Castilho (Ed.), *História do português paulista, Vol. 1* (pp. 151–157). Campinas: Instituto de Estudos da Linguagem/Unicamp.
- Rodrigues, Â. C. S. and Ferreira Netto, W. (2000). Transcrição de Inquéritos: Problemas e Sugestões. In Megale, H. (Ed.), *Filologia Bandeirante* (pp.171-193). São Paulo: Humanitas.
- Rodríguez, J. and Busso, G. (2009). *Migración interna y desarrollo en América Latina entre 1980 y 2005: Un estudio comparativo con perspectiva regional basada en siete países* (1. ed., Vol. 102). *Publicación de las Naciones Unidas, 102*. Santiago de Chile: Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL).
- Rolnik, R. (2008). A lógica da desordem. *Le Monde Diplomatique Brasil*, 10–11.
- Rolnik, R. and Frúgoli, H., JR. (2001). Reestruturação urbana da metrópole paulistana: a Zona Leste como território de rupturas e permanências. *Cadernos Metrópole*, 43–66, from [http://web.observatoriodasmetrosoles.net/download/cm\\_artigos/cm6\\_25.pdf](http://web.observatoriodasmetrosoles.net/download/cm_artigos/cm6_25.pdf).
- Rowntree, D. (2004). *Statistics without tears: A primer for non-mathematicians* (Classic ed.). Boston: Pearson/A, and B.
- Sachdev, I., Arnold, D. Y. and Yapita, J. D. de (2006). Indigenous Identity and Language: Some Considerations from Bolivia and Canada. *BISAL*, 1, 107–128.
- Sankoff, G. (2008). Linguistic Outcomes of Language Contact. In J. K. Chambers, P. Trudgill and N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change* (pp. 638–668). John Wiley and Sons Ltd.
- Saraiva, C. and Marques, E. (2007). A dinâmica social das Favelas da Região Metropolitana de São Paulo. *Pensamento & Realidade (Revistas PUC-SP)*, 27, 12-41.
- Schmid, S. (1995). Morphological naturalness in Spanish-Italian interlanguages. In H. Pishwa (Ed.), *Tübinger Beiträge zur Linguistik: Vol. 399. The development of morphological systematicity. A cross-linguistic perspective* (pp. 263–289). Tübingen: Narr.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Teaching*, 10, 209–241.
- Siebel, W. (1997). Die Stadt und die Zuwanderer. In H. Häußermann and I. Oswald (Eds.), *Leviathan : Sonderheft. Zuwanderung und Stadtentwicklung* (pp. 30–41). Opladen: Westdt. Verlag.
- Siegel, J. (1985). Koines and koineization. *Language in Society*, 14, 357–378.

- Siegel, J. (2003). Substrate influence in creoles and the role of transfer in second language acquisition. *Studies in Second Language Acquisition*, 25(2), 185–209.
- Siegel, J. (2008). *The emergence of Pidgin and Creole languages*. [Oxford linguistics]. Oxford: Oxford Univ. Press.
- Simmel, G. (1908). *Soziologie - Untersuchungen über die Formen der Vergesellschaftung*. Berlin: Duncker & Humblot.
- Santos, A., G. d. et al. (2004). *A dura vida dos imigrantes bolivianos em São Paulo*, from Universidade Anhembi Morumbi: [www.anhembi.br/publicue/cgi/cgilua.exe/sys/start.htm?infoid=17002&sid=1932](http://www.anhembi.br/publicue/cgi/cgilua.exe/sys/start.htm?infoid=17002&sid=1932).
- Silva, S. A. da (2005). *Bolivianos: A presença da cultura andina* (1a ed.). *Séria Lazuli, imigrantes no Brasil*. São Paulo: Companhia Ed. Nacional.
- Silva, S. A. da (2006). Bolivianos em São Paulo: entre o sonho e a realidade. *Estudos Avançados*, 20(57), 157–170.
- Silva, S. A. da (2008). *Faces da Latinidade Hispano-Americanos em São Paulo* (Vol. 55). *Textos Nepo*, 55. Campinas.
- Silverstein, M. (1976). Shifters, Linguistic Categories and Cultural Description. In K. Basso and H. A. Selby (Eds.), *Meaning in Anthropology* (pp. 11–56). Albuquerque: New Mexico Press.
- Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, 23, 193–229.
- Simões, A. R., Carvalho, A. M. and Wiedemann, L. (Eds.) (2004). *Português para falantes de espanhol: Artigos selecionados escritos em português e inglês*. Campinas, Spain: Pontes.
- Simões, A. R. and Kelm, O. R. (1991). O processo de aquisição das vogais semi-abertas "é, ó" /ε, ɔ/ do português (brasileiro) como lingua estrangeira. *Hispania*, 74(3), 654–665.
- Sivertsen, E. (1960). *Cockney Phonology*. Oslo: Oslo University Press.
- Smits, C. (1998). Two Models for the study of language contact: A psycho-linguistic perspective versus a socio-cultural perspective. In M. Schmid, J. Austin and D. Stein (Eds.), *Amsterdam studies in the theory and history of linguistic science : Series 4, Current issues in linguistic theory. Historical linguistics 1997. Selected papers from the 13th International Conference on Historical Linguistics Düsseldorf 10 - 17 August 1997* (pp. 377–390). Amsterdam: Benjamins.
- Souza Campos, O. G. de and Rodrigues, Â. C. (2002). Flexão nominal: Indicação de Pluralidade no Sintagma Nominal. In R. Ilari (Ed.), *Gramática do Português Falado: Volume II: Níveis de análise lingüística* (4th ed., pp. 101–122). Campinas: Editora Unicamp.
- Sydow, E. (2003). *Bolivianos escravizados em São Paulo*, from Rede social de justiça e direitos humanos: [www.social.org.br/relatorio2003/relatorio021.htm](http://www.social.org.br/relatorio2003/relatorio021.htm).
- Tarallo, F. (1996). Diagnosticando uma gramática brasileira: O português d'aquém e d'além-mar ao final do século XIX. In I. Roberts and M. A. Kato (Eds.), *Coleção Repertórios: . Português brasileiro - uma viagem diacrônica. Homenagem a Fernando Tarallo* (2nd ed., pp. 69–106). Campinas, SP, Brasil: Editora da Unicamp.
- Tarallo, F. (2001). *A pesquisa sociolingüística* (7. ed., 3 impr., Vol. 9). *Série princípios*, 9. São Paulo: Ed. Ática.
- Tarone, E. (2007). Sociolinguistic Approaches to Second Language Acquisition Research - 1997–2007. *The Modern Language Journal*, 91: 837–848.

- Taschner, S. P. (2003). Um olhar sobre a habitação em São Paulo. *Cadernos Metrópole*, (9), 81–117, from [http://web.observatoriodasmetrolopes.net/download/cm\\_artigos/cm9\\_41.pdf](http://web.observatoriodasmetrolopes.net/download/cm_artigos/cm9_41.pdf).
- Taschner, S. P. and Bógus, L. M. M. (1999). São Paulo como patchwork: unindo fragmentos de uma cidade segregada. *Cadernos Metrópole*, (1), 33–81. Retrieved June 12, 2010, from [http://web.observatoriodasmetrolopes.net/download/cm\\_artigos/cm1\\_2.pdf](http://web.observatoriodasmetrolopes.net/download/cm_artigos/cm1_2.pdf).
- Teddlie, C. and Tashakkori, A. (2003). Major Issues and Controversies in the Use of Mixed Methods in the Social and Behavioral Sciences. In A. Tashakkori and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 3–50). Thousand Oaks, Calif.: SAGE.
- Tedlock, B. (2003). Ethnography and Ethnographic Representation. In N. K. Denzin and Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 455–485). Thousand Oaks, Calif.: SAGE.
- Thakerar, J. N., Giles H. and Cheshire, J. (1982). Psychological and linguistic parameters of speech accommodation theory. In C. Fraser and K. R. Scherer (Eds.), *Advances in the social psychology of language* (pp.205-255). Cambridge: Cambridge University Press.
- Thomason, S. (2001). *Language contact: An introduction*. Washington, D.C: Georgetown Univ. Press.
- Thomason, S. (2006). Contact as a Source of Language Change. In B. D. Joseph and R. D. Janda (Eds.), *Blackwell handbooks in linguistics. The handbook of historical linguistics* (pp. 687–712). Carlton: Blackwell Publishing.
- Thomason, S. and Kaufman, T. (1988). *Language contact, creolization and genetic linguistics*. Berkeley/Los Angeles/Oxford: University of California Press.
- Thun, H. (1986). Zum Status der spanisch-portugiesischen Sprachmischung im Norden Uruguays. *Neue Romania*, 4, 37–74.
- Thun, H. (2000). El yeísmo y el leísmo portugués en el Norte del Uruguay y en el Sur del Brasil. Un cambio lingüístico circular. In C. Klettke, A. C. Franco, G. Hammermüller and J. Schmidt-Radefeldt (Eds.), *Ästhetik der Texte - Varietät von Sprache. Beiträge zu Paul Valéry und zur romanischen Philologie ; Festschrift für Jürgen Schmidt-Radefeldt zum 60. Geburtstag* (pp. 289–302). Tübingen: Narr.
- Thun, H., Elizaincín, A. and Boller, F. (2000). *Atlas lingüístico diatópico y diastrático del Uruguay: (ADDU). Dialectologia pluridimensionalis Romanica: Vol. 10*. Kiel: Westensee-Verl.
- Traugott, E. C. (Juli 1997). *The discourse connective "after all": A historical pragmatic account*. Paper presented at the 10th International Congress of Linguists. Paris.
- Trudgill, P. (1986). *Dialects in Contact*. Oxford: Blackwell.
- Trudgill, P. (2010). Contact and Sociolinguistic Typology. In R. Hickey (Ed.), *Blackwell handbooks in linguistics. The handbook of language contact* (1st ed., pp. 299–319). Malden, Mass: Wiley-Blackwell.
- UNDP. *International Human Development Indicators*. Retrieved July 05, 2011, from <http://hdrstats.undp.org/en/countries/profiles/BOL.html>.
- Urban, G. (1991). The Semiotics of State-Indian Linguistic Relationships: Peru, Paraguay and Brazil. In G. Urban and J. Sherzer (Eds.), *Nation-States and Indians in Latin America* (pp. 306–330). Austin: University of Texas Press.
- Urbano, H. (2002). Aspectos basicamente interacionais dos marcadores discursivos. In M. H. de Moura Neves (Ed.), *Gramática do Português Falado: Volume VII: Novos estudos* (2nd ed., pp. 195–258). Campinas: Editora Unicamp.
- Urbano, H. (2003). Marcadores conversacionais. In D. Preti (Ed.), *Análise de Textos Oraís* (6th ed., pp. 93–116). São Paulo: Humanitas.

- Vandresen, P. (2000). A pesquisa sociolingüística no Brasil - uma retrospectiva dos últimos 25 anos. In E. Gärtner, C. Hundt and A. Schönberger (Eds.), *Biblioteca luso-brasileira: Vol. 15. Estudos de sociolingüística brasileira e portuguesa* (pp. 15–39). Frankfurt am Main: TFM Teo Ferrer de Mesquita.
- Varella, T. (2004). Imigrantes bolivianos vivem como escravos em São Paulo. *Espaco Cidadania*, (26), from [www.metodista.br/cidadania/numero-26](http://www.metodista.br/cidadania/numero-26).
- Vasconcelos, P. A. de (2006). São Paulo e Salvador: Duas metrópoles brasileiras na longa duração (1549/1554-2004). In Carlos, A.F.A. (Ed.), *Geografias das metrópoles* (pp. 45–63). São Paulo: Contexto.
- Véras, M. Pardini Bicudo (1999). Territorialidade e cidadania em tempos globais: imigrantes em São Paulo. *Cadernos Metrópole*, (2), 73–119.
- Vilela, S. (15.3.2010). *Bairro de São Paulo reflete interseção entre diferentes culturas no Brasil*, from [www.dw-world.de/dw/article/0,,5300584,00.html](http://www.dw-world.de/dw/article/0,,5300584,00.html).
- Villa, M. and Rodriguez, J. (1996). Demographic trends in Latin America's metropolises, 1950 - 1990. In A. Gilbert (Ed.), *The mega-city in Latin America* (pp. 26–52). Tokyo, New York, NY, Paris: United Nations Univ. Press.
- Villaça, F. J. Magalhães and Zioni, S. M. (2005). *OS TRANSPORTES SOBRE TRILHOS NA REGIÃO METROPOLITANA DE SÃO PAULO: O Poder público acentuando a Desigualdade*. Retrieved June 12, 2010, from <http://www.cbtu.gov.br/monografia/2005/publicacao/monografia07.pdf>.
- Villaça, F. (1997). *Efeitos do Espaço sobre o Social*. Recife.
- Waiselfisz, J. J. (2010). *Mapa da Violência 2010: Anatomia dos homicídios no Brasil*. São Paulo: Instituto Sangari.
- Wiedemann, L. (Ed.) (2008). *Português para Falantes de Espanhol: Ensino e Aquisição*. Artigos selecionados escritos em Português e Inglês. Campinas: Pontes.
- Williams, A. and Kerswill, P. (1999). Dialect levelling: change and continuity in Milton Keynes, Reading and Hull. In P. Foulkes and G. Docherty (Eds.), *Urban Voices. Accents Studies in the British Isles*. (pp. 141–162). London: Arnold.
- Wurzel, W. (1987). System-Dependent Morphological Naturalness in Inflection. In W. U. Dressler (Ed.), *Studies in language: Companion series: Vol. 10. Leitmotifs in Natural Morphology* (pp. 59–96). Amsterdam and Philadelphia: Benjamins.

## Annex I: Overview of the informants

Table I.1: Overview of the informants' social characteristics. In order to obtain the links to the transcriptions and audio files (Exmeralda), please write an email to [steffi.niehoff@web.de](mailto:steffi.niehoff@web.de).

Speaker	Age at the time of the interview	Sex	Year of the arrival to São Paulo	Profession	Place of residence
A1	23	m	2006	Sales assistant	<i>Centro</i>
A2	18	f	2005	Sales assistant	<i>Zona Leste</i>
A3	34	f	2004	Sales assistant	<i>Zona Leste</i>
A4	21	m	2006	Seamster	<i>Zona Leste</i>
A5	28	f	2000	Dona	<i>Centro</i>
A6	34	m	1999	Seamster	<i>Zona Leste</i>
A7	32	f	1994	Seamstress	<i>Centro</i>
A8	27	f	2005	Receptionist	<i>Centro</i>
A9	42	m	1994	Dono	<i>Zona Leste</i>
A10	22	m	2006	Seamster	<i>Periferia</i>
A11	20	m	2002	Seamster	<i>Periferia</i>
A12	39	m	1990	Dono	<i>Zona Leste</i>
A13	33	m	2000	Dono	<i>Zona Leste</i>
A14	28	m	2006	Seamster	<i>Zona Norte</i>
A15	28	m	2000	Dono	<i>Centro</i>
A16	32	f	2004	Physician	<i>Zona Leste</i>
A17	31	m	2007	Physician	<i>Centro</i>
A18	30	m	2004	Physician	<i>Centro</i>
A19	42	m	1988	Physician	<i>Centro</i>
A20	27	f	2006	Physician	<i>Guarulhos</i>
A21	62	m	1964	Engineer	<i>Centro</i>
A22	27	f	2008	Physician	<i>Zona Sul</i>
A23	45	f	1990	Pharmacist	<i>Zona Oeste</i>
A24	25	m	2004	Seamster/ Sales assistant	<i>Centro</i>
A25	60	f	1989	Seamstress/ Sales assistant	<i>Zona Leste</i>
A26	25	f	2005	Seamstress	<i>Zona Leste</i>
A27	35	f	2006	Sales assistant	<i>Zona Leste</i>
A28	28	m	2002	Seamster	<i>Zona Leste</i>
B1	27	m	2006	Seamster	<i>Centro</i>

<b>B2</b>	30	m	2004	Seamster	<i>Zona Leste</i>
<b>B3</b>	21	m	2007	Seamster	<i>Zona Leste</i>
<b>B4</b>	27	m	2003	Seamster	<i>Centro</i>
<b>B5</b>	19	m	2008	Seamster	<i>Zona Leste</i>
<b>B6</b>	30	m	2003	Seamster	<i>Zona Leste</i>
<b>B7</b>	40	m	1986	Journalist	<i>Centro</i>
<b>B8</b>	48	f	1989	Physician	<i>Osasco</i>
<b>B9</b>	24	m	2005	Seamster	<i>Zona Leste</i>
<b>B10</b>	23	f	2007	Seamstress	<i>Periferia</i>
<b>B11</b>	29	f	2008	Seamstress	<i>Periferia</i>
<b>B12</b>	21	m	2005	Seamster	<i>Centro</i>
<b>S I</b>	25	m	2008	Seamster	<i>Centro</i>
<b>SII</b>	26	m	1999	Seamster	<i>Centro</i>
<b>SIII</b>	28	m	2003	Seamster	<i>Centro</i>

The social characteristics are added in the same order as in Table I.1 when the informants are cited in the study. The code used for the characterization:

Table I.2: Speaker's attributes indicated by citation in the text

Attribute	Code	Decoded
Corpus group	A	Speech example is used for linguistic analysis.
	B	Speech example is only used in non-linguistic contexts.
Age	(23)	Age of informant at time of the interview: 22.
Gender	m	Speaker's sex: masculine
	f	Speaker's sex: feminine
Year of arrival to São Paulo	2002	Speaker arrived to São Paulo in 2002.
Profession	D	<i>Dono/dona</i>
	E	Engineer
	J	Journalist
	P	Physician
	Ph	Pharmacist
	R	Receptionist
	S	Seamstress
	SA	Sales assistant
Place of residence	Ce	<i>Centro</i> (Center of São Paulo)
	Gu	<i>Guarulhos</i> (City in São Paulo's northeastern periphery)
	Os	<i>Osasco</i> (District in São Paulo's northwest)
	ZL	<i>Zona Leste</i> (Eastern zone of São Paulo)
	ZN	<i>Zona Norte</i> (Northern zone of São Paulo)
	ZO	<i>Zona Oeste</i> (Western zone of São Paulo)
	ZS	<i>Zona Sul</i> (Southern zone of São Paulo)
	Pe	<i>Periféria</i> (Periphery of São Paulo)

Accordingly, A1 is classified as A1 (23) m. 2006 SA Ce, i.e. he is male, 23 years old, has come to São Paulo in the year 2006, works as a sales assistant and lives in the center of São Paulo.

## Annex II: Transcription protocol

- As the sound files are available, no phonetic transcription was intended, but still I tried to capture the particularities of the Bolivians' speech, e.g. by transcribing *tô* in accordance with the phonetic realization.
- The interviews were transcribed normatively, from a Portuguese point of view. In the case of ambiguous words, Portuguese orthography was used.
- Only if the word was unambiguously identified as Castilian or found in a Castilian context (word before and after unambiguously identified as Castilian), it would be written according to Castilian orthography.
- To differentiate between orthographical homonyms of Castilian and Portuguese, they were spelled differently (Castilian *no*, for example, was transcribed as *nô*).
- Punctuation marks were used to facilitate the understanding of the interviews, indicating questions by question marks or the end of a phrase and short pauses by full stop.
- To reflect the orality of the interviews in the written representation as accurately as possible, pauses, interruptions, etc. were included in the transcription.
- Table II.1 gives an overview of further signs that were used in the transcriptions:

Table II.1: Explanation of the signs used in the transcription of the interviews

Occurrence	Sign	Example
Incomprehensible word(s)	(?)	Con coreanos estoy (?)
Suggestion for the meaning	(suggestion?)	Ajudante geral tô (trabalhando?)
Prolonged realization of the vowel or the consonant	::	Mh:::
Interruption of the sentence or the word	...	Co... conozco mui... mui... muita gente
Descriptive comments	either: (( )) or: separately in a further line	((laughing)) A client enters the shop.
Pause	((duration of pause in seconds))	Para ganhar un poquito más. ((0,8s))

Annex III: The questionnaire

Figure III.1: Picture story (cf. <http://www.4teachers.de/?action=material&id=32456#top>)

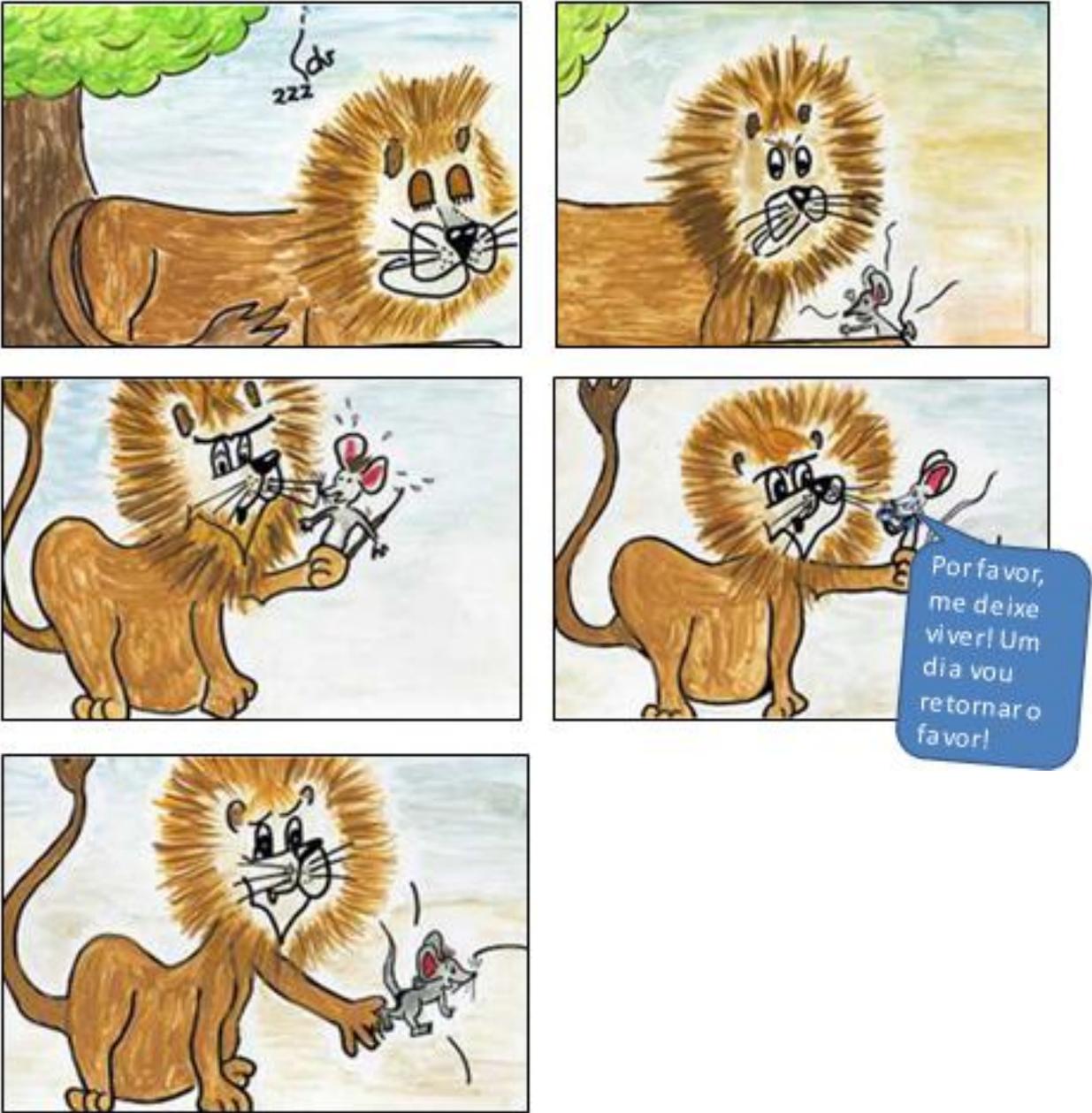
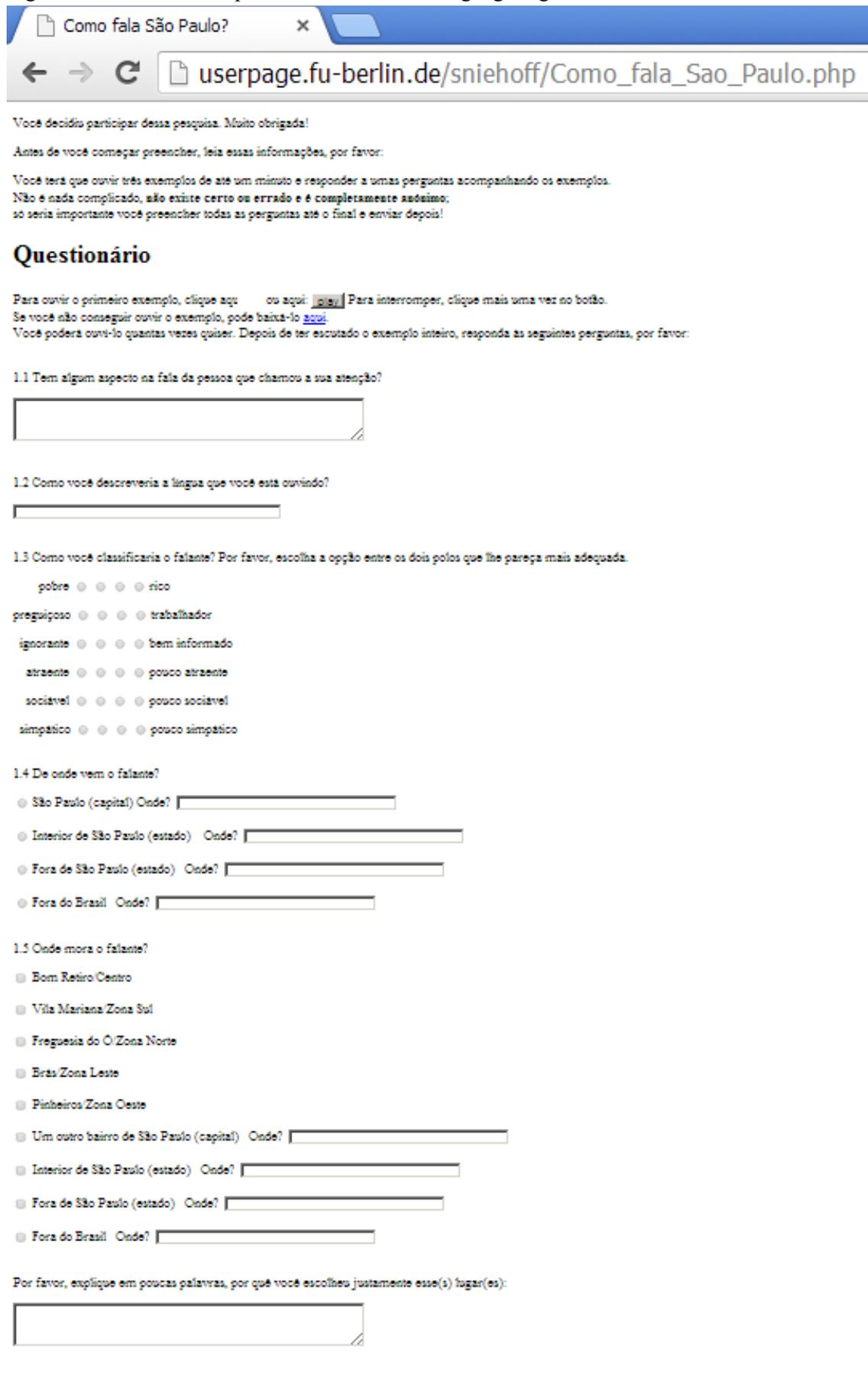




Figure III.2: The online questionnaire / test of language regard: Screenshots



---

Para ouvir o segundo exemplo, clique aqui [ou aqui](#) [play](#) Para interromper, clique mais uma vez no botão.

Se você não conseguir ouvir o exemplo, pode baixá-lo [aqui](#).

Você poderá ouvi-lo quantas vezes quiser. Depois de ter escutado o exemplo inteiro, responda as seguintes perguntas, por favor:

2.1 Tem algum aspecto na fala da pessoa que chamou a sua atenção?

2.2 Como você descreveria a língua que você está ouvindo?

2.3 Como você classificaria o falante? Por favor, escolha a opção entre os dois polos que lhe pareça mais adequada.

pobre     rico

preguiçoso     trabalhador

ignorante     bem informado

atraente     pouco atraente

sociável     pouco sociável

simpático     pouco simpático

2.4 De onde vem o falante?

São Paulo (capital) Onde?

Interior de São Paulo (estado) Onde?

Fora de São Paulo (estado) Onde?

Fora do Brasil Onde?

2.5 Onde mora o falante?

Bom Retiro/Centro

Vila Mariana/Zona Sul

Freguesia do Ó/Zona Norte

Brás/Zona Leste

Pinheiros/Zona Oeste

Um outro bairro de São Paulo (capital) Onde?

Interior de São Paulo (estado) Onde?

Fora de São Paulo (estado) Onde?

Fora do Brasil Onde?

Por favor, explique em poucas palavras, por que você escolheu justamente esse(s) lugar(es):

Para ouvir o último exemplo, clique aqui [ou aqui](#). Para interromper clique mais uma vez no botão. Se você não conseguir ouvir o exemplo, pode baixá-lo [aqui](#). Você poderá ouvi-lo quantas vezes quiser. Depois de ter escutado o exemplo inteiro, responda as seguintes perguntas, por favor:

3.1 Tem algum aspecto na fala da pessoa que chamou a sua atenção?

3.2 Como você descreveria a língua que você está ouvindo?

3.3 Como você classificaria o falante? Por favor, escolha a opção entre os dois polos que lhe pareça mais adequada.

- pobre     rico
- preguiçoso     trabalhador
- ignorante     bem informado
- atraente     pouco atraente
- sociável     pouco sociável
- simpático     pouco simpático

3.4 De onde vem o falante?

- São Paulo (capital) Onde?
- Interior de São Paulo (estado) Onde?
- Fora de São Paulo (estado) Onde?
- Fora do Brasil Onde?

3.5 Onde mora o falante?

- Bom Retiro Centro
- Vila Mariana Zona Sul
- Freguesia do Ó Zona Norte
- Brás Zona Leste
- Pinheiros Zona Oeste
- Um outro bairro de São Paulo (capital) Onde?
- Interior de São Paulo (estado) Onde?
- Fora de São Paulo (estado) Onde?
- Fora do Brasil Onde?

Por favor, explique em poucas palavras, por que você escolheu justamente esse(s) lugar(es):

4. O que você acha sobre a recente imigração para São Paulo, particularmente dos países vizinhos do Brasil? Você tem experiências próprias com imigrantes?

5. Dentro da sociedade Paulistana, existe discriminação contra imigrante?

Por favor, explique em poucas palavras, por que você escolheu justamente esse(s) lugar(es):

4. O que você acha sobre a recente imigração para São Paulo, particularmente dos países vizinhos do Brasil? Você tem experiências próprias com imigrantes?

5. Dentro da sociedade Paulistana, existe discriminação contra imigrante?

---

Agora faltam só umas informações estatísticas:

Você é  masculino  feminino

A sua idade?

A sua profissão?

Você mora em qual bairro?

Você conhece outras línguas além do português?

Sim  Não

Quais?

---