Urban youth in Rio de Janeiro
Contemporary linguistic variation in Brazilian Portuguese

Christian P. Thompson & Sky Onosso

Introduction
Our poster reports on an ongoing project investigating linguistic variation in school children in Rio de Janeiro (2015 metro pop. 12,280,702), attended mostly by children from lower socioeconomic status households.
For the initial stage, we have focused on the creation of the first corpus of spontaneous speech from students on the island-district of Isha do Governo, 152 students (32 male, 70 female) between the ages 11–16 from two schools on the island took part in Labovian sociolinguistic interviews in 2015 & 2016.
We hope that our project can help illuminate some of the unique linguistic characteristics of this speech community, which is not generally given much social prestige within the wider Brazilian Portuguese community.

Geographic & Social Context
Rio de Janeiro is the former capital of Brazil, its 2nd largest metropolitan centre, and home of the largest school district in Latin America (pop. 654,454). Our study focuses on middle-school students at two public schools on Ilha do Governo (Governor's Island) (pop. 400,000), which has restricted connections with the mainland.

The primary data source (75% of all interviewees, and the six currently transcribed interviews) is a newly built (2015) athletics-focused school; students must compete for entrance and maintain high performance levels & grades. Part of a special government project, the school has many athletics-focused facilities including training rooms, dormitories, male and female outdoor pools.
The second school is a typical public school which lacks any special amenities or resources. Despite these differences, the children at both schools, which are only 5 km distance from each other, come from similar or identical neighbourhoods.

Research Context
Here we briefly summarize the demographics and data types of two public corpora covering Rio de Janeiro. Our corpus aims to contribute to research in this speech community.
NURC (Projeto da Norma Urbana Oral Culta do Rio de Janeiro)
Era of data collection: 1970s, 1990s
Participants: 152 educated adults (1970s)
In the 1990s, additional participants were added and the original participants were re-interviewed.
PEUL (Programa de Estudos sobre o Uso da Língua)
Carico (Rio) data collected in 2005.
Oral interviews were conducted in first two periods; newspaper text data was added covering 2000-2004.

Study Participants
Students and staff at both schools were very receptive to our study, leading to high recruitment rates; of the students who we were introduced to at the two schools, 33% & 57% (respectively) participated.

Participants:
Interviews were conducted in pairs (or as triads in a few cases). Subjects selected their interview partners, allowing natural and spontaneous conversation to emerge. The researcher (author #1) is a native of Isha do Governo, further facilitating naturalness of speech.

Data:
Both video and audio data were captured. The full corpus comprises 46 interviews which range between 45-60 minutes in length. In total, approximately 40 hours of data. 6 full interviews have been transcribed to date, documenting the speech of 12 female subjects ages 12-14.

Age of Participants (all)

N=108 instances of quotative constructions occur in our transcribed data. There are four observed patterns which follow Cameron (1998), with one addition (Type III).

Type I: Direct reported speech. The subject and a verb of speech precede the quotation. This type constitutes 34% of quotatives. Of the six speech verbs which occur, falé ‘speak, talk’ accounts for the vast majority of tokens, at 81%.

E eu falé, Môe, cãdê ou Max? [I] got home, ‘Mom, where’s Max?’ [215]

Type II: ‘Rare noun phrase’ (Cameron 1998). A noun phrase (pronoun or full NP) without a verb precedes the quotation. Cameron describes a very similar pattern for Puerto Rican Spanish, although it is unknown how widespread this pattern is in other varieties of Spanish. This type constitutes the majority of quotatives in our data, at 44%.

Elo, ‘Calm down.’ She, ‘Calm down.’ [115]

Type III: Scene-setting. A sentence or clause precedes the quotation but does not directly indicate the speaker or include a speech verb, typically serving instead to set the scene in some way. This type is the least common, constituting 9% of quotatives.

Cheguei em casa, ‘Mom, cãdê ou Max?’ (I) got home, ‘Mom, where’s Max?’ [215]

Type IV: ‘Test-rerouting’ (Clark & Gerrig 1990). The quoted speech is part of a sequence of uninterrupted consecutive quotations, absent any introduction. This type is relatively infrequent, constituting 12% of quotatives in our data.

Coda lenition (debeubalization)
Carico (Rio denomyn) speech is well-known for “chádo,” a process of backlingual/alatalization of coronal cauda fricatives, e.g. [r] → [ɾ]/[ɹ] and affrication of [tʃ] (Whitney & DeBattista 1998). This latter process is also known for wide variation in the phonemic quality of /r/ which includes both backed and debuccalized (non-oral) varieties, e.g. [ɾ] and [ɹ]/[ɹ].

Our data documents instances of both backlingual and debuccalization of coronal cauda fricatives (which can occur in [ɾ]/[ɹ]) and affrication of [tʃ]. This is also known for wide variation in the phonemic quality of /r/ which includes both backed and debuccalized (non-oral) varieties, e.g. [ɾ] and [ɹ]/[ɹ].

A spectral slice comparison of the fricative portion of tokens of mesmo [meimu] (left) vs. [meimu] (right) speaker 145 reveals strong acoustic differences between the two variants; the debuccalized version is devoid of upper frequency information.

Future Directions
Future work in our corpus will run in two main directions - further analyses of the speech data, and pilot work for the wider project. This latter will aim to investigate the role of such factors as speaker sex, sociocultural status (on the basis of parental employment information) and geographic location (children who live in similar or different neighbourhoods). Additionally, all transcripts at present are from the population at just one of the two schools where we collected data. Inclusion of transcripts from the second school will permit us to further investigate differences in usage which may be associated with the different social environments at each school.

Acknowlegements
This study was designed and carried out under the supervision of Dr. Alexandre Stârchi, University of Victoria, whom we wish to thank for her valuable feedback and advice.
Our data was collected under the approval of the University of Veterinary Medical/Protocol 2020/135; 2014/8, the University of Veterinary Medical and the Municipal Secretary of Education for the Prefecture of the City of Rio de Janeiro, Protocol 407/004-640/2015.
We are extremely grateful to the City of Rio de Janeiro for granting us permission to carry out this study, and the students and staff at Escola Municipal Jornalista Orlando Dantas and Gislando Experimental Olímpica Nelson Prudêncio for their participation.
We hope that our research contributes positively towards the community who graciously allowed us access, and that our work reflects well on their openness and kindness.
We are open to any comments, suggestions, or other feedback from the sociolinguistic community.

References