

Documentation and User Guide to the Portuguese Regional Accent Databank (RADbank)

Simone Ashby

Institute for Theoretical and Computational Linguistics (ILTEC)

Translated by Sílvia Barbosa and Catarina Silva

Lisbon, Portugal

Table of Contents

1	RADbank.....	3
2	How to use the Portuguese RADbank	3
2.1	Browsing by country	3
2.2	Browsing by informant.....	7
2.3	Browsing by word/phrase	7
	Acknowledgments.....	9

1 RADbank

The Portuguese RADbank is available to the research community in the form of a free, searchable, online databank for: accessing recordings, testing the results of different speech processing systems, conducting empirical analyses across multiple Portuguese accents, monitoring contact language effects, and facilitating the entry of lesser documented regional variants into the digital domain.

The RADbank contains all of the read-speech audio recordings collected as part of the LUPo project, along with their orthographic transcriptions; recordings of sociolinguistic interviews; and summary information about each of the informants recorded, their linguistic background, and some of the linguistic attitudes they expressed during the interview.

2 How to use the Portuguese RADbank

The Portuguese RADbank is available on the Web at <http://194.117.2.11/radbank/>¹, and has dual English and Portuguese interfaces. Users can browse this resource by selecting a country, an informant, or by choosing among the words and phrases collected as part of LUPo's read speech corpus.

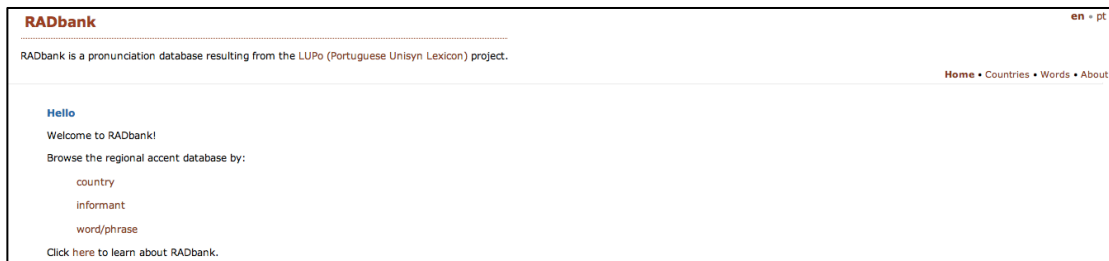


Figure 1: RADbank homepage.

2.1 Browsing by country

Within a country, the user can browse the data by informant or regional variety. For example, clicking on 'informants' under Cape Verde results in a display of all of the informants recorded in that country, listed by informant ID and town.

¹ Note that the Portuguese RADbank can also be accessed via the Portal da Língua Portuguesa website (www.portaldalinguaportuguesa.org) by clicking 'RADbank' in the right-hand menu under 'Consultar'.

Select an informant	
Country: Cape Verde	
Informant 072	Lameirão, São Vicente
Informant 033	Mindelo, São Vicente
Informant 035	Mindelo, São Vicente
Informant 074	Mindelo, São Vicente
Informant 075	Mindelo, São Vicente
Informant 078	Mindelo, São Vicente
Informant 085	Mindelo, São Vicente
Informant 089	Mosteiros, Fogo
Informant 084	Paul, Santo Antão
Informant 079	Porto Novo, Santo Antão
Informant 086	Praia Branca, São Nicolau
Informant 100	Praia, Santiago
Informant 102	Praia, Santiago
Informant 007	Praia, Santiago
Informant 090	Praia, Santiago
Informant 094	Praia, Santiago
Informant 097	Praia, Santiago
Informant 098	Praia, Santiago
Informant 099	Praia, Santiago
Informant 080	Ribeira Brava, Santo Antão
Informant 081	Ribeira Brava, São Nicolau
Informant 082	Ribeira Brava, São Nicolau
Informant 083	Ribeira Brava, São Nicolau
Informant 087	Ribeira Brava, São Nicolau
Informant 034	Ribeira Grande, Santo Antão
Informant 069	Ribeira Grande, Santo Antão
Informant 070	Ribeira Grande, Santo Antão
Informant 071	Ribeira Grande, Santo Antão
Informant 076	Ribeira Grande, Santo Antão
Informant 077	Ribeira Grande, Santo Antão

Figure 2: Listing of the informants recorded in Cape Verde.

Clicking on one of the above informant IDs results in a display of summary information about the recording, along with options for hearing the interview or read-speech portion of the recording.

Informant 81	
Age:	18
Gender:	Feminino
Education:	Frequenta o 1º ano de direito
Nationality:	cabo-verdiana (Ribeira Brava, S. Nicolau)
Recording data:	01/12/2011, Mindelo, Cabo Verde
Interviewer:	Simone Ashby / José Pedro Ferreira
Technical data:	Prompts: Marantz PMD 670, PCM, mono, int. mic., 44.1 kHz; INQ.: Zoom HD2, PCM, 44.1 kHz, int. mic. (bi-directional, stereo)
Select a type of resource	
interview	
read speech	

Figure 3: Informant display.

Clicking on ‘interview’ results in a display containing information about the informant, a link to the audio file containing the complete interview (with the informant’s name extracted), and a textual summary of the key sociolinguistic questions asked and the responses given by the informant.

Interview

Informant 81

Age: 18
 Gender: Feminino
 Education: Frequenta o 1º ano de direito
 Nationality: cabo-verdiana (Ribeira Brava, S. Nicolau)
 Recording data: 01/12/2011, Mindelo, Cabo Verde
 Interviewer: Simone Ashby / José Pedro Ferreira
 Technical data: Prompts: Marantz PMD 670, PCM, mono, int. mic., 44.1 kHz; INQ.: Zoom HD2, PCM, 44.1 kHz, int. mic. (bi-directional, stereo)

[Listen to the full interview](#)

Questions and responses

Perfil sociolinguístico

1. Que línguas fala?
 crioulo de cv, português, um pouco de espanhol

2. Com que idade começou a falar cada uma delas?
 espanhol: há 10 anos (curso); crioulo de Cabo Verde: língua materna

3. Com que idade começou a falar português?
 Com 6 anos.

4. Que língua acha mais fácil de aprender a falar? Porquê?
 crioulo: língua materna; imersão.

5. Que línguas fala em casa?
 crioulo.

6. Que línguas fala no trabalho?

Figure 4: Sample interview display, including a link to the audio file containing the complete interview.

Alternatively, by clicking ‘read speech’, a complete list of words and phrases uttered in isolation is presented alphabetically for the given informant. Here, each word or phrase is linked to an audio file, enabling the user to browse the actual recordings.

Listen to read speech	
Informant 81: Ribeira Brava, São Nicolau, Cape Verde	
listen	B03_217 abasteciam
listen	B02_131 abundante
listen	B03_203 accionam
listen	B02_105 acena
listen	B05_340 actividade
listen	B06_400 actual
listen	B02_085 acumulação
listen	B03_180 acção
listen	B06_376 adaptar
listen	B01_066 adopção
listen	B06_383 adquirida
listen	B06_382 adversário
listen	B08_589 agora bebemos
listen	B08_569 agora fugimos
listen	B06_401 agricultura
listen	B06_389 aia
listen	B08_591 Aldeia Galega
listen	B05_351 algoz
listen	B05_345 Almada
listen	B03_166 ama-lo
listen	B06_373 amara
listen	B05_357 amará
listen	B03_159 ambidestro
listen	B02_115 americano

Figure 5: Sample read speech display, with links to the audio files for each recorded word and phrase.

Note that clicking on one of the words or phrases displayed in the read speech page results in a display of that word or phrase, along with audio links for the complete set of informants recorded, listed by informant ID, country, and town.

listen	095 Cape Verde	Santa Catarina, Fogo
listen	101 Cape Verde	Santa Catarina, Fogo
listen	006 Cape Verde	Santa Cruz, Santiago
listen	092 Cape Verde	São Domingos, Santiago
listen	008 Cape Verde	São Domingos
listen	091 Cape Verde	São Filipe, Fogo
listen	096 Cape Verde	São Filipe, Fogo
listen	088 Cape Verde	São Salvador do Mundo, Santiago
listen	073 Cape Verde	São Vicente
listen	018 Guiné-Bissau	Bissau
listen	051 Mozambique	Beira, Sofala
listen	065 Mozambique	Beira, Sofala
listen	068 Mozambique	Beira, Sofala
listen	060 Mozambique	Cambine, Mafola, Inhambane
listen	044 Mozambique	Carropeira, Nampula
listen	058 Mozambique	Cumbane, Inhambane
listen	059 Mozambique	Dombe, Sussundenga, Manica
listen	053 Mozambique	Gongola, Manica
listen	017 Mozambique	Inhambane

Figure 6: Sample word/phrase display, with links to the audio files for the same word or phrase, across regional varieties.

Note that in the above word/phrase display, the user can choose to browse a different informant or country by simply clicking the desired informant ID or country name.

2.2 Browsing by informant

Browsing by informant enables the user to see the complete list of informants, listed by informant ID, country, and town.

Select an informant		
Informant 011	Angola	Luanda
Informant 015	Brazil	Alfenas, Minas Gerais
Informant 014	Brazil	Belo Horizonte, Minas Gerais
Informant 004	Brazil	Belém
Informant 021	Brazil	Campinas
Informant 005	Brazil	Pará
Informant 012	Brazil	Piracicaba, São Paulo
Informant 001	Brazil	Rio de Janeiro
Informant 022	Brazil	Rio de Janeiro
Informant 023	Brazil	Rio de Janeiro
Informant 106	Brazil	São Paulo
Informant 003	Brazil	São Paulo
Informant 072	Cape Verde	Lameirão, São Vicente
Informant 033	Cape Verde	Mindelo, São Vicente
Informant 035	Cape Verde	Mindelo, São Vicente
Informant 074	Cape Verde	Mindelo, São Vicente
Informant 075	Cape Verde	Mindelo, São Vicente
Informant 078	Cape Verde	Mindelo, São Vicente
Informant 085	Cape Verde	Mindelo, São Vicente
Informant 089	Cape Verde	Mosteiros, Fogo
Informant 084	Cape Verde	Paul, Santo Antão
Informant 079	Cape Verde	Porto Novo, Santo Antão
Informant 086	Cape Verde	Praia Branca, São Nicolau
Informant 100	Cape Verde	Praia, Santiago
Informant 102	Cape Verde	Praia, Santiago

Figure 7: Browsing the RADbank by informant.

As described in Section 2.1, clicking on an informant ID results in the display of summary information about the recording, along with options for hearing the interview or read-speech portion of the recording..

2.3 Browsing by word/phrase

Browsing by word or phrase enables the user to see a complete list of the words and phrases recorded.

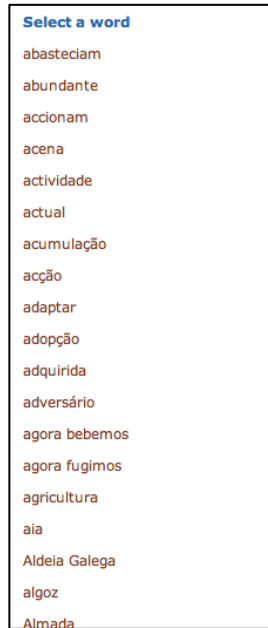


Figure 8: Browsing the RADbank by word/phrase.

As described in Section 2.1, each word or phrase is clickable and links to a display of the word/phrase, along with audio links for the complete set of informants recorded, listed by informant ID, country, and town.

Acknowledgments

The author gratefully acknowledges the support of the Portuguese Science and Technology Foundation (Ref: **PTDC/CLE-LIN/100335/2008**), and the cooperation of Susan Fitt, whose development of the original English Unisyn Lexicon is the inspiration for this work.

This work would not have been possible without the immense contributions of LUPo's research team members, assistants, consultants, and collaborators. The author particularly wishes to thank: Silvia Brandão and João Antônio de Moraes (Federal University of Rio de Janeiro); Mário Eduardo Viaro (University of São Paulo); Sílvia Barbosa, Ana Catarina Silva, and José Pedro Ferreira (ILTEC); Maarten Janssen (Universitat Pompeu Fabra, Barcelona); Paulino Fumo (Eduardo Mondlane University); Celeste Rodrigues (University of Lisbon); Davi Albuquerque (University of Brasília); Aline Ponciano, Caio Castro, and Priscila Francisca dos Santos (Federal University of Rio de Janeiro); Juliana Fonte (Federal University of São Paulo); Joana Santos, Neuza Campaniço and Rita Monteiro (formerly of ILTEC); and Lauren Phillips (University of California Berkeley).

The author is also extremely grateful to Inês Machungo and colleagues at the Eduardo Mondlane University; Dora Pires (University of Cape Verde, Mindelo); Adelaide Monteiro (University of Cape Verde, Praia and Cultural Ministry of Cape Verde); João Neves (Camões Portuguese Cultural Center, Praia); and Manuel de Pina (University of Cape Verde, Praia). Sincere thanks go to all of the informants who participated in this study.