Dynamic Correspondences:

An Object-Oriented Approach to Tracking Sound Reconstructions

Tyler Peterson – University of British Columbia (Canada)
Gessiane Picanço – Universidade Federal do Pará (Brazil)

Introduction:

- Databases: have transitioned from searchable repositories of tokens, to a tool that can organize vast amounts of data, and execute complex statistical functions and queries on the data it stores.
- Many options: commercial databases (MS Access/Excel, Filemaker Pro; OpenOffice Calc and Base); Academic/Research databases: ShoeBox, various applications from the MPI.
- How can we put these innovations to work on Comparative/Historical linguistics?

The BDEC-T Database

- *Base de Dados para Estudos Comparativos – Tupí* (Database for Comparative Studies – Tupí)
- Implemented within MS Access 2003, providing the user an interface for entering language data, which is then externally linked to tab-delimited text files.
- Provides a visual interface for entering segmented data which can then be cross-tabulated in the form of *pivot tables* -- a recent innovation in the implementation SQL queries.
- Why? We believe cross-tabulation and pivot tables are an effective tool for tracking several one-to-one and one-to-many correspondences simultaneously across several languages; the ability to dynamically survey the language-internal distribution of segments and their features.
- What are cross-tabulations? They track the joint distribution of two variables – for our purposes, these variables will be phonemes.

The Data: The Tupí langauges

- The Tupí stock of language families is concentrated in the Amazon river basin of Brazil, comprising of 10 language families totaling approximately 64 languages.
- At present, the BDEC-T for the Tupí stock contains a glossary of 813 words and up to 3,785 entries distributed across 15 Tupían languages.
- Approximately 18% of this 813-word list appear to have cognates in the majority of languages entered so far, and which can be used as reference for a reliable set of robust cognates across the entire Tupí stock.

Promising results

- BDEC-T has predicted larger sets of correspondences than those posited by previous, manual methods. ⟶
- BDEC-T has challenged previous claims: A previous analysis mistakenly posited the correspondence set *∅/? for both Mundurukú and Mekéns. The segmentation function in BDECT-T tracked a more suitable correspondence * ∅ /? for Mekéns but *(C)VPV/(C)V for Mundurukú. ⟶

Rodrigues		BDEC-T		Rodrigues		BDEC-T	
P-T	Mund.	P-T	Mund.	P-T	Mund.	P-T	Mund.
*p	p	*p	p	*ʧ	ʃ	*ʧ	ʃ
			∅		ʧ		ʧ
			ps		ʤ		ʤ
			p/b				
*p^{ʔ}	b	*p^{ʔ}	b	*ʧ^{ʔ}	t	*ʧ^{ʔ}	t
			p		d		d

	S1	S2	S3	S4	S5	S6
Proto-Tupí: *upiʔa	∅	u	p	i	ʔ	a
Mundurukú: topsa̲	t	o	ps	∅	∅	a̲
Mekéns: upia	∅	u	p	i	∅	a

Table 5: *(C)V?V corresponding with (C)V̲

*k	k	*k	k	*k^{ʔ}	ʔ	*k^{ʔ}	ʔ
			ʃ				

Table 4: The correspondence sets as proposed by Rodrigues (1995) compared with those generated by the BDEC-T.

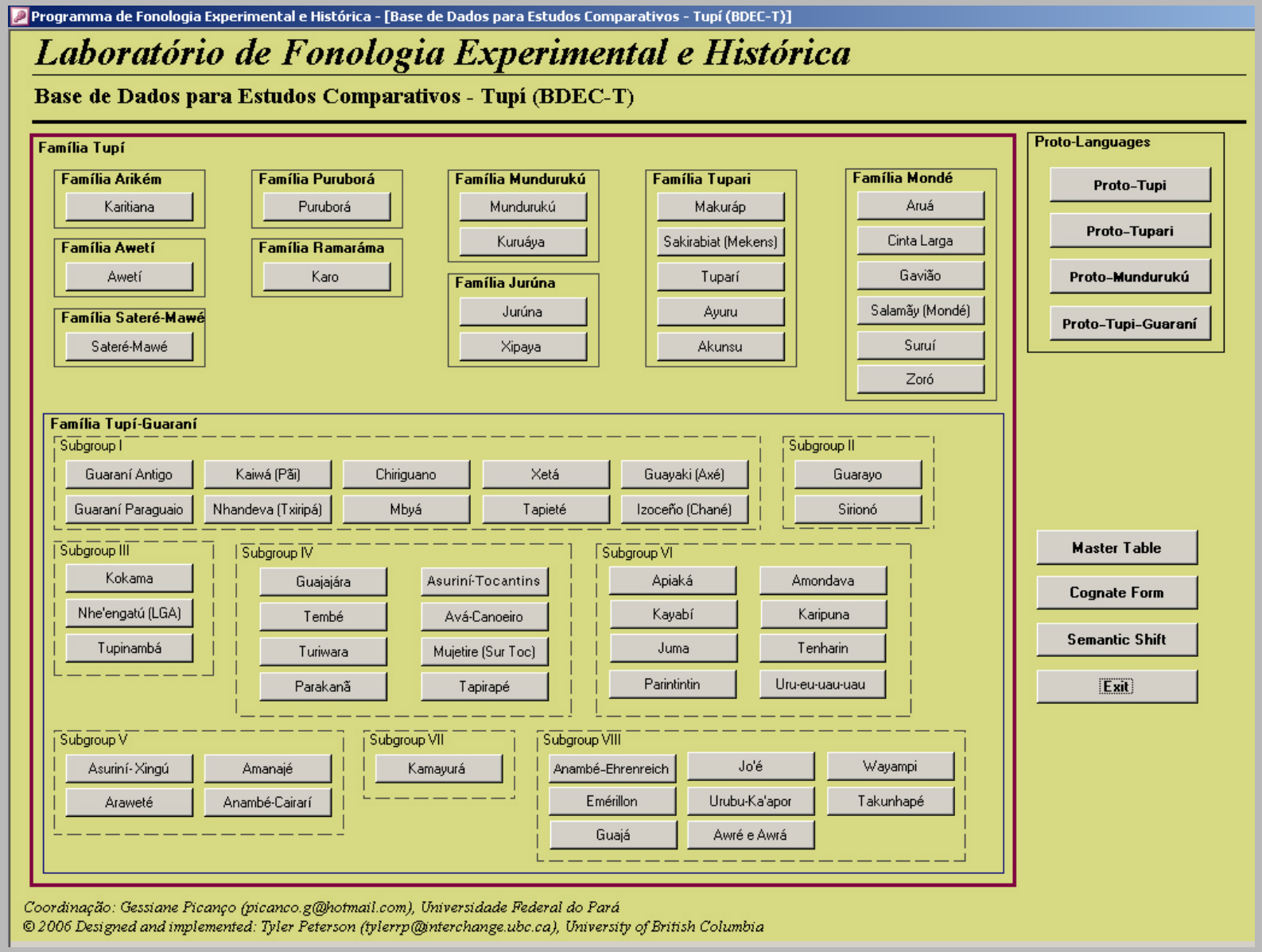
Objectives and Plan

- To demonstrate the core functions of BDEC-T: segmentation and correspondence analysis utilizing cross-tabulation and pivot tables.
- To present a powerful yet practical use of an ‘off-the-shelf’ database application that can be implemented with little or no previous programming experience.

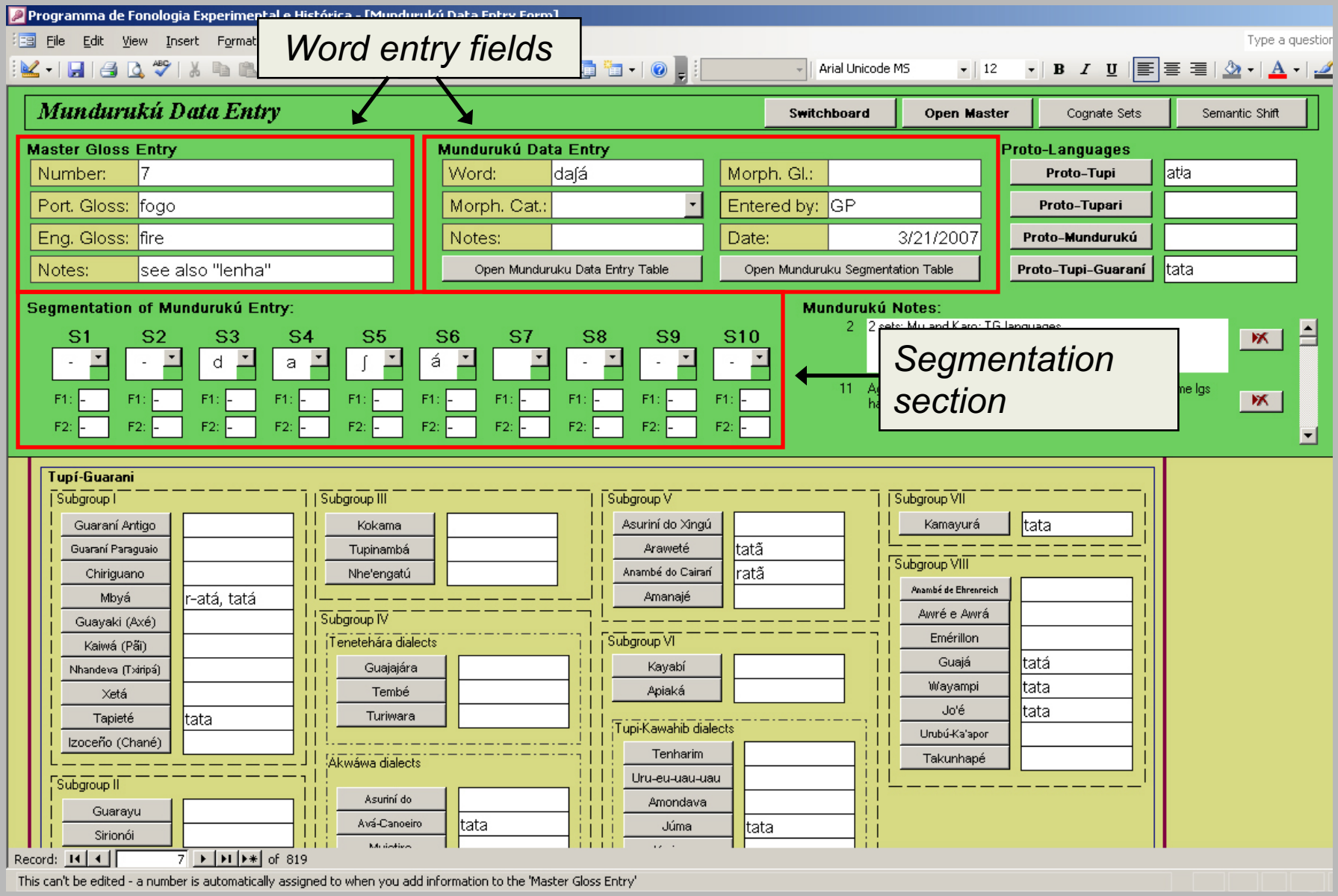
Procedure:

I. Data entry
II. Segmentation
III. Queries
IV. Analysis

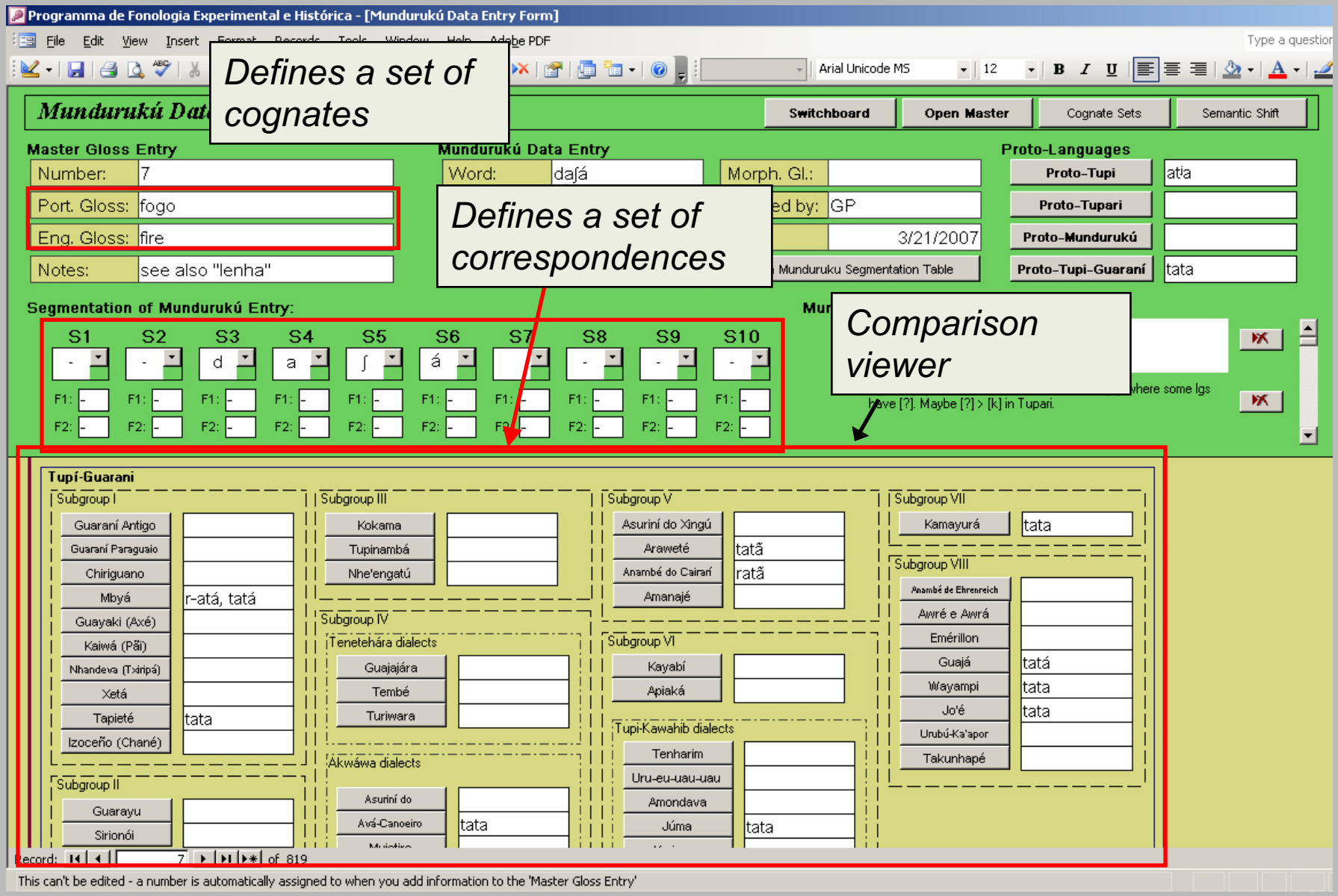
Master Switchboard:



Data Entry Form:



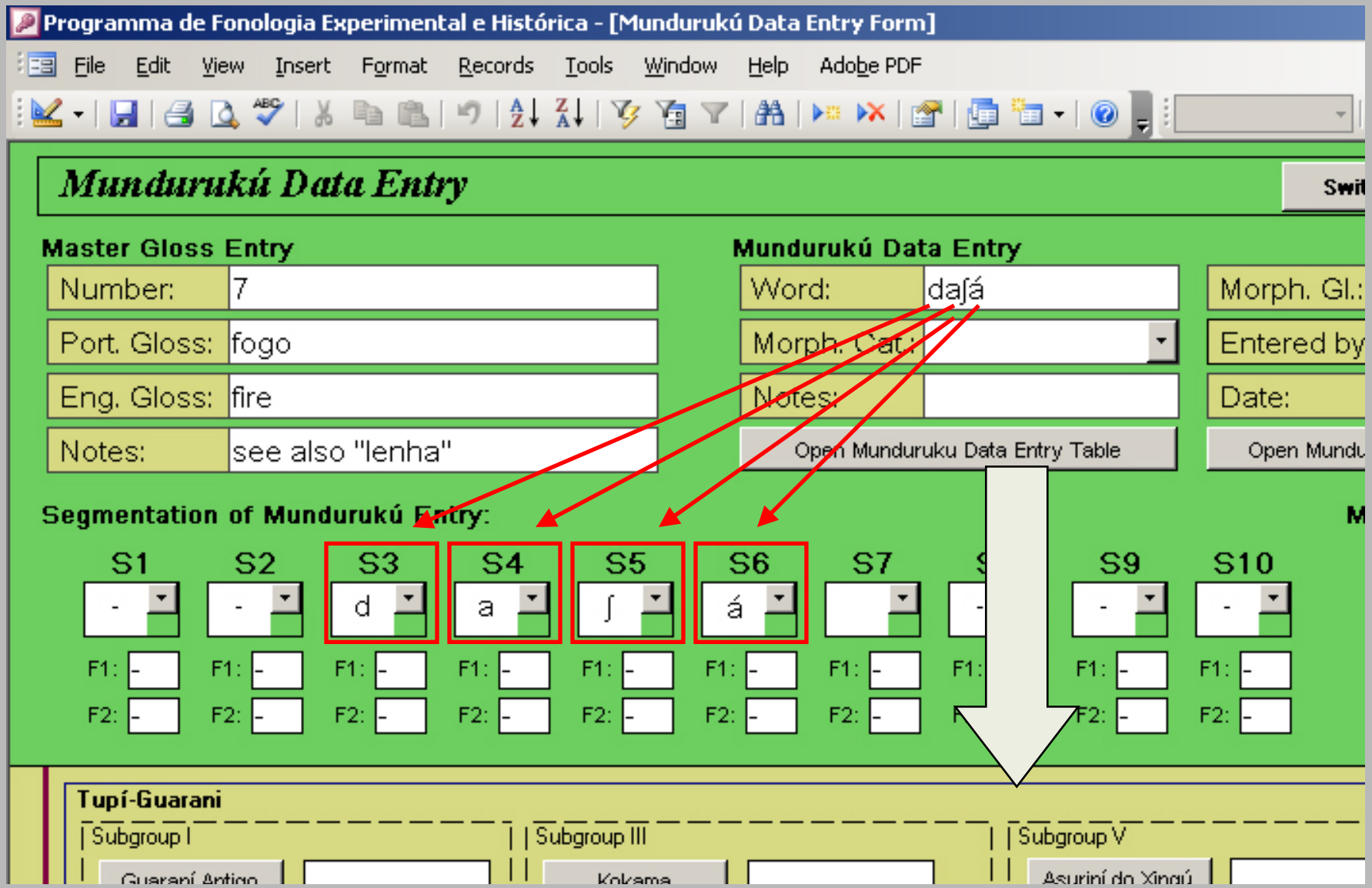
Data Entry Form:



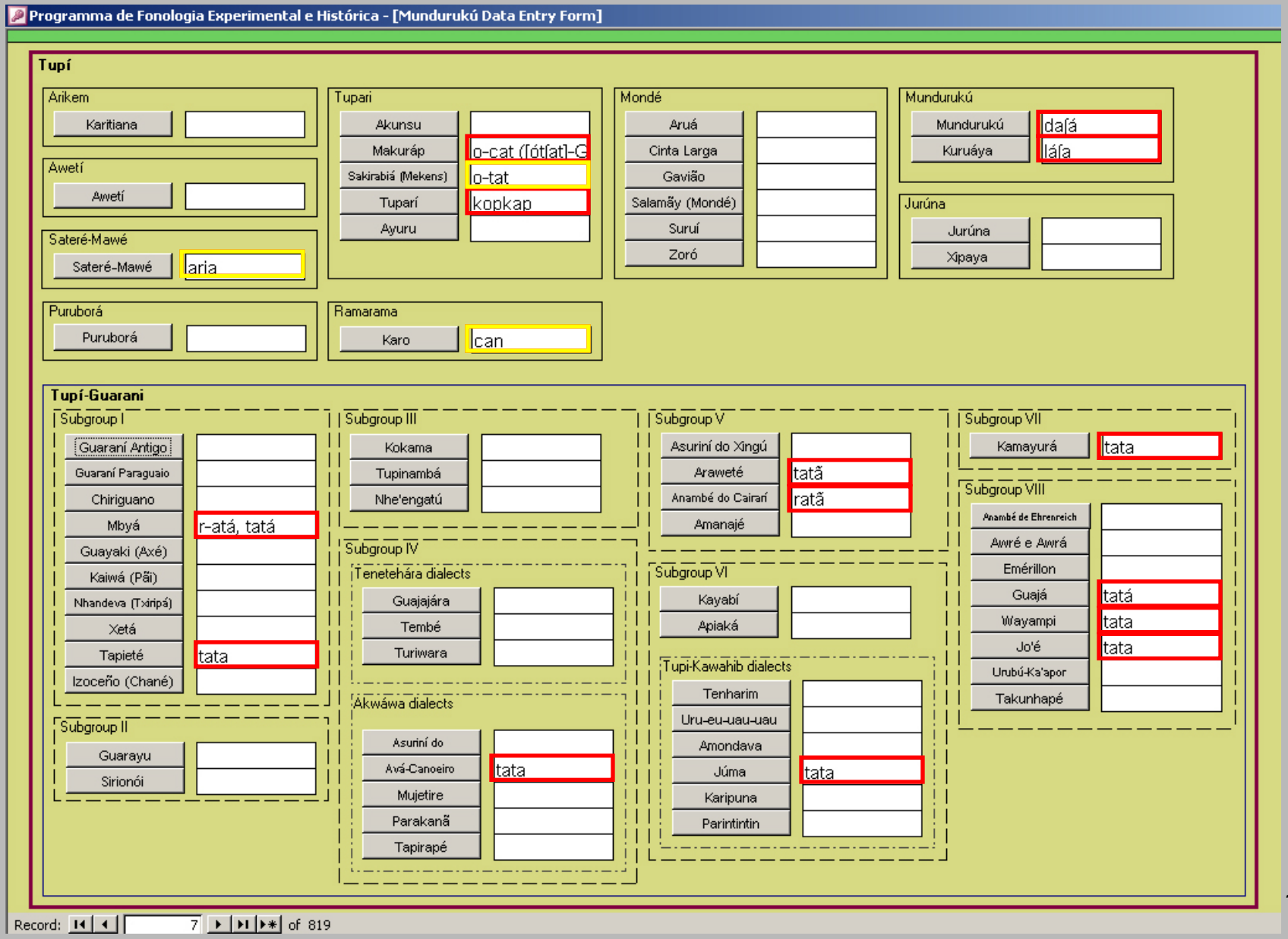
Forms and Data storage

- Each of the 64 languages and 4 proto-languages in BDEC-T is associated with its own data entry form.
- The data entered in these forms is stored in a master table where all of the languages are represented as columns.
- Glosses are the rows, where each gloss is assigned a unique, autogenerated number in the master record when it is entered into the database.
- This serves as the primary key for all the translations of that gloss across all of the languages.

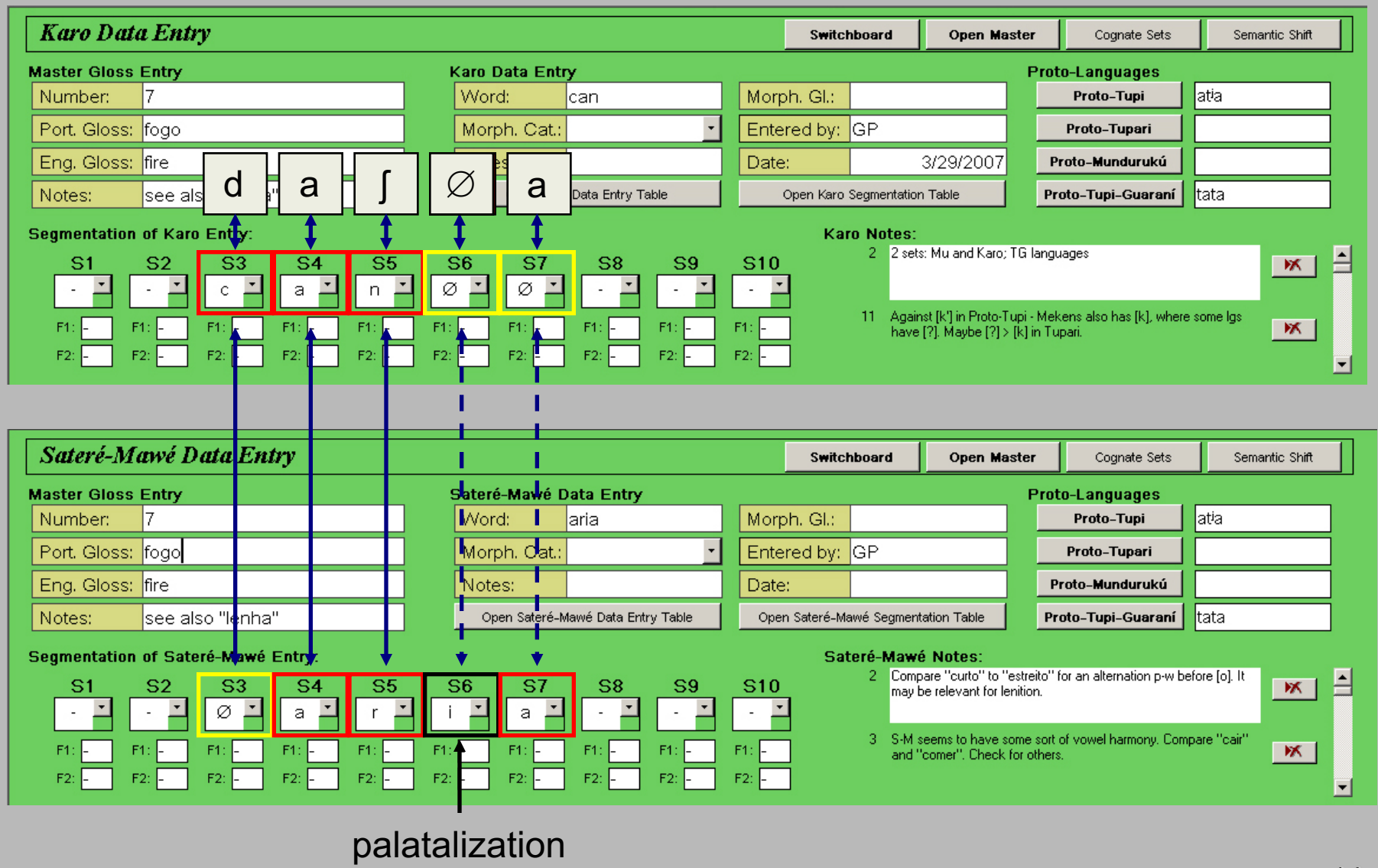
Segmentation:

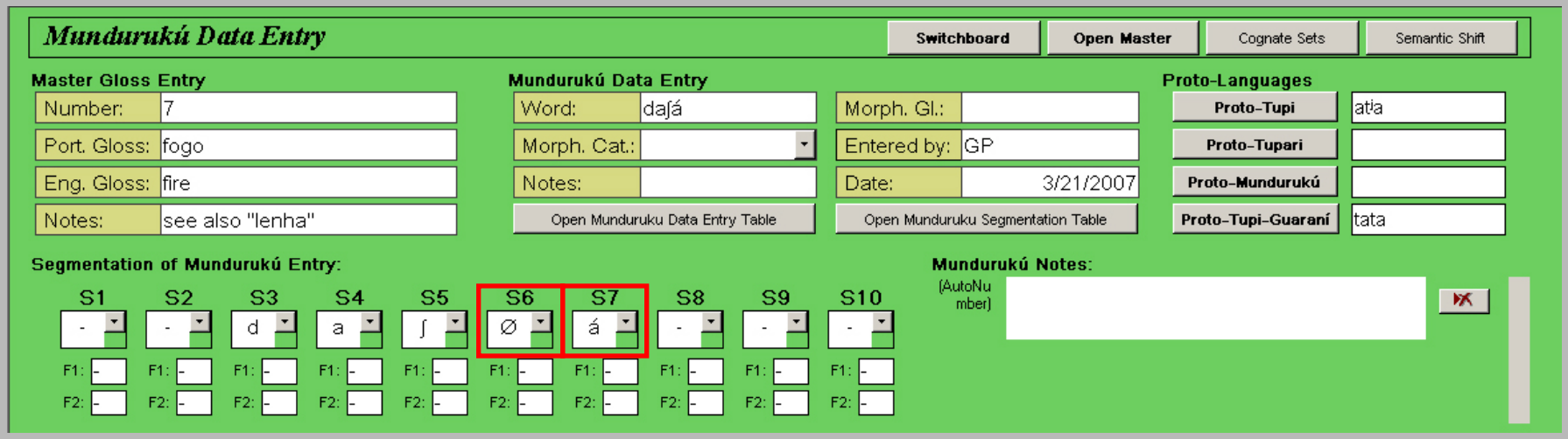


Segmentation and comparison:



Aligning correspondences:





	S2	S3	S4	S5	S6	S7
Mundurukú		d	a	ʃ	∅	a
Karo		c	a	n	∅	∅
Sateré Mawé		∅	a	r	i	a
Makuráp	o	c	a	t	∅	∅
Sakirabiá	o	t	a	t	∅	∅

Segmentation slot	S1	S2	S3	S4	S5
Avá-Canoeiro	∅	i	∅	a	p
Guajá	∅	u	ʔ	i	∅
Mbyá	h –	u	ʔ	i	∅
Kamayurá	h	i	ʔ	i	p

Table 2: Segmentation of *'arrow'*

Segmentation slot	S1	S2	S3	S4	S5
Avá-Canoeiro	∅	∅	i	t	i
Guajá	∅	w	i	t	i
Araweté	i	w	i	t	i

Table 1: Segmentation of *'wind'*

What can be done with this data?

- Cross-tabulations of the segment data recorded in each column can be used to:
  - 1) monitor the language-internal distribution of segments; and
  - 2) track correspondences between languages for a particular cognate or segment slot.
- This is visually implemented through *pivot tables*.

Cross-tabulation and Pivot tables

- Cross-tabulation: displays the joint distribution of two or more variables, usually presented as a contingency table which plots the distribution of two or more variables simultaneously.
- This allows us to examine frequencies of observations that belong to specific categories on more than one variable. By examining these frequencies, we can identify relations between crosstabulated variables.
- Access 2003 includes a graphical implementation of SQL statements in the form of cross tabulations, or pivot tables.
- Pivot tables can be described as an ‘object-oriented’ representation of SQL statements: columns of data are treated as objects, which allow the user to create multidimensional views of the data by ‘dragging and dropping’ columns into various sorting arrangements.
- Different levels of detail or organize data can be used by dragging the fields and items or by showing and hiding items in the drop-down lists for the fields.

I. Language-internal distributions

- Phonotactic restrictions may, in many cases, be gaps left behind by historical changes.
- Pivot tables allow the analyst is able to easily monitor and track distributional gaps or contrasts and so provide a more systematic diachronic analysis.

Sample Case study: Mundurukú

i. Root-initial distribution of phonemes.

ii. Distribution of root-intial labials.

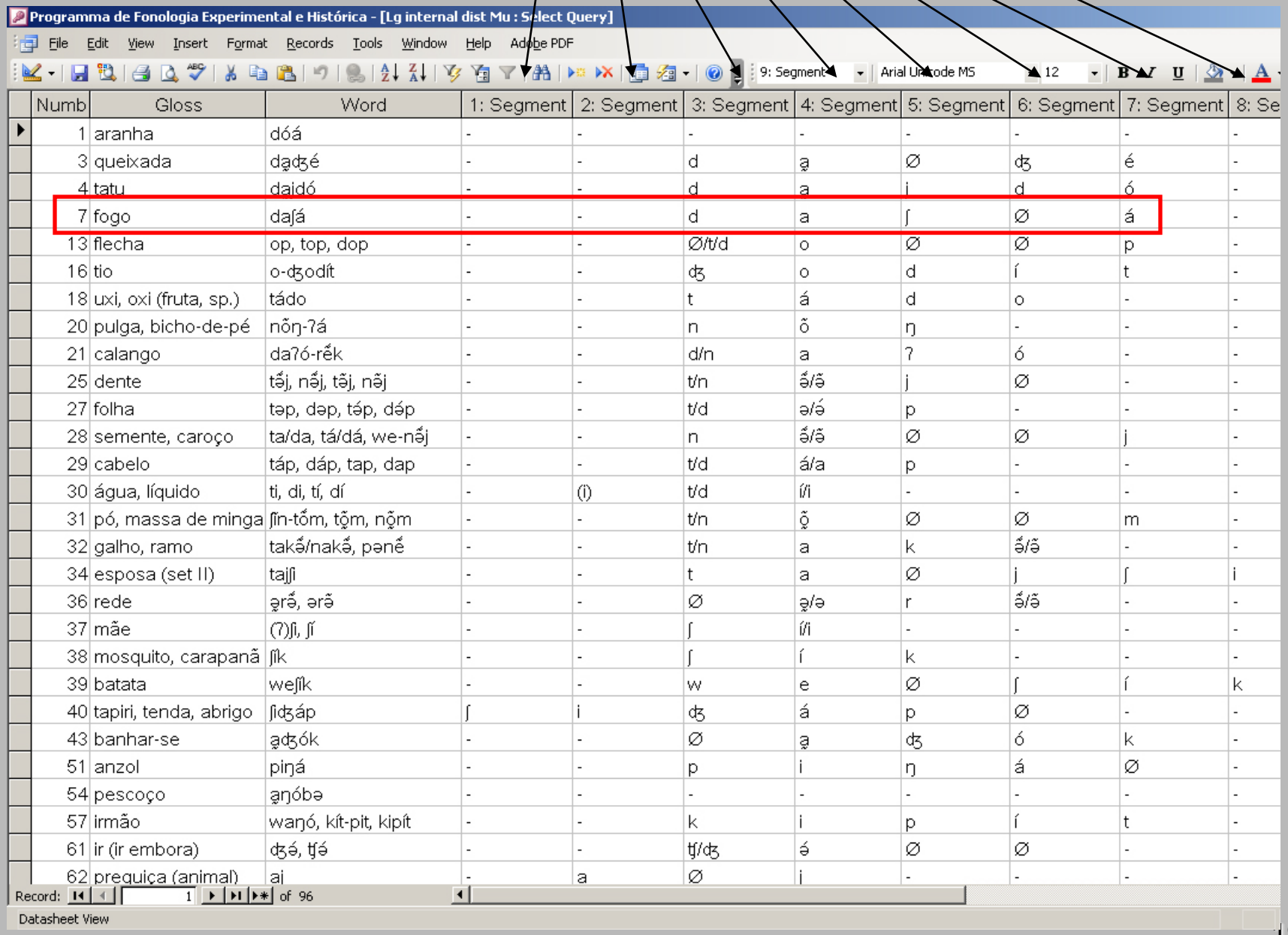
iii. Root-initial consonants and their following vowels.

iv. Root-initial distribution of creaky and nasal vowels.

v. Root-initial syllables: vowels and what consonants precede them.

Munduruku data table (entered on the Data entry form)

Language Segmentation Form → S1 S2 S3 S4 S5 S6 S7 S8 S9 S10



Programma de Fonologia Experimental e Histórica - [Lg internal dist Mu : Select Query]

File Edit View Insert Format Records Tools Window Help Adobe PDF

9: Segment Arial Unicode MS 12

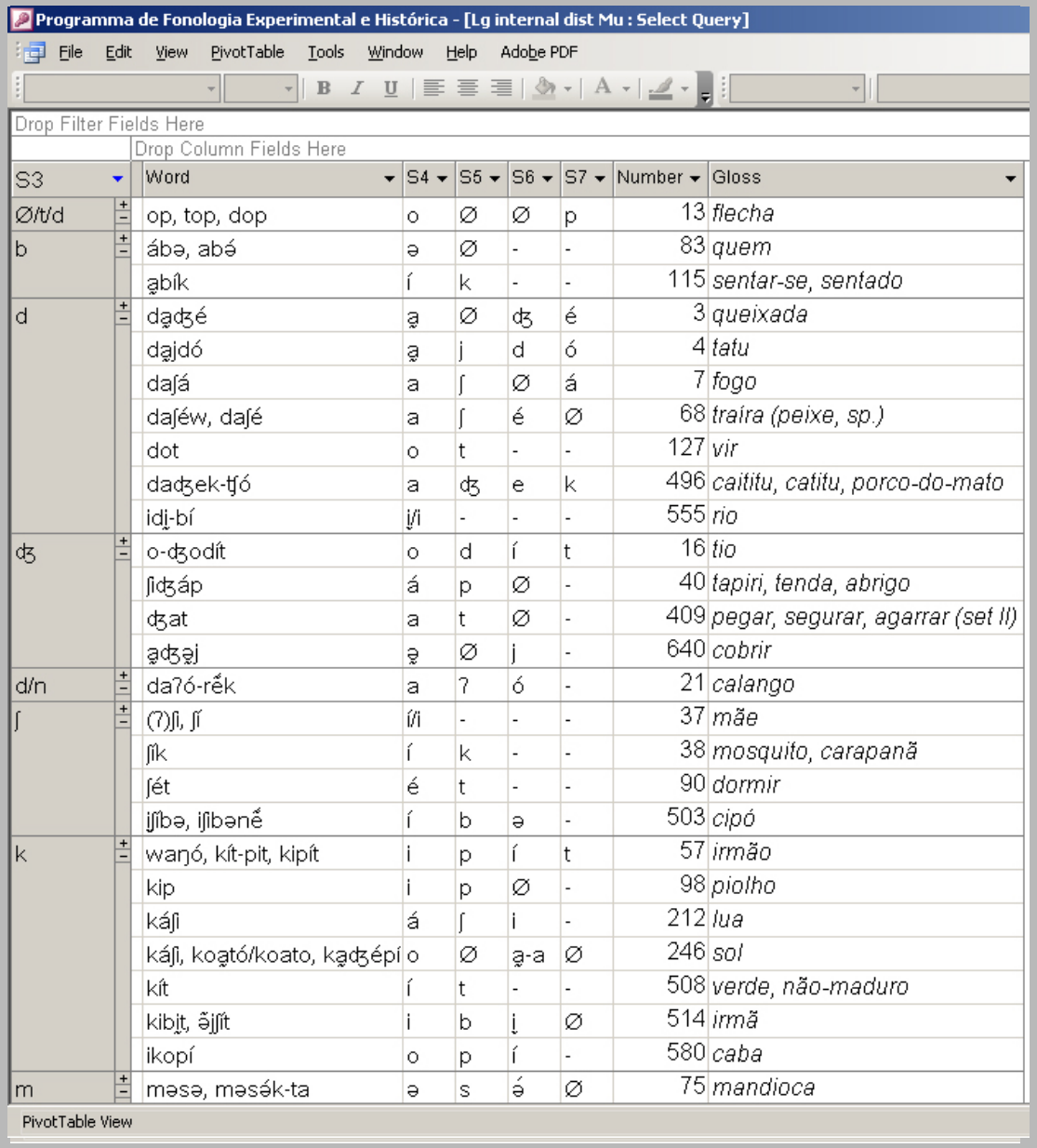
Numb	Gloss	Word	1: Segment	2: Segment	3: Segment	4: Segment	5: Segment	6: Segment	7: Segment	8: Se
1	aranha	dóá	-	-	-	-	-	-	-	-
3	queixada	dạdʒé	-	-	d	ạ	Ø	dʒ	é	-
4	tatu	daidó	-	-	d	a	i	d	ó	-
7	fogo	daʃá	-	-	d	a	ʃ	Ø	á	-
13	flecha	op, top, dop	-	-	Ø/t/d	o	Ø	Ø	p	-
16	tio	o-dʒodít	-	-	dʒ	o	d	í	t	-
18	uxi, oxi (fruta, sp.)	tádo	-	-	t	á	d	o	-	-
20	pulga, bicho-de-pé	nõŋ-ʔá	-	-	n	õ	ŋ	-	-	-
21	calango	daʔó-rẽk	-	-	d/n	a	ʔ	ó	-	-
25	dente	tə̃j, nə̃j, tə̃j, nə̃j	-	-	t/n	ə̃/ə̃	j	Ø	-	-
27	folha	təp, dəp, táp, dáp	-	-	t/d	ə/á	p	-	-	-
28	semente, caroço	ta/da, tá/dá, we-nə̃j	-	-	n	ə̃/ə̃	Ø	Ø	j	-
29	cabelo	táp, dáp, tap, dap	-	-	t/d	á/a	p	-	-	-
30	água, líquido	ti, di, tí, dí	-	(i)	t/d	í/i	-	-	-	-
31	pó, massa de minga	ʃín-tõm, tõ̞m, nõ̞m	-	-	t/n	õ̞	Ø	Ø	m	-
32	galho, ramo	takə̃/nakə̃, pənẽ	-	-	t/n	a	k	ə̃/ə̃	-	-
34	esposa (set II)	tajʃi	-	-	t	a	Ø	j	ʃ	i
36	rede	ə̗rə̃, ərə̃	-	-	Ø	ə̗/ə	r	ə̃/ə̃	-	-
37	mãe	(ʔ)ʃi, ʃí	-	-	ʃ	í/i	-	-	-	-
38	mosquito, carapanã	ʃík	-	-	ʃ	í	k	-	-	-
39	batata	weʃík	-	-	w	e	Ø	ʃ	í	k
40	tapiri, tenda, abrigo	ʃidʒáp	ʃ	i	dʒ	á	p	Ø	-	-
43	banhar-se	ạdʒók	-	-	Ø	ạ	dʒ	ó	k	-
51	anzol	piŋá	-	-	p	i	ŋ	á	Ø	-
54	pescoço	ạŋóbə	-	-	-	-	-	-	-	-
57	irmão	waŋó, kít-pit, kipít	-	-	k	i	p	í	t	-
61	ir (ir embora)	dʒə́, tʃə́	-	-	tʃ/dʒ	ə́	Ø	Ø	-	-
62	preguiça (animal)	aj	-	a	Ø	j	-	-	-	-

Record: 1 of 96

Datasheet View

All of the columns in the treats as Pivot objects. A cross-tab query is achieved by dragging and dropping column segments to the shaded ‘Drop Row Fields Here’ field on the left.

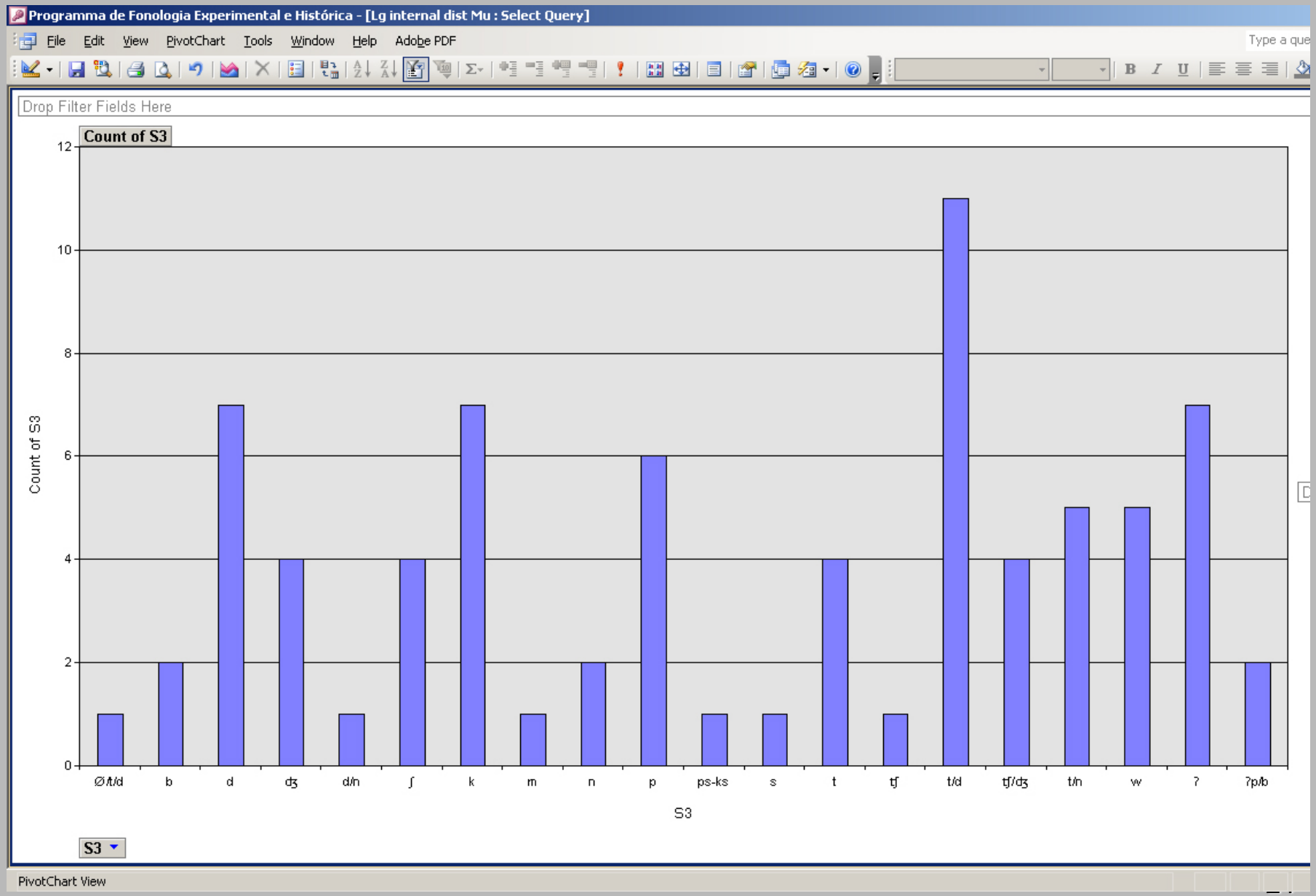
The phonemes in S3 act as a sort function: all of the words that begin with that phoneme are organized according to S3.



S3	Word	S4	S5	S6	S7	Number	Gloss
Ø/t/d	op, top, dop	o	Ø	Ø	p	13	*flecha*
b	ábə, abá	ə	Ø	-	-	83	*quem*
	ą̇bík	í	k	-	-	115	*sentar-se, sentado*
d	dą̇ʤé	ą̇	Ø	ʤ	é	3	*queixada*
	dą̇jdó	ą̇	j	d	ó	4	*tatu*
	daʃá	a	ʃ	Ø	á	7	*fogo*
	daʃéw, daʃé	a	ʃ	é	Ø	68	*traíra (peixe, sp.)*
	dot	o	t	-	-	127	*vir*
	daʤek-ʧó	a	ʤ	e	k	496	*caititu, catitu, porco-do-mato*
	idi̱-bí	i̱/i	-	-	-	555	*rio*
ʤ	o-ʤodít	o	d	í	t	16	*tio*
	ʃiʤáp	á	p	Ø	-	40	*tapiri, tenda, abrigo*
	ʤat	a	t	Ø	-	409	*pegar, segurar, agarrar (set II)*
	ą̇ʤą̇j	ą̇	Ø	j	-	640	*cobrir*
d/n	da?ó-rẽ́k	a	?	ó	-	21	*calango*
ʃ	(?)ʃi, ʃí	í/i	-	-	-	37	*mãe*
	ʃík	í	k	-	-	38	*mosquito, carapanã*
	ʃét	é	t	-	-	90	*dormir*
	ɨʃíbə, ɨʃíbənẽ́	í	b	ə	-	503	*cipó*
k	waŋó, kít-pit, kipít	i	p	í	t	57	*irmão*
	kip	i	p	Ø	-	98	*piolho*
	káʃi	á	ʃ	i	-	212	*lua*
	káʃi, koą̇tó/koato, ką̇ʤépí	o	Ø	ą̇-a	Ø	246	*sol*
	kít	í	t	-	-	508	*verde, não-maduro*
	kibi̱t, ə̃jɨ́t	i	b	i̱	Ø	514	*irmã*
	ikopí	o	p	í	-	580	*caba*
m	məsə, məsək-ta	ə	s	ə́	Ø	75	*mandioca*

PivotTable View

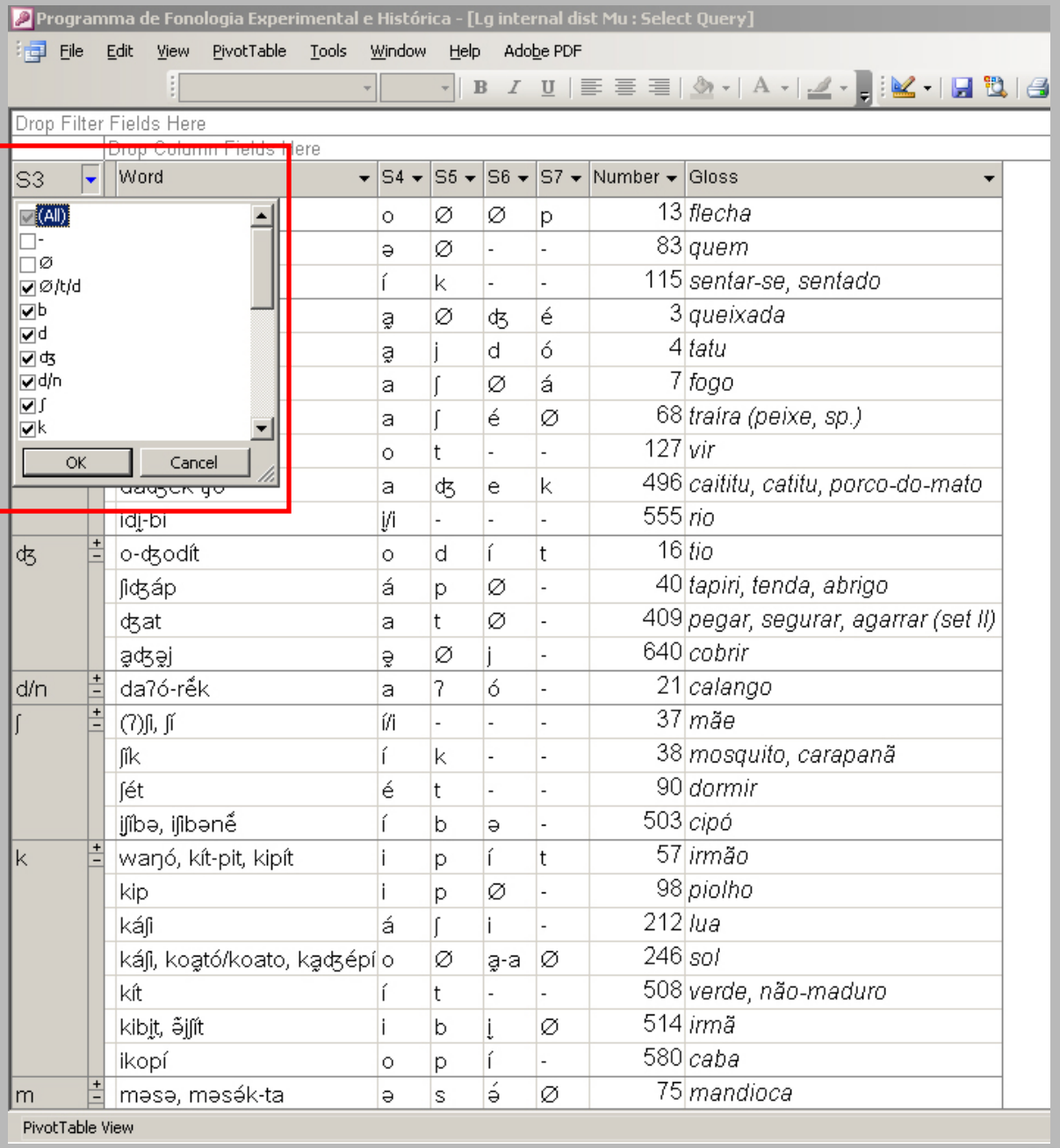
Pivot tables can be converted into Pivot charts, which can perform basic statistical procedures, in this case, counting the distribution of root-initial consonants:



Boolean searches are replaced by a filtering function.

Rather than specifying segments/phonemes, they are ‘filtered’ out by deselecting them from a pull down box attached to the sorting column.

For example, if we are interested in tracking the distribution of labials, we simply delesect all non-labial phonemes.



The sort function re-applies, with only the labial phonemes left, along with all the words root-initial labials are found in.

We believe this has an advantage over standard string queries: rather than spend time searching for phonemes that might not be there, you select what's available and filter out the rest.

Programma de Fonologia Experimental e Histórica - [Lg internal dist Mu : Select Query]

File Edit View PivotTable Tools Window Help Adobe PDF

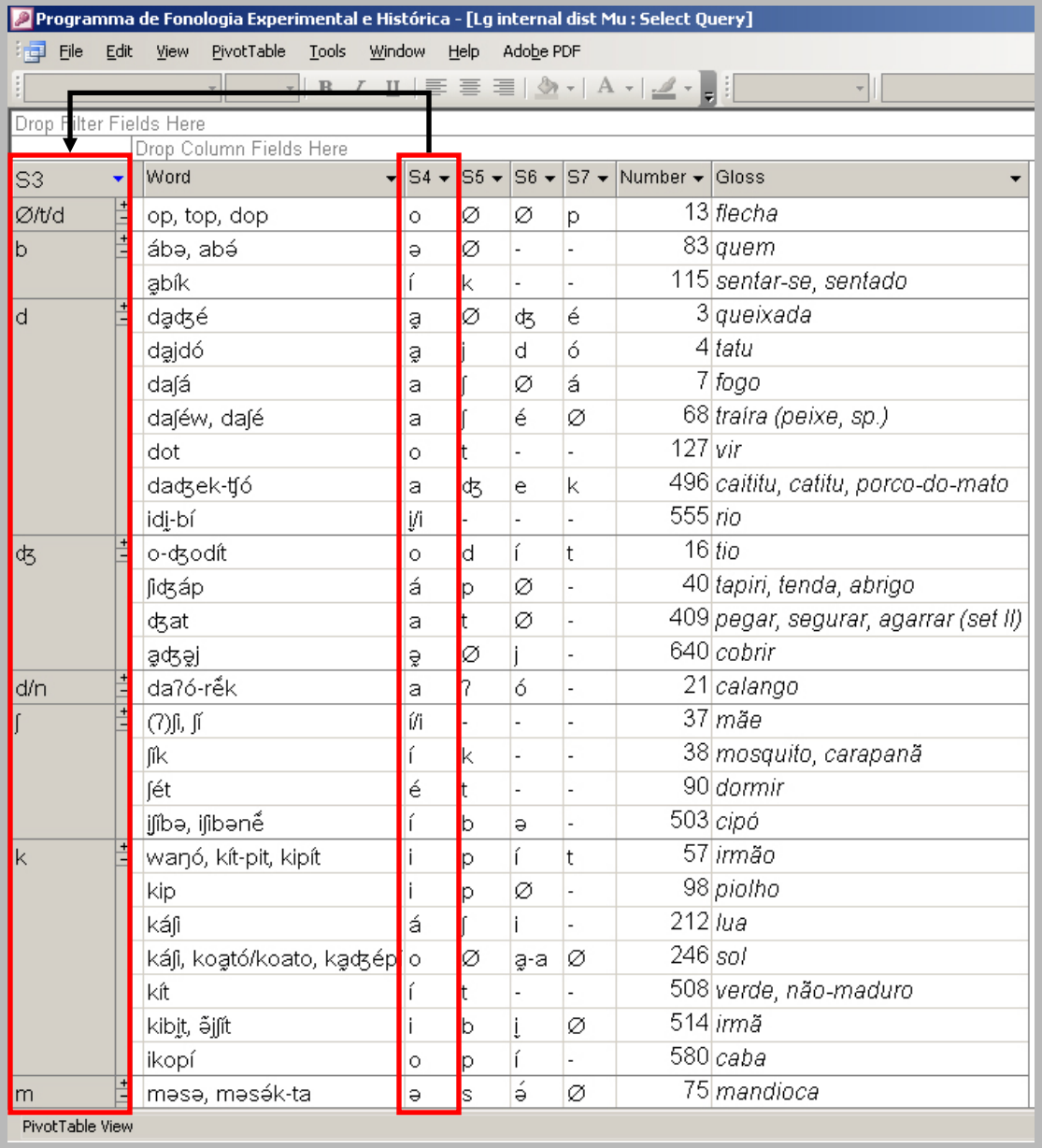
Drop Filter Fields Here

Drop Column Fields Here

S3	Word	S4	S5	S6	S7	Number	Gloss
b	ábə, abə́	ə	∅	-	-	83	*quem*
	ą̇bík	í	k	-	-	115	*sentar-se, sentado*
m	məsə, məsə́k-ta	ə	s	ə́	∅	75	*mandioca*
p	piŋá	i	ŋ	á	∅	51	*anzol*
	pə̨j-bə́, pəj-bə	ə	j	-	-	76	*cobra (gen.)*
	poʃí	o	ʃ	í	∅	133	*pesado*
	pəw	ə	w	-	-	247	*soprar, assoprar*
	pósə̃ŋ	ó	s	ə̃	ŋ	369	*remédio*
	pəkasó, ʧoót	ə	k	a	s	502	*pombo, rolinha*
w	weʃík	e	∅	ʃ	í	39	*batata*
	wjʃą̇	i̯	ʃ	∅	∅	91	*saúva, formiga (sp.)*
	witő	i	t	ő	-	219	*mutum*
	witá-ʔa	i	t	á	-	265	*pedra*
	waʔiʔa	a	ʔ	í	ʔ	489	*cabaça*
Grand T							

The drop field can accommodate more than one column. We can use this to track what kind of vowel follows the root-initial onsets.

Because S4 will typically contain these vowels, it is dragged over to the drop area and placed to the right of S3

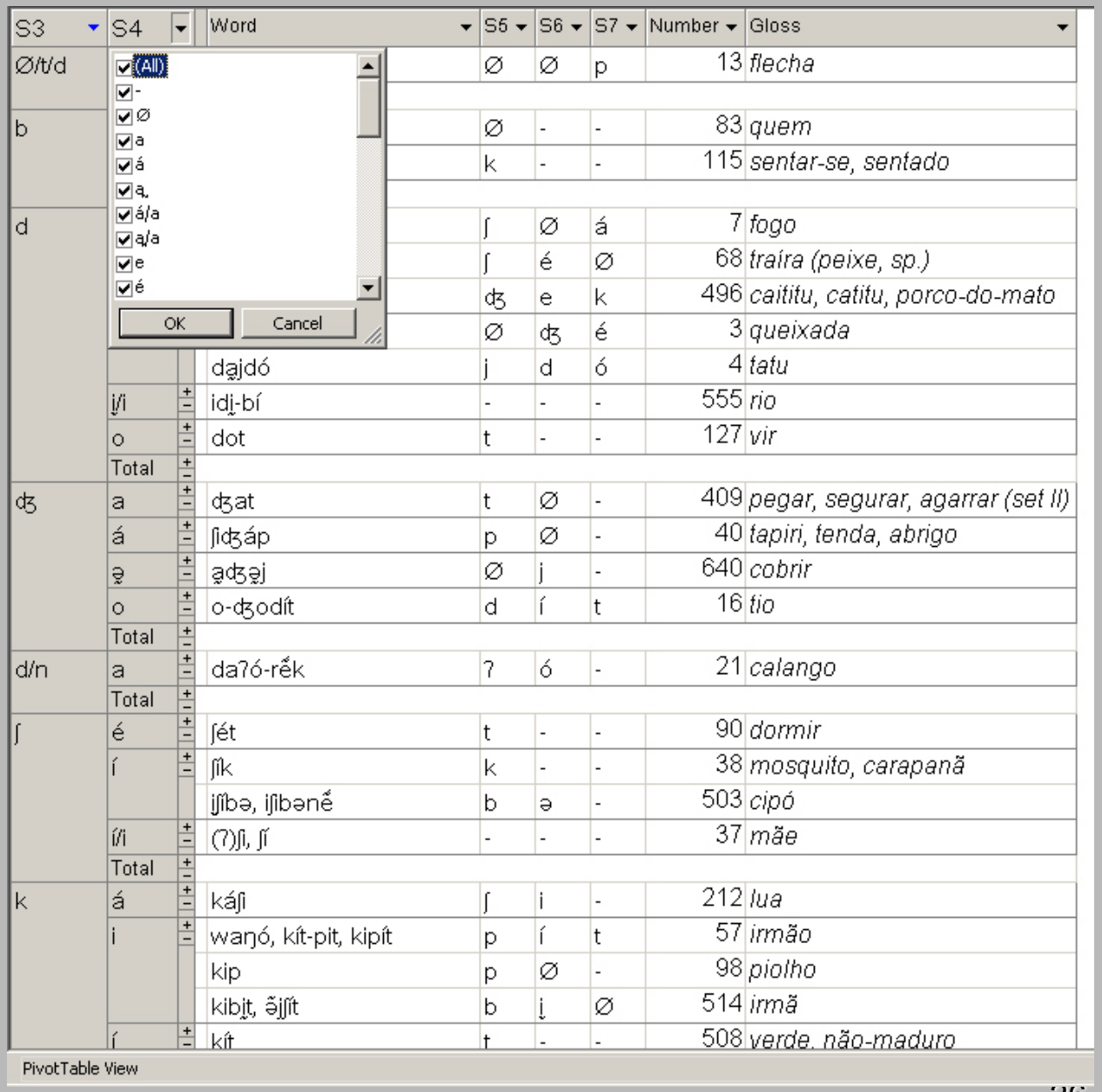


S3	Word	S4	S5	S6	S7	Number	Gloss
Ø/t/d	op, top, dop	o	Ø	Ø	p	13	*flecha*
b	ábə, abə́	ə	Ø	-	-	83	*quem*
	ą̇bík	í	k	-	-	115	*sentar-se, sentado*
d	dą̇ʤé	ą̇	Ø	ʤ	é	3	*queixada*
	dą̇jdó	ą̇	j	d	ó	4	*tatu*
	daʃá	a	ʃ	Ø	á	7	*fogo*
	daʃéw, daʃé	a	ʃ	é	Ø	68	*traíra (peixe, sp.)*
	dot	o	t	-	-	127	*vir*
	dadʒek-ʧó	a	ʤ	e	k	496	*caititu, catitu, porco-do-mato*
	idi̯-bí	i̯/i	-	-	-	555	*rio*
ʤ	o-ʤodít	o	d	í	t	16	*tio*
	ʃiʤáp	á	p	Ø	-	40	*tapiri, tenda, abrigo*
	ʤat	a	t	Ø	-	409	*pegar, segurar, agarrar (set II)*
	ą̇ʤą̇j	ą̇	Ø	j	-	640	*cobrir*
d/n	da?ó-rẽ́k	a	?	ó	-	21	*calango*
ʃ	(?)ʃi, ʃí	í/i	-	-	-	37	*mãe*
	ʃík	í	k	-	-	38	*mosquito, carapanã*
	ʃét	é	t	-	-	90	*dormir*
	iʃíbə, iʃibənẽ́	í	b	ə	-	503	*cipó*
k	waŋó, kít-pit, kipít	i	p	í	t	57	*irmão*
	kip	i	p	Ø	-	98	*piolho*
	káʃi	á	ʃ	i	-	212	*lua*
	káʃi, koą̇tó/koato, kąʤép	o	Ø	ą̇-a	Ø	246	*sol*
	kít	í	t	-	-	508	*verde, não-maduro*
	kibi̯t, ẽjí̯t	i	b	i̯	Ø	514	*irmã*
	ikopí	o	p	í	-	580	*caba*
m	məsə, məsə́k-ta	ə	s	á	Ø	75	*mandioca*

The effect is one of hierarchical sorting: S3 sorts the vowels in S4, the join of which sorts the words that are associated with the combinations:

S3	S4	Word	S5	S6	S7	Number	Gloss
Ø/t/d	o	op, top, dop	Ø	Ø	p	13	*flecha*
	Total						
b	ə	ábə, abə́	Ø	-	-	83	*quem*
	í	a̱bík	k	-	-	115	*sentar-se, sentado*
	Total						
d	a	daʃá	ʃ	Ø	á	7	*fogo*
		daʃéw, daʃé	ʃ	é	Ø	68	*traíra (peixe, sp.)*
		daʤek-ʧó	ʤ	e	k	496	*caititu, catitu, porco-do-mato*
	a̱	da̱ʤé	Ø	ʤ	é	3	*queixada*
		da̱jdó	j	d	ó	4	*tatu*
	j̨/i	idj̨-bí	-	-	-	555	*rio*
	o	dot	t	-	-	127	*vir*
	Total						
ʤ	a	ʤat	t	Ø	-	409	*pegar, segurar, agarrar (set II)*
	á	ʃiʤáp	p	Ø	-	40	*tapiri, tenda, abrigo*
	ə̨	a̱ʤə̨j	Ø	j	-	640	*cobrir*
	o	o-ʤodít	d	í	t	16	*tio*
	Total						
d/n	a	da7ó-rẽ́k	?	ó	-	21	*calango*
	Total						
ʃ	é	ʃét	t	-	-	90	*dormir*
	í	ʃík	k	-	-	38	*mosquito, carapanã*
		jʃíbə, iʃíbənẽ́	b	ə	-	503	*cipó*
	í/i	(7)ʃi, ʃí	-	-	-	37	*mãe*
	Total						
k	á	káʃi	ʃ	i	-	212	*lua*
	i	waŋó, kít-pit, kipít	p	í	t	57	*irmão*
		kip	p	Ø	-	98	*piolho*
		kibj̨t, ə̃jjít	b	j̨	Ø	514	*irmã*
	í	kít	t	-	-	508	*verde, não-maduro*

Filtering can again be applied, but this time, to select only creaky and nasal vowels which follow the S3 consonants:



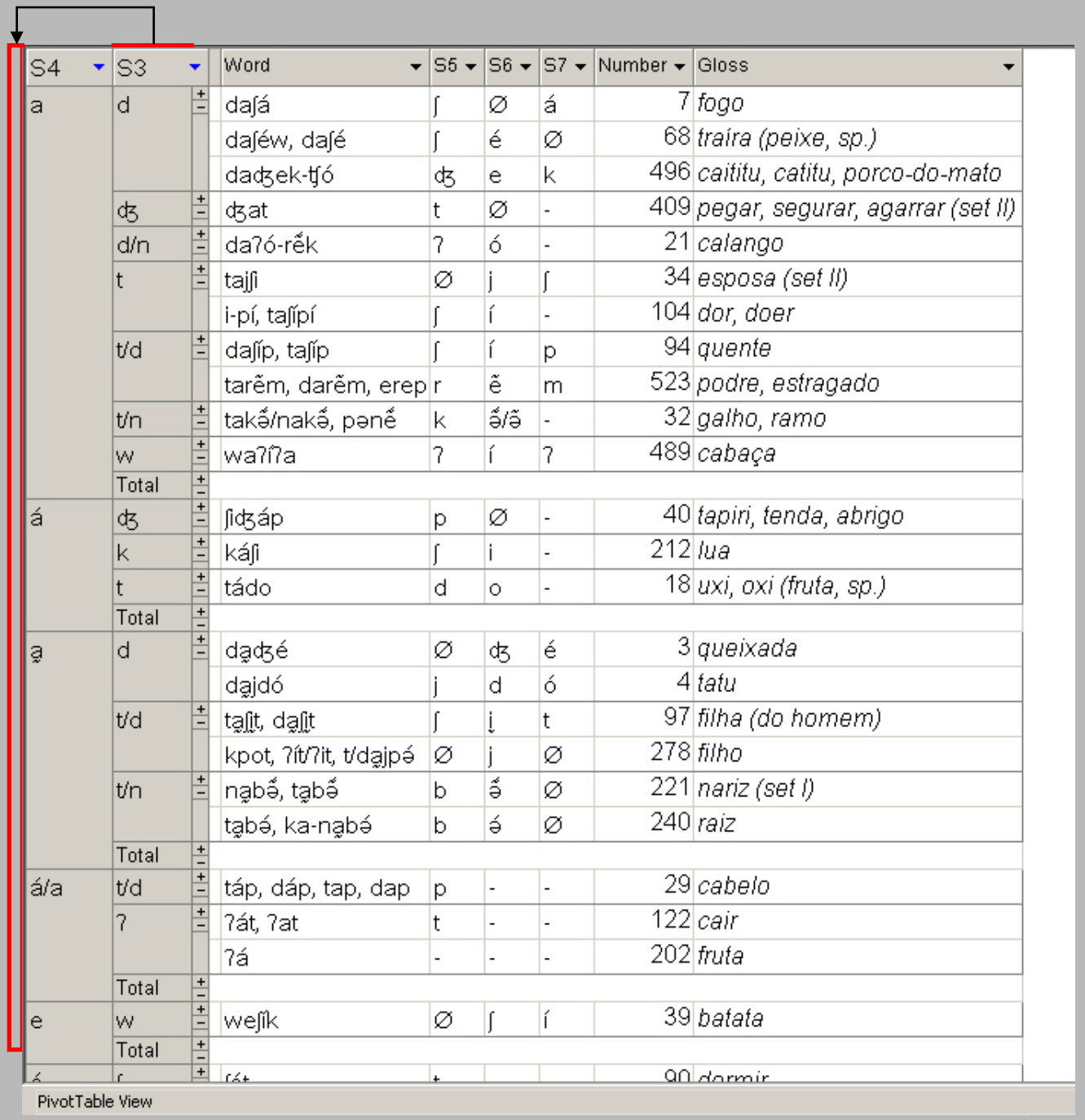
S3	S4	Word	S5	S6	S7	Number	Gloss
Ø/t/d			Ø	Ø	p	13	*flecha*
b			Ø	-	-	83	*quem*
			k	-	-	115	*sentar-se, sentado*
d			ʃ	Ø	á	7	*fogo*
			ʃ	é	Ø	68	*traíra (peixe, sp.)*
			ʤ	e	k	496	*caititu, catitu, porco-do-mato*
			Ø	ʤ	é	3	*queixada*
		dą̨jdó	j	d	ó	4	*tatu*
	į̨/i	idį̨-bí	-	-	-	555	*rio*
	o	dot	t	-	-	127	*vir*
	Total						
ʤ	a	ʤat	t	Ø	-	409	*pegar, segurar, agarrar (set II)*
	á	ʃiʤáp	p	Ø	-	40	*tapiri, tenda, abrigo*
	ą̨	ą̨ʤą̨j	Ø	j	-	640	*cobrir*
	o	o-ʤodít	d	í	t	16	*tio*
	Total						
d/n	a	da7ó-rẽ́k	7	ó	-	21	*calango*
	Total						
ʃ	é	ʃét	t	-	-	90	*dormir*
	í	ʃík	k	-	-	38	*mosquito, carapanã*
		jʃíbə, iʃíbənẽ́	b	ə	-	503	*cipó*
	í/i	(7)ʃi, ʃí	-	-	-	37	*mãe*
	Total						
k	á	káʃi	ʃ	i	-	212	*lua*
	i	waŋó, kít-pit, kipít	p	í	t	57	*irmão*
		kip	p	Ø	-	98	*piolho*
		kibį̨t, ə̃jʃít	b	į̨	Ø	514	*irmã*
	í	kít	t	-	-	508	*verde, não-maduro*

PivotTable View

Filtering is not hierarchical: any filters applied will eliminate the values in any other sort column.

S3	S4	Word	S5	S6	S7	Number	Gloss
d	ạ	dạʤé	Ø	ʤ	é	3	*queixada*
		dạjdó	j	d	ó	4	*tatu*
	ị/i	idị-bí	-	-	-	555	*rio*
	Total						
ʤ	ẹ	ạdʒẹj	Ø	j	-	640	*cobrir*
	Total						
n	ə̃/ə̃	ta/da, tá/dá, we-nə̃j	Ø	Ø	j	28	*semente, caroço*
	õ	nõŋ-ʔá	ŋ	-	-	20	*pulga, bicho-de-pé*
	Total						
s	ə̃	sə̃sə̃n (sə̃t)	t	-	-	410	*ter vergonha, envergonhar-se*
	Total						
t/d	ạ	tạʃɪt, dạʃɪt	ʃ	ɪ	t	97	*filha (do homem)*
		kpot, ʔít/ʔit, t/dạjpə́	Ø	j	Ø	278	*filho*
	Total						
ʧ/ʤ	ḛ	ʤḛ́m, ʧḛ́m	m	-	-	500	*sair*
	ẹ	ʧẹ̀k	Ø	Ø	k	82	*frio*
	Total						
t/n	ạ	nạbə̃́, tạbə̃́	b	ə̃́	Ø	221	*nariz (set I)*
		tạbə́, ka-nạbə́	b	ə́	Ø	240	*raiz*
	ə̃/ə̃	tə̃́j, nə̃́j, tə̃j, nə̃j	j	Ø	-	25	*dente*
	ọ̃	ʃín-tọ̃́m, tọ̃m, nọ̃m	Ø	Ø	m	31	*pó, massa de mingau*
	Total						
w	ị	wịʃạ	ʃ	Ø	Ø	91	*saúva, formiga (sp.)*
	Total						
Grand Total							

S4 can be dragged over to precede S3. In effect, we can observe the distribution of which consonants serve as onsets to which vowels, as sorted by those vowels.



S4	S3	Word	S5	S6	S7	Number	Gloss
a	d	daʃá	ʃ	Ø	á	7	*fogo*
		daʃéw, daʃé	ʃ	é	Ø	68	*traíra (peixe, sp.)*
		dadʒek-ʧó	dʒ	e	k	496	*caititu, catitu, porco-do-mato*
	dʒ	dʒat	t	Ø	-	409	*pegar, segurar, agarrar (set II)*
	d/n	da?ó-rẽ́k	?	ó	-	21	*calango*
	t	tajʃi	Ø	j	ʃ	34	*esposa (set II)*
		i-pí, taʃípí	ʃ	í	-	104	*dor, doer*
	t/d	daʃíp, taʃíp	ʃ	í	p	94	*quente*
		tarẽm, darẽm, erep	r	ẽ	m	523	*podre, estragado*
	t/n	takẽ́/nakẽ́, pənẽ́	k	ẽ́/ə	-	32	*galho, ramo*
	w	wa?i?a	?	í	?	489	*cabaça*
	Total						
á	dʒ	ʃidʒáp	p	Ø	-	40	*tapiri, tenda, abrigo*
	k	kájʃi	ʃ	i	-	212	*lua*
	t	tádo	d	o	-	18	*uxi, oxi (fruta, sp.)*
	Total						
ą	d	dądʒé	Ø	dʒ	é	3	*queixada*
		dąjdó	j	d	ó	4	*tatu*
	t/d	tąʃi̯t, dąʃi̯t	ʃ	i̯	t	97	*filha (do homem)*
		kpot, ?ít/?it, t/dąjpé	Ø	j	Ø	278	*filho*
	t/n	nąbẽ́, tąbẽ́	b	ẽ́	Ø	221	*nariz (set I)*
		tąbé, ka-nąbé	b	é	Ø	240	*raiz*
	Total						
á/a	t/d	táp, dáp, tap, dap	p	-	-	29	*cabelo*
	?	?át, ?at	t	-	-	122	*cair*
		?á	-	-	-	202	*fruta*
	Total						
e	w	weʃík	Ø	ʃ	í	39	*batata*
	Total						
é	ʃ	ʃét	t			90	*dormir*

PivotTable View

In sum so far:

- Flexibility: This methodology of ‘dragging and dropping’, swapping, filtering and pivot charting can be applied to all the S1-S10 segment columns.
- Intervocalic consonants, root-final segments etc. can all be tracked.
- Corrections can be monitored for and made: all modifications made in the language data entry forms is automatically updated in the pivot tables.

II. Mundurukú and Kuruáya: Comparing S3 onsets

- Because of the relational capabilities of Access, we can join the S1-S10 data from several languages and track their correspondences.
- All of the same functions demonstrated above can be applied, and this will be an effective tool in working out phonemic reconstruction.

A SQL query is generated that joins the S3 data from both Mundurukú (S3-M) and Kuruáya (S3-K).

By dragging S3-M over to the drop area, the corresponding S3-K and their words are sorted to S3-M.

S3-M	S3-K	Mundurukú	Kuruáya	Número	Gloss
ps-ks	b	o-psa̱, -ksa̱	ó-bia̱	87	*fígado*
s		sə̃sə̃n (sə̃t)		410	*ter vergonha, envergonhar-se*
t	t	tádo	talɔ́	18	*uxí, oxí (fruta, sp.)*
	t	tajʃi	ó-taitʃi	34	*esposa (set II)*
	t	pə-tét, bə-tét	o-bɨ́-tet	81	*nome*
		i-pí, taʃí-pí		104	*dor, doer*
tʃ	tʃ	tʃókə̃n	tʃokã̱n	64	*tucano*
t/d	Ø/t/l	təp, dəp, tə́p, də́p	tɨp, lɨp, ɨp	27	*folha*
	t/l	táp, dáp, tap, dap	tap, lap	29	*cabelo*
	t/l	ti/di, tí/dí, idi	títi, ti, li	30	*água, líquido*
	t	dajíp, tajíp	tákip	94	*quente*
		ta̱ʃi̱t, da̱ʃi̱t		97	*filha (do homem)*
	t-l	toj, doj, tój, dój	loj, toj	101	*sangue*
	t-l	tiŋ, diŋ	tĩŋ, lĩŋ	116	*fumaça*
		topsa̱, dopsa̱	tóbia̱, lóbia̱	229	*ovo*
	t	toa̱j-bə́ /doa̱j-bə́	táj-bɨ	239	*rabo*
		kpot, ʔít/ʔit, t/da̱j-pə́		278	*filho*
		tarə̃m, darə̃m, erep	í-patek	523	*podre, estragado*
tʃ/ʤ	d	ʤə́, tʃə́	o-dɨ́	61	*ir (ir embora)*
	tʃ-d	tʃa̱k, ka-ʤa̱k	í-tʃɨk, ká-dɨk	82	*frio*
		ʤó, tʃó	o-dódó	103	*ver*
		ʤə̃́m, tʃə̃́m		500	*saír*
t/n	t/l	tə̃j, nə̃j, tə̃j, nə̃j	tə̃j, lə̃j	25	*dente*
	t/l	ʃĩn-tə̃m, tõ̱m, nõ̱m	tõ̱m, lõ̱m	31	*pó, massa*
		takə̃́/nakə̃́, pənə̃́		32	*galho, ramo*
	t-l	na̱bə̃́, ta̱bə̃́	lomĩ, tomĩ, lõbɨ	221	*nariz (set I)*
		ta̱bə́, ka-na̱bə́		240	*raíz*
w	w	weʃík	wedík	39	*batata*
	w	wiʃa̱	wíʃa-laj	91	*saúva, formiga (sp.)*

PivotTable View

This correspondence can be easily switched by simply swapping S3-M for S3-K:

By inserting S4-M after S3-M, we can observe the correspondence between Mundurukú root-intial CV sequences with those in Kuruáya:

S3-K	S3-M	Mundurukú	Kuruáya	Número	Gloss
b	ps-ks	o-psa̱, -ksa̱	ó-bia̱	87	*fígado*
d	ʃ	(ʔ)ʃi, ʃí	i-di	37	*mãe*
	ʃ	ʃík	dik	38	*mosquito, carapanã*
	ʤ	ʃidʤáp	kidap	40	*tapíri, tenda, abrigo*
	ʧ/ʤ	ʤá, ʧá	o-dɨ́	61	*ir (ir embora)*
	ʃ	ʃíbə, ʃíbənɛ̃	ídi-bɨ, i-bɨ, kóro-bɨ	503	*cipó*
k	k	kip	kip	98	*piolho*
l	d	daʃá	láʃa	7	*fogo*
	n	nõŋ-ʔá	lõŋ	20	*pulga, bicho-de-pé*
	d/n	daʔó-rɛ̃k, naʔó-rɛ̃k	láʔo	21	*calango*
	n	ta/da, tá/dá, we-nɛ̃j	ta, la, we-lɛ̃-ĩ	28	*semente, caroço*
	d	daʃé-w, daʃé	láʃoj	68	*traíra (peixe, sp.)*
m	m	məsə, məsək-ta	másik	75	*mandioca*
p	p	piŋá	pinã	51	*anzol*
	p	pa̱j-bá, pəj-bə	pɨ́i	76	*cobra (gen.)*
t	t	tádo	talɔ́	18	*uxí, oxí (fruta, sp.)*
	t	tajʃi	ó-taiʧi	34	*esposa (set II)*
	t	pə-tét, bə-tét	o-bɨ́-tet	81	*nome*
	t/d	daʃíp, taʃíp	tákip	94	*quente*
	t/d	toa̱j-bá /doa̱j-bá	táj-bɨ	239	*rabo*
ʧ	ʧ	ʧókɔ̃n	ʧokãn	64	*tucano*
t/l	t/n	tɛ̃j, nɛ̃j, tɛ̃j, nɛ̃j	tãj, lãj	25	*dente*
	t/d	táp, dáp, tap, dap	tap, lap	29	*cabelo*
	t/d	ti/di, tí/dí, idi	títi, ti, li	30	*água, líquido*
	t/n	ʃĩn-tɔ̃m, tɔ̃m, nɔ̃m	tɔ̃m, lɔ̃m	31	*pó, massa*
ʧ-d	ʧ/ʤ	ʧa̱k, ka-ʤa̱k	í-ʧɨk, ká-dɨk	82	*frio*
t-l	t/d	toj, doj, tój, dój	loj, toj	101	*sangue*
	t/d	tiŋ, diŋ	tĩŋ, lĩŋ	116	*fumaça*
	t/n	nabɛ̃, tabɛ̃	lomĩ, tomĩ, lõbɨ	221	*naríz (set I)*

PivotTable View

S3-M	S4-M	S3-K	S4-K	Mundurukú	Kuruáya	Número	Gloss
t	a	t	a	tajʃi	ó-taitʃi	34	*esposa (set II)*
	á	t	a	tádo	talɔ́	18	*uxí, oxí (fruta, sp.)*
	é	t	e	pə-tét, bə-tét	o-bɨ́-tet	81	*nome*
	Total						
tʃ	ó	tʃ	o	tʃókɔ̃n	tʃokãn	64	*tucano*
	Total						
t/d	a	t	á	daʃíp, taʃíp	tákip	94	*quente*
	ą						
	á/a	t/l	a	táp, dáp, tap, dap	tap, lap	29	*cabelo*
	ə/ə́	Ø/t/l	ɨ	təp, dəp, táp, dáp	tip, lip, ɨp	27	*folha*
	i	t-l	ĩ	tiŋ, diŋ	tĩŋ, lĩŋ	116	*fumaça*
	í/i	t/l	i	ti/di, tí/dí, idi	títi, ti, li	30	*água, líquido*
	o	t	Ø	toą̱j-bə́ /doą̱j-bə́	táj-bɨ	239	*rabo*
	ó-o	t-l	o	toj, doj, tój, dój	loj, toj	101	*sangue*
	Total						
tʃ/ʤ	ɛ́						
	ə́	d	ɨ́	ʤə́, tʃə́	o-dɨ́	61	*ir (ir embora)*
	ą̱	tʃ-d	ɨ	tʃą̱k, ka-ʤą̱k	í-tʃɨk, ká-dɨk	82	*frio*
	ó						
	Total						
t/n	a						
	ą	t-l	o-õ	nąbə̃, tąbə̃	lomĩ, tomĩ, lõbɨ	221	*nariz (set I)*
	ə̃/ə̃	t/l	ã	tə̃j, nə̃j, tə̃j, nə̃j	tə̃j, lə̃j	25	*dente*
	ǫ̱	t/l	ǫ̱	ʃĩn-tǫ̃m, tǫ̱̃m, nǫ̱̃m	tǫ̱̃m, lǫ̱̃m	31	*pó, massa*
	Total						
w	a						
	e	w	e	weʃík	wedík	39	*batata*
	i						
	į	w	í	wiʃą̱	wíʃa-laj	91	*saúva, formiga (sp.)*
	Total						

PivotTable View

III. S3 across several languages

- There is no inherent limit to the number of segments or languages.
- Five languages: Mundurukú, Tapieté, Wayampi, Avá-Canoeiro, Kamayurá.

i. Test S3 correspondences across all of these languages simultaneously.

ii. Test CV correspondences.

S3-T (Tapieté) corresponding with the other four languages:

S3-T	S3-M	S3-W	S3-A	S3-K	Mundurukú	Tapieté	Wayampi	Avá-Canoeiro	Kamayurá	Gloss
k		k	k	k	kisé	kɨse	kɨse (empr.)	kɨe	kɨe-ʔi	*faca*
		k	k	k	taó, daó, tao, dao	t̃ɨ-kãwẽ	karʲʷe / karʲʷer	karŋ	karŋ	*osso*
				k	pariwát	karai		maira	karaʔip	*homem branco, civ*
			k	k	ikopí	kaw-apua, kaw-usu,		kaw	kap	*caba*
m		m	m		-ktop, i-top	ʃe-me	mẽ / men	men	i-ʔirũ	*marido*
			Ø		ʔ̃ɨt, jo-p̃ɨt	mĩj̃ĩ		atɨ		*pequeno*
mb				p		mbɨte			-pɨter-ɨp	*meio*
p		p	p	p	ŋebaᷠ	tɨ-pepo	pepũ-kã	i-pepo	i-pepo	*asa*
	p	p	p	p	poʃí	pohɨ	poɨj / po(w)ɨj (WA)	poɨj	i-powɨj	*pesado*
		p	p	p	kabi-ók	p̃ɨtu	pɨa, pɨsajé, pɨtũ	ɨpɨaj, pɨtun	ɨpɨtun	*noite, escuridão*
		p	p	p	ajpãn, ísə́	taʔɨ, pɨʔahu	pɨau	i-pɨau	i-pɨau	*novo, jovem*
	Ø	p	p	m/p	i, í	te-pɨ	pɨ	pɨŋ	mɨ, i-pɨ	*pé*
	ʔp/b	p	p	p	(ʔ)pə, pə́, bə, bə́	po, te-po	po	kʷ̃a, i-po	hʷ̃a, po	*mão*
		p	p		-opsaᷠ, -ksaᷠ	pɨʔa	pɨa-kʷe / pɨa-kʷer	pɨa	i-pere	*fígado*
r/h			Ø	Ø	tópa, dópa	rowa, howa		owa	owa	*face, rosto*
	t/n	Ø	Ø/r	Ø	tə̃j, nə̃j, tə̃j, nə̃j	rãĩ, hãĩ	ãj	i-ãj, ni-rãj	aĩ	*dente*
	t/d	Ø	Ø	ʔ	táp, dáp, tap, dap	ha, ra	ap-ira	ap	ʔap	*cabelo*
s		s	ʧ	Ø	á, idə́	suʔu, hemo	suʔu	ʧu	ne-uʔu	*morder, picar*
	ʔʃ	Ø	Ø	Ø	(ʔ)ʃi, ʃí	sɨ	ɨ	ɨ	ij-ɨ	*mãe*
s/h/r		Ø	Ø/ʁ	h	ẽn	soʔo, hoʔo, roʔo	oʔo	o, -ʁo	haʔo	*carne*
t	ʧ		t	t	ʧókõn	tũka		tukan	tukan	*tucano*
				t	taᷠʃ̃t, daᷠʃ̃t	tadʒɨ			tajɨt	*filha*
ʧ		k	k	k	ʃét	o-ʧe	a-ke	ker / kɨr	e-ket	*dormir*
t/r		Ø	Ø	t/r	baj	tu, ru	u	papaj, uw	apa, tup, je=rup	*pai*
w		s	w		wáʔodá	wasu-wasu	soʔo	watʃu		*veado*
		w/β	w	w	ɨpí	ɨwɨ	ɨwɨ / ɨβɨ (WA)	ɨwɨ, ɨj-a	ɨwɨ	*terra, solo, chão*
			Ø	Ø	-j̃ĩ̃, -j̃ĩrí	-kʷ̃e, wasu		-uʁu, -u	-u	*AUMentativo*
					aᷠdʒə́m	wãhe		ɨk	o-ɨk	*chegar*
										passarinho, pássar

PivotTable View

S3-M (Mundurukú) corresponding with the other four languages:

S3-M	S3-T	S3-W	S3-A	S3-K	Mundurukú	Tapieté	Wayampi	Avá-Canoeiro	Kamayurá	Gloss
p	p	p	p	p	poʃí	pohɨ	poɨj / po(w)ɨj (WA)	poɨj	i-powɨj	*pesado*
					pəw	pedʒu	peju	o-peju, pɨ	ja-pɨ	*soprar, assoprar*
					pósə̃ŋ	mõha			moaŋ, i-hoaŋ	*remédio*
s					sə̃sə̃n (sə̃t)	mãrã, pɨhu			otsĩ	*ter vergonha, enverg*
t	h		Ø	Ø	i-pí, taʃí-pí	hasɨ		aɨ, i-aɨ	-aɨ	*dor, doer*
tʃ	t		t	t	tʃókə̃n	tũka		tukan	tukan	*tucano*
t/d					topsa̱, dopsa̱	piʔa	upiʔa	upia, ʁupia	-rupiʔa, upiʔa	*ovo*
	r/h	Ø	Ø	ʔ	táp, dáp, tap, dap	ha, ra	ap-ira	ap	ʔap	*cabelo*
					ti, di, tí, dí	ɨ	ɨ	ɨ	ʔɨ	*água, líquido*
	h	Ø	Ø	h	daʃíp, taʃíp	haku	aku	akup, i-akup	hakup	*quente*
		s	t		tiŋ, diŋ		tata-sĩ	tata-tiŋ		*fumaça*
tʃ/dʒ	h	Ø		r	dʒó, tʃó	heʃa	esa	a-mae	ne=retsak	*ver*
t/n	r/h	Ø	Ø/r	Ø	tə̃j, nə̃j, tə̃j, nə̃j	rãĩ, hãĩ	ãj	i-ãj, ni-rãj	aĩ	*dente*
w	dʒ		Ø	Ø	witá-ʔa	dʒita-kɨ		ita	ita	*pedra*
					witə̃́		mɨtũ	mɨtũ	mɨtũ	*mutum*
					weʃík	ʃure, dʒetɨ		jitika	jetik	*batata*
					waʔíʔa	tɨʔɨwa		ɨa	ɨʔa	*cabaça*
ʔ					ʔá, ʔa	iʔa	ʔa	a	i-ʔa	*fruta*
	ɨ	ɨ	(ɨ)	ɨ	ʔip, ʔíp	ɨwɨra	ʔɨ, ɨwɨra / ɨβɨra (WA)	ɨwɨra, wɨra	ɨwɨra	*árvore, madeira, pau*
	Ø	ʔ	Ø	ʔ	ʔát, ʔat	a	ʔa	o-ɨʁ, o-jɨpɨra	o-kuj, ʔat	*cair*
	h				ʔó/ʔo, kə̃n	karu, hoʔu		u, o-u	o-ʔu, a-karu	*comer*
					ʔot-pá					*larva, bicho-de-pau,*
					ʔé	mbeʔu		e	ja-ʔe	*dizer*
ʔʃ	s	Ø	Ø	Ø	(ʔ)ʃi, ʃí	sɨ	ɨ	ɨ	ij-ɨ	*mãe*
ʔp/b	p	p	p	p	(ʔ)pə, pá, bə, bá	po, te-po	po	kʷ̃ã, i-po	hʷ̃ã, po	*mão*
					(ʔ)pído, bído	tũtũhẽ	ɨwitu / ɨβitu (WA)		mitu, i-pitu	*ar, respiração, respi*
Grand T										

PivotTable View

S3-W (Wayampi) corresponding with the other four languages:

S3-W	S3-M	S3-T	S3-A	S3-K	Mundurukú	Tapieté	Wayampi	Avá-Canoeiro	Kamayurá	Gloss
k			k	k	ʃep		ka	i-kaw	i-kap	*banha, gordura*
		k	k	k	kisé	kise	kise (empr.)	kie	kie-ʔi	*faca*
		k	k	k	taó, daó, tao, dao	t̃i-kãwẽ	karʲ̃e / karʲ̃er	karŋ	karŋ	*osso*
		ʧ	k	k	ʃét	o-ʧe	a-ke	ker / kir	e-ket	*dormir*
	Count of S3-T	3								
m		m	m		-ktop, i-top	ʃe-me	mẽ / men	men	i-ʔirũ	*marido*
	Count of S3-T	1								
p		p	p	p	ŋebaꞡ	ti-pepo	pepũ-kã	i-pepo	i-pepo	*asa*
	p	p	p	p	poʃí	pohɨ	poij / po(w)ij (WA)	poij	i-powij	*pesado*
		p	p	p	kabi-ók	p̃itu	pia, pisajé, pitũ	ɨpiaj, pitun	ɨpitun	*noite, escuridão*
		p	p	p	ajpãn, ísé	taʔɨ, pɨʔahu	piau	i-piau	i-piau	*novo, jovem*
	Ø	p	p	m/p	i, í	te-pɨ	pɨ	pɨŋ	mɨ, i-pɨ	*pé*
		Ø	p	p		hũwa, hũ	piũ, pijũ	pitun	pitsun	*preto*
	ʔp/b	p	p	p	(ʔ)pə, pé, bə, bé	po, te-po	po	kʷ̃ã, i-po	hʷ̃ã, po	*mão*
		p	p		-opsaꞡ, -ksaꞡ	pɨʔa	pia-kʷ̃e / pia-kʷ̃er	pia	i-pere	*fígado*
			p	p	tit, tít, dit, dít		poti / potir	potiʁ	i-potir-a	*flor*
	Count of S3-T	8								
s		w	w		wáʔodá	wasu-wasu	soʔo	watʃu		*veado*
		s	ʧ	Ø	á, idé	suʔu, hemo	suʔu	ʧu	ne-uʔu	*morder, picar*
	ʧ/d		t		tiŋ, diŋ		tata-sĩ	tata-tiŋ		*fumaça*
	Count of S3-T	2								
w		w	w	w	wásẽ	wɨra	wɨra	wɨra-miri, wɨra	wɨra	*passarinho, pássaro*
	Count of S3-T	1								
w/β		w	w	w	ɟpí	ɨwɨ	ɨwɨ / ɨβɨ (WA)	ɨwɨ, ɨj-a	ɨwɨ	*terra, solo, chão*
	Count of S3-T	1								
ʔ	ʔ	Ø	Ø	ʔ	ʔát, ʔat	a	ʔa	o-ɨʁ, o-ɟɨpɨra	o-kuj, ʔat	*cair*
	Count of S3-T	1								

PivotTable View

S3-A (Avá-Canoeiro) corresponding with the other four languages:

S3-A	S3-M	S3-T	S3-W	S3-K	Mundurukú	Tapieté	Wayampi	Avá-Canoeiro	Kamayurá	Gloss
k			**k**	**k**	ʃep		ka	i-kaw	i-kap	*banha, gordura*
		k	**k**	**k**	kisé	kɨse	kɨse (empr.)	kɨe	kɨe-ʔi	*faca*
		k	**k**	**k**	taó, daó, tao, dao	ɬɨ-kãwẽ	karɲ^{w}e / karɲ^{w}er	karŋ	karŋ	*osso*
		ʧ	**k**	**k**	ʃét	o-ʧe	a-ke	ker / kɨr	e-ket	*dormir*
		k		**k**	ikopí	kaw-apua, kaw-usu,		kaw	kap	*caba*
	Count of S3-T	4								
m		**m**	**m**		-ktop, i-top	ʃe-me	mẽ / men	men	i-ʔirũ	*marido*
	Count of S3-T	1								
p		**p**	**p**	**p**	ŋebaʂ	ti-pepo	pepũ-kã	i-pepo	i-pepo	*asa*
	p	**p**	**p**	**p**	poʃí	pohɨ	poɨj / po(w)ɨj (WA)	poɨj	i-powɨj	*pesado*
		p	**p**	**p**	kabi-ók	pɨ̃tu	pɨa, pɨsajé, pɨtũ	ɨpɨaj, pɨtun	ɨpɨtun	*noite, escuridão*
		p	**p**	**p**	ajpãn, ísə	taʔɨ, pɨʔahu	pɨau	i-pɨau	i-pɨau	*novo, jovem*
	Ø	**p**	**p**	**m/p**	i, í	te-pɨ	pɨ	pɨŋ	mɨ, i-pɨ	*pé*
		Ø	**p**	**p**		hũwa, hũ	piũ, pijũ	pitun	pɨtsun	*preto*
	ʔp/b	**p**	**p**	**p**	(ʔ)pə, pá, bə, bá	po, te-po	po	k^{w}ã, i-po	h^{w}ã, po	*mão*
		p	**p**		-opsaʂ, -ksaʂ	pɨʔa	pɨa-k^{w}e / pɨa-k^{w}er	pɨa	i-pere	*fígado*
			p	**p**	tit, tít, dit, dít		potɨ / potɨr	potɨʁ	i-potɨr-a	*flor*
	Count of S3-T	8								
r		**Ø**		**Ø**		ti-reko		ʧi-remireko	emireko	*esposa (set I)*
	Count of S3-T	1								
t	**ʧ**	**t**		**t**	ʧókɔ̃n	tũka		tukan	tukan	*tucano*
	t/d		**s**		tiŋ, diŋ		tata-sĩ	tata-tiŋ		*fumaça*
	Count of S3-T	1								
ʧ		**s**	**s**	**Ø**	á, idə	su^{ʔ}u, hemo	su^{ʔ}u	ʧu	ne-u^{ʔ}u	*morder, picar*
	Count of S3-T	1								
w		**w**	**s**		wá^{ʔ}odá	wasu-wasu	so^{ʔ}o	watʃu		*veado*
		w	**w/β**	**w**	ɨpí	ɨwɨ	ɨwɨ / ɨβɨ (WA)	ɨwɨ, ɨj-a	ɨwɨ	*terra, solo, chão*

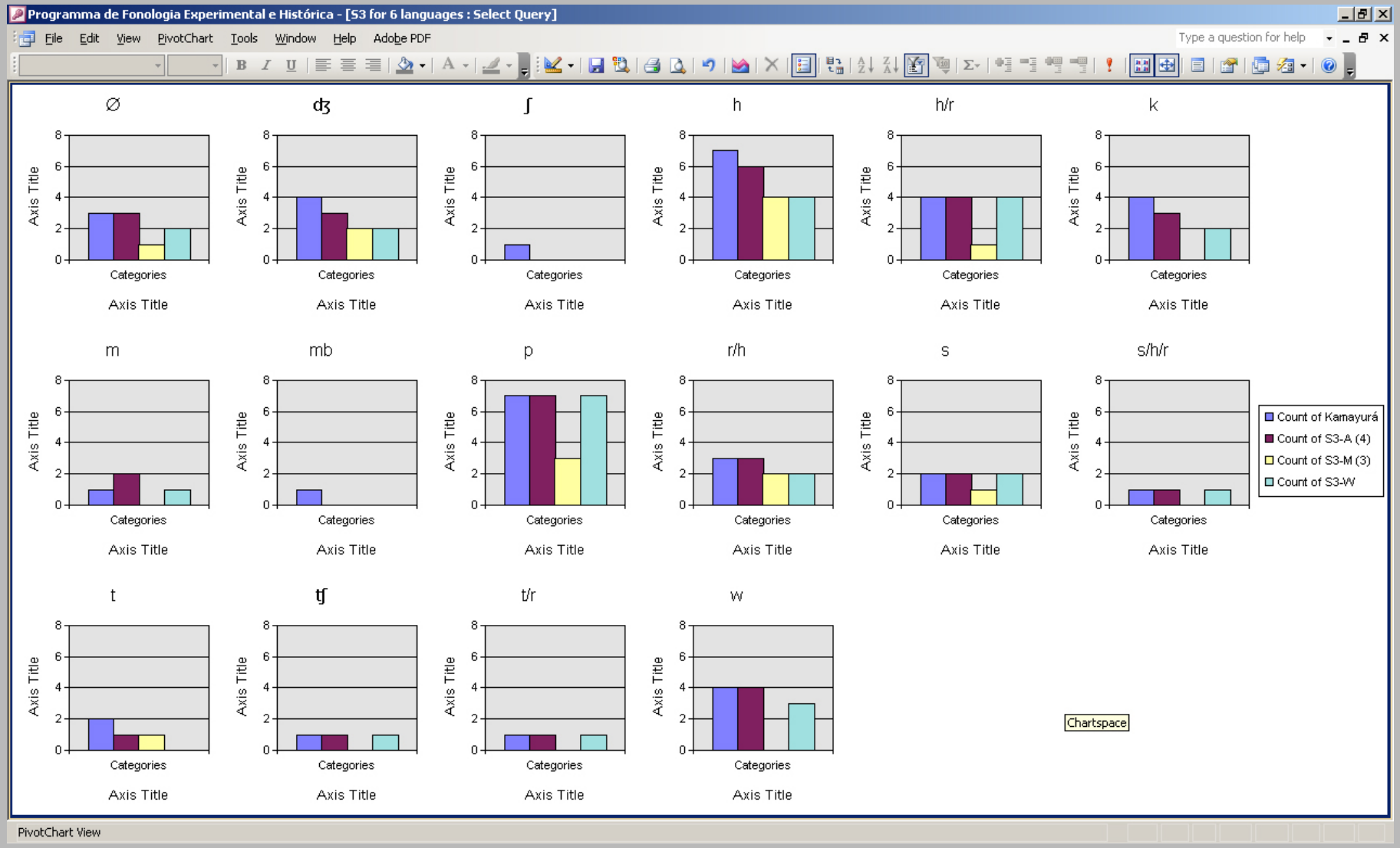
PivotTable View

S3-K (Kamayurá) corresponding with the other four languages:

S3-K	S3-M	S3-T	S3-W	S3-A	Mundurukú	Tapieté	Wayampi	Avá-Canoeiro	Kamayurá	Gloss
k			**k**	**k**	ʃep		ka	i-kaw	i-kap	*banha, gordura*
		k	**k**	**k**	kisé	kɨse	kɨse (empr.)	kɨe	kɨe-ʔi	*faca*
		k	**k**	**k**	taó, daó, tao, dao	t̃i-kãwẽ	karʲʷe / karʲʷer	karŋ	karŋ	*osso*
		tʃ	**k**	**k**	ʃét	o-tʃe	a-ke	ker / kɨr	e-ket	*dormir*
		k			pariwát	karai		maira	karaʔip	*homem branco, civ*
		k		**k**	ikopí	kaw-apua, kaw-usu,		kaw	kap	*caba*
	Count of S3-T 5									
m/p	**Ø**	**p**	**p**	**p**	i, í	te-pɨ	pɨ	pɨŋ	mɨ, i-pɨ	*pé*
	Count of S3-T 1									
p		**p**	**p**	**p**	ŋebaˌ	ti-pepo	pepũ-kã	i-pepo	i-pepo	*asa*
	p	**p**	**p**	**p**	poʃí	pohɨ	poɨj / po(w)ɨj (WA)	poɨj	i-powɨj	*pesado*
		p	**p**	**p**	kabi-ók	p̃itu	pɨa, pɨsajé, pɨtũ	ɨpɨaj, pɨtun	ɨpɨtun	*noite, escuridão*
		p	**p**	**p**	ajpãn, ísé	taʔɨ, pɨʔahu	pɨau	i-pɨau	i-pɨau	*novo, jovem*
		Ø	**p**	**p**		hũwa, hũ	piũ, pijũ	pitun	pɨtsun	*preto*
	ʔp/b	**p**	**p**	**p**	(ʔ)pə, pé, bə, bé	po, te-po	po	kʷʷã, i-po	ɦʷʷã, po	*mão*
			p	**p**	tit, tít, dit, dít		potɨ / potɨr	potɨʁ	i-potɨr-a	*flor*
		mb				mbɨte			-pɨter-ɨp	*meio*
	Count of S3-T 7									
r	**tʃ/dʒ**	**h**	**Ø**		dʒó, tʃó	heʃa	esa	a-mae	ne=reɨsak	*ver*
	Count of S3-T 1									
t	**Ø**	**h/r**	**Ø**	**Ø**	é	hape, rape	pee ~ ape	ape	tape	*caminho*
	tʃ	**t**		**t**	tʃókõn	tũka		tukan	tukan	*tucano*
		h/r	**Ø**	**Ø**	pətét, bətét	he, ti-re-ʔe	e / er	er-a	tet	*nome*
		t			taʃ̲ɨt, daʃ̲ɨt	tadʒɨ			tajɨt	*filha*
	Count of S3-T 4									
t/r		**t/r**	**Ø**	**Ø**	baj	tu, ru	u	papaj, uw	apa, tup, je=rup	*pai*
	Count of S3-T 1									

PivotTable View

These correspondences can be pivot charted: S3-T (Tapieté) corresponding with the other four languages, along with their counts:



CV correspondences across all of the languages can be achieved: S3-M and S4-M is sorts all potential correspondences in the other four languages:

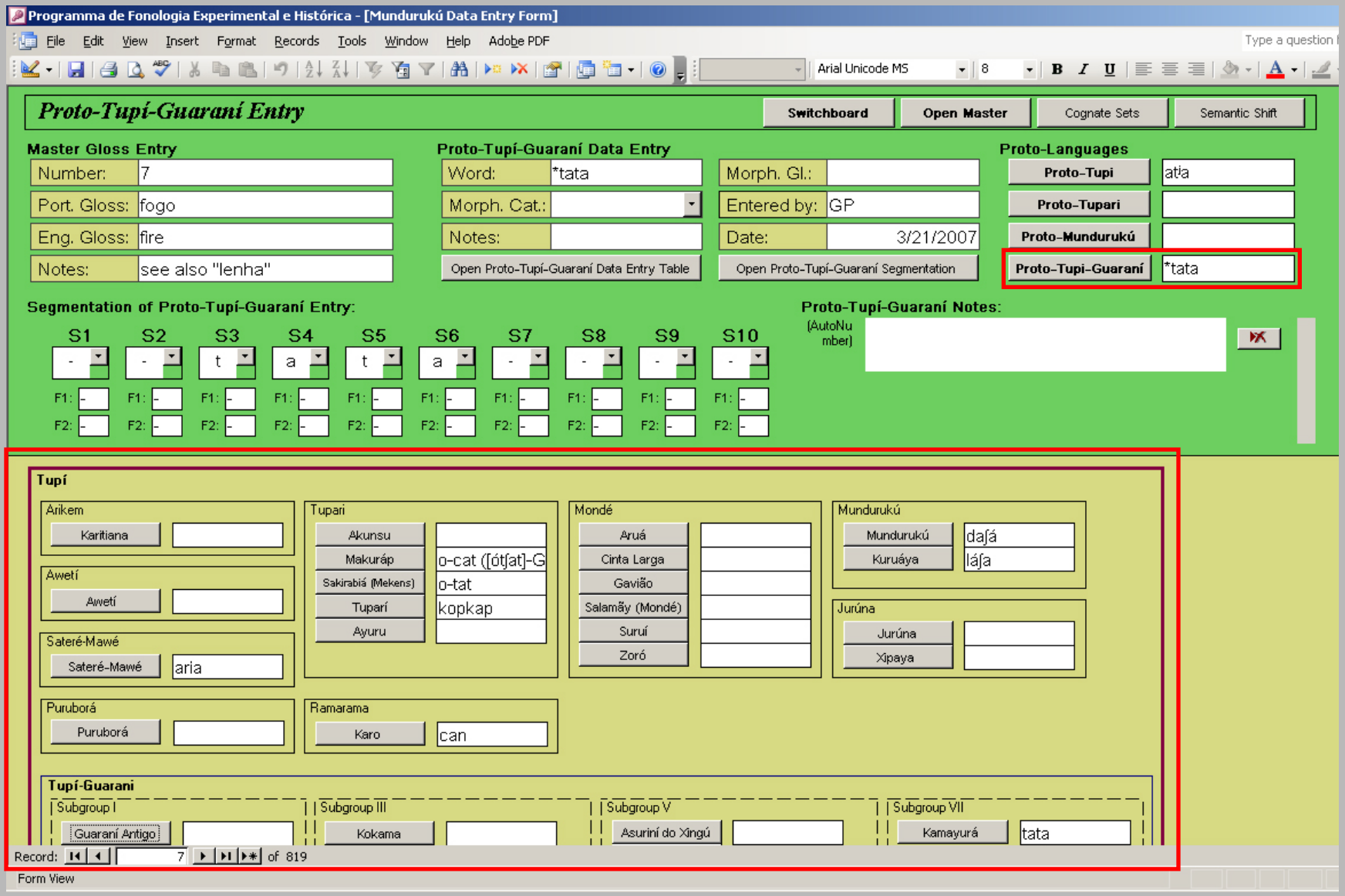
S3-M	S4-M	Mundurukú	S3-T	S4-T	Tapieté	S3-W	S4-W	Wayampi	S3-K	S4-K	Kamayurá	S3-A	S4-A	Avá-Canoeiro
p	ə													
	o	poʃí	p	o	pohɨ	p	o	poɨj / po(w)ɨj (WA)	p	o	i-powɨj	p	o	poɨj
	ó													
	Total													
s	ə́													
	Total													
t	a													
	Total													
ʧ	ó	ʧókõn	t	ũ	tũka				t	u	tukan	t	u	tukan
	Total													
t/d	a	táp, dáp, tap, dap	r/h	a	ha, ra	Ø	a	ap-ira	ʔ	a	ʔap	Ø	a	ap
		daʃíp, taʃíp	h	a	haku	Ø	a	aku	h	a	hakup	Ø	a	akup, i-akup
	i													
	í/i													
	o													
	Total													
ʧ/ʤ	ó	ʤó, ʧó	h	e	heʃa	Ø	e	esa	r	e	ne=reɬsak			a-mae
	Total													
t/n	ə́/ə̃													
	Total													
ʔ	á/a	ʔát, ʔat	Ø	a	a	ʔ	a	ʔa	ʔ	a	o-kuj, ʔat	Ø	ɨ	o-ɨʁ, o-jɨpɨra
	é													
	í/i													
	o													
	ó/o													
	Total													
ʔp/b	ə́/ə	(ʔ)pə, pə́, bə, bə́	p	o	po, te-po	p	o	po	p	o	hʷ̃ã, po	p	o	kʷ̃ã, i-po
	í													
	Total													
Grand Total														

PivotTable View

IV. Approaching phonemic reconstruction

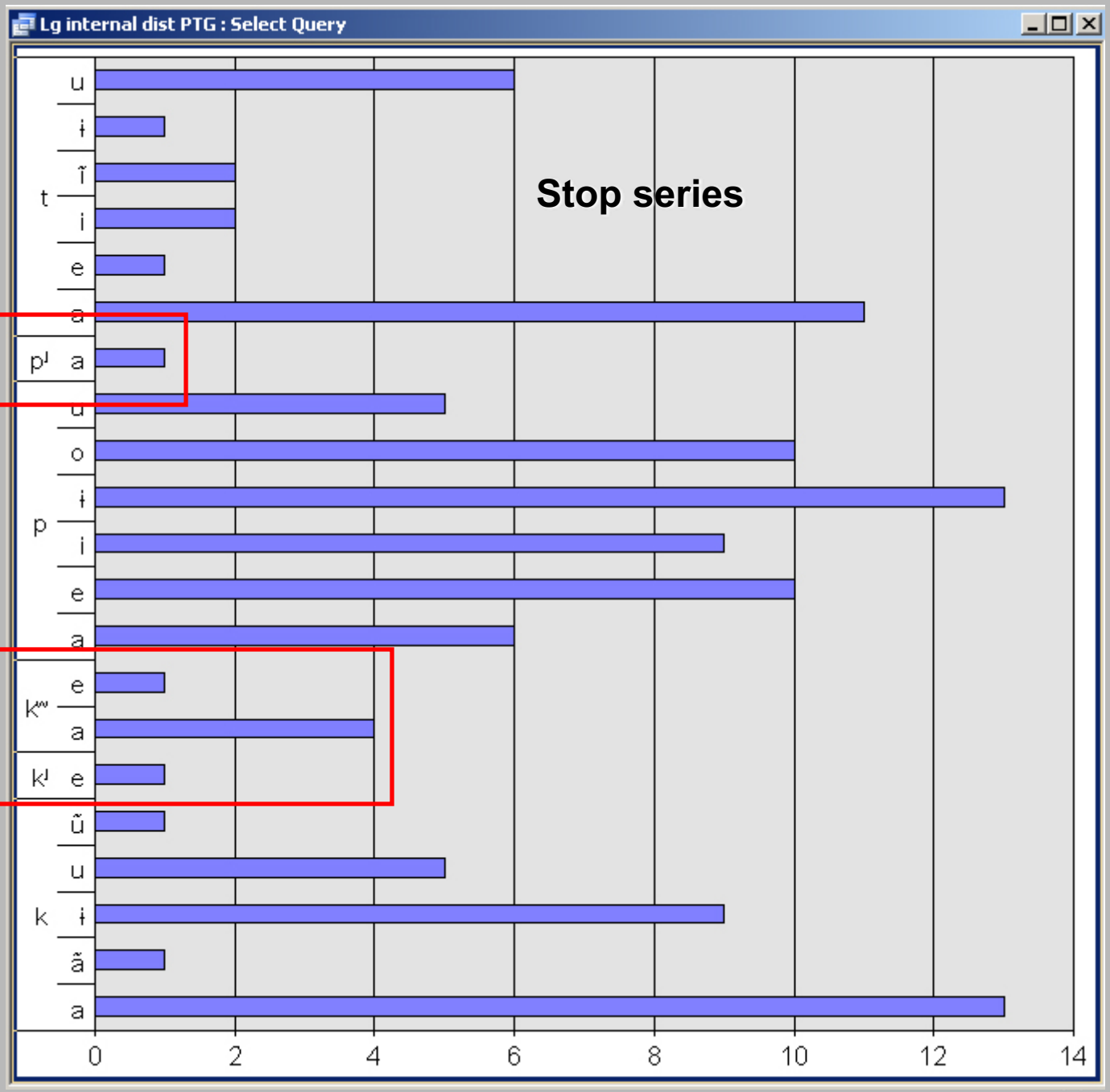
- With this methodology, we have a set of useful tools that are particularly applicable to tracking sound changes.
- Once the analyst works out a phonemic reconstruction based on the results obtained from the methods demonstrated above, they can be entered in their own proto-language data entry form.

Proto-language Data Entry Form:



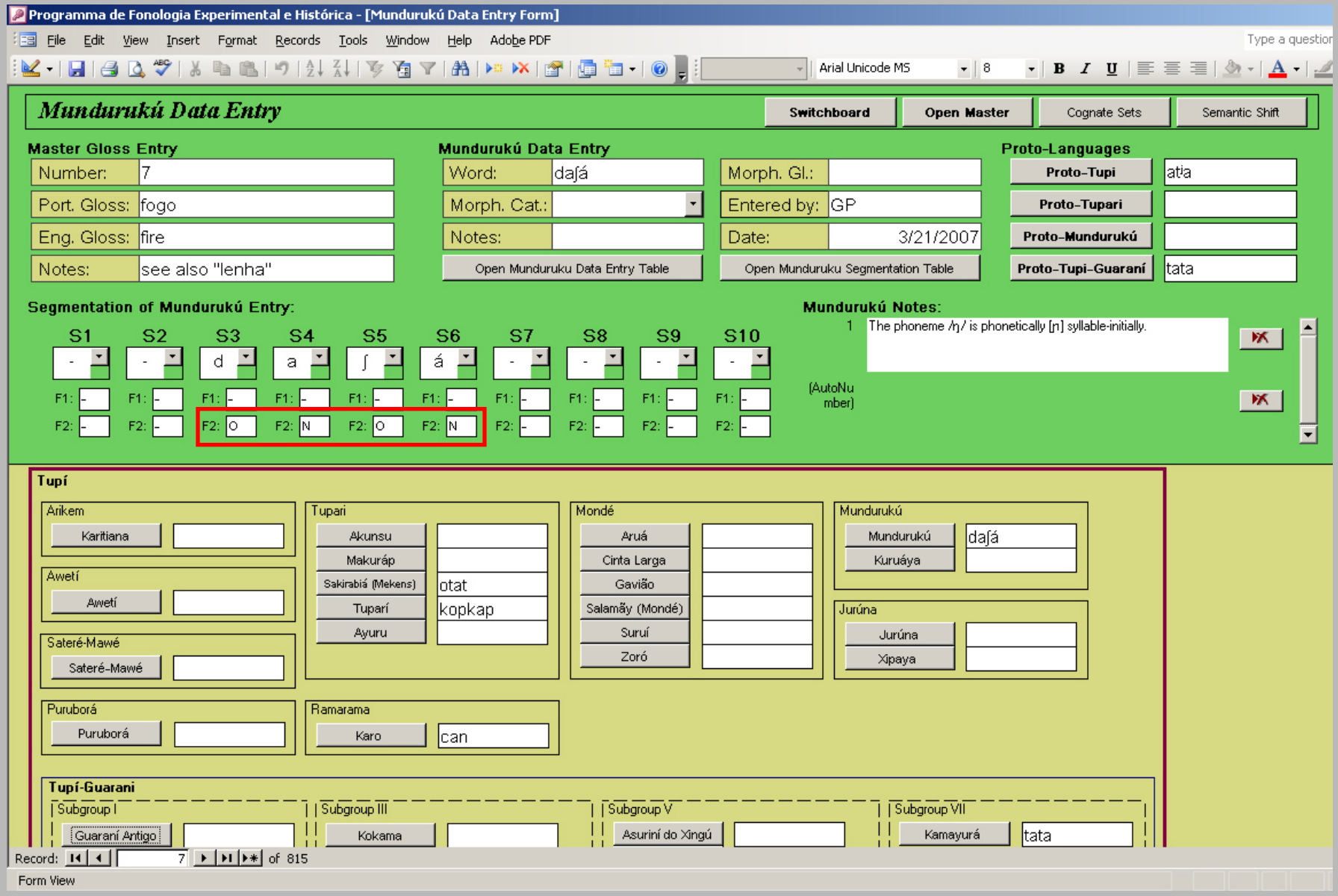
A pivot table and chart can be constructed in the same way as shown above:

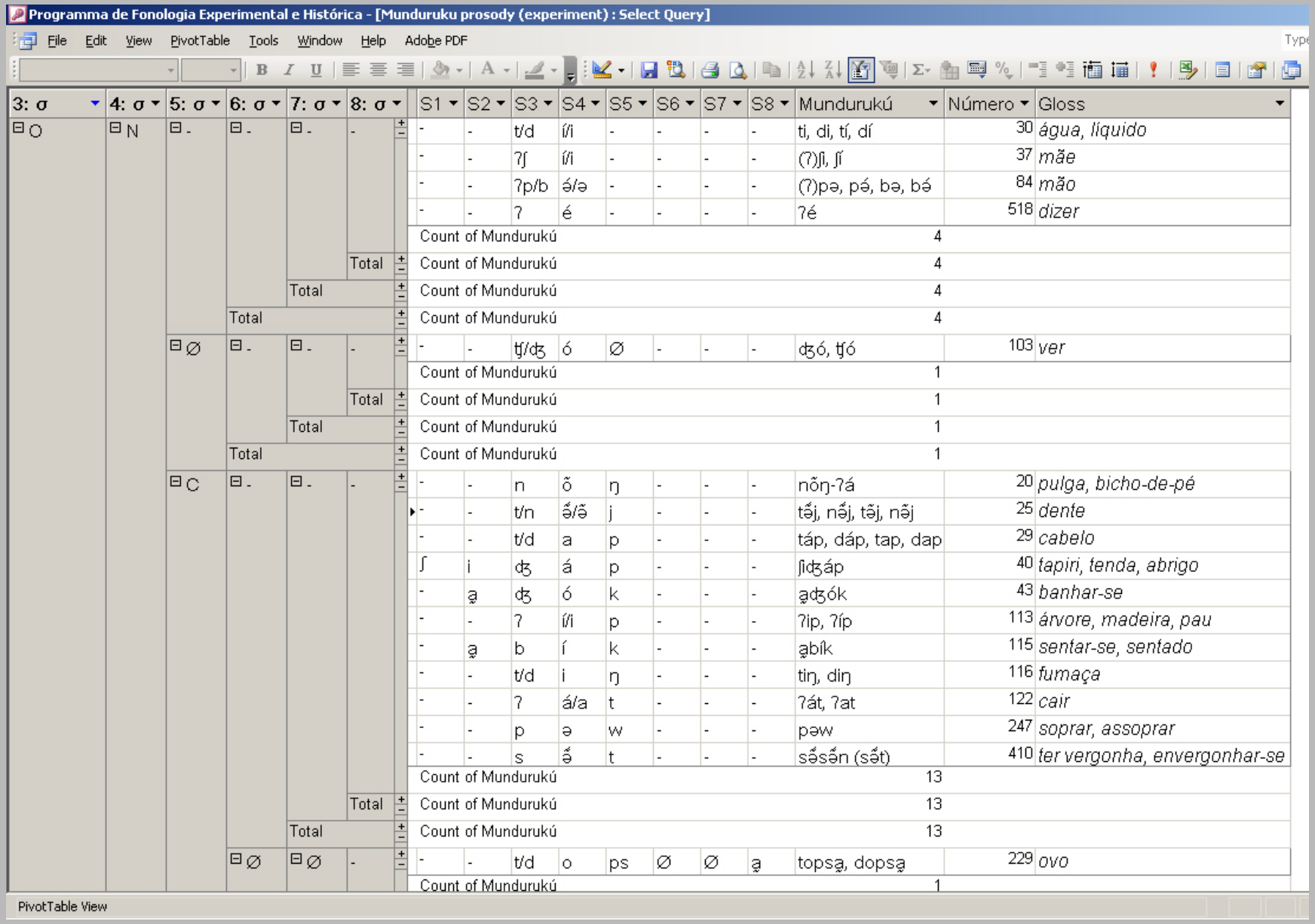
We can track the progress of our reconstruction, and check the language-internal distribution of the (proto-) segments which will aid in preventing possible cases of skewed occurrences.



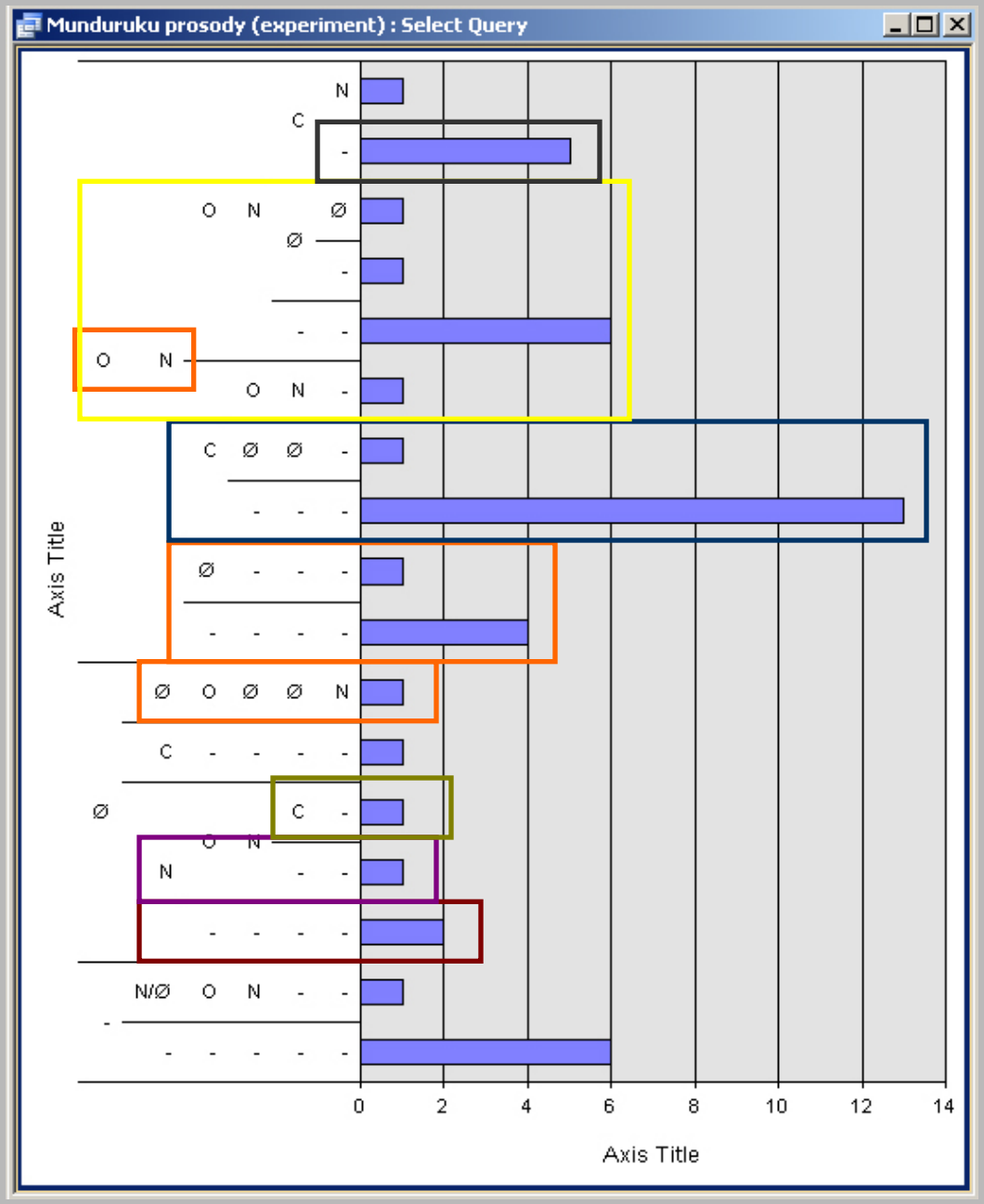
V. An experiment: Features

- Each segment has two fields associated with it (F1 and F2) that can store any type of feature the user finds relevant.
- This has been used to generate syllable templates and track the distribution of nasality.


Programma de Fonologia Experimental e Histórica - [Mundurukú Data Entry Form]
File Edit View Insert Format Records Tools Window Help Adobe PDF
Type a question
Arial Unicode MS
8
Mundurukú Data Entry
Switchboard
Open Master
Cognate Sets
Semantic Shift
Master Gloss Entry
Number: 7
Port. Gloss: fogo
Eng. Gloss: fire
Notes: see also "lenha"
Mundurukú Data Entry
Word: daʃá
Morph. Cat.:
Notes:
Open Munduruku Data Entry Table
Morph. Gl.:
Entered by: GP
Date: 3/21/2007
Open Munduruku Segmentation Table
Proto-Languages
Proto-Tupi atʲa
Proto-Tupari
Proto-Mundurukú
Proto-Tupi-Guaraní tata
Segmentation of Mundurukú Entry:
S1 S2 S3 S4 S5 S6 S7 S8 S9 S10
- - d a ʃ á - - - -
F1: F2: O F2: N F2: O F2: N
Mundurukú Notes:
1 The phoneme /ɲ/ is phonetically [n] syllable-initially.
(AutoNumber)
Tupí
Arikem
Karitiana
Awetí
Awetí
Sateré-Mawé
Sateré-Mawé
Puruborá
Puruborá
Tupari
Akunsu
Makuráp
Sakirabiá (Mekens) otat
Tuparí kopkap
Ayuru
Ramarama
Karo can
Mondé
Aruá
Cinta Larga
Gavião
Salamãy (Mondé)
Suruí
Zoró
Mundurukú
Mundurukú daʃá
Kuruáya
Jurúna
Jurúna
Xipaya
Tupí-Guarani
Subgroup I
Guarani Antigo
Subgroup III
Kokama
Subgroup V
Asuriní do Xingú
Subgroup VII
Kamayurá tata
Record: 7 of 815
Form View


Programma de Fonologia Experimental e Histórica - [Munduruku prosody (experiment) : Select Query]
File Edit View PivotTable Tools Window Help Adobe PDF
3: σ 4: σ 5: σ 6: σ 7: σ 8: σ S1 S2 S3 S4 S5 S6 S7 S8 Mundurukú Número Gloss
O N - - - - - - t/d í/i - - - - ti, di, tí, dí 30 água, líquido
- - ʔʃ í/i - - - - (ʔ)ʃi, ʃí 37 mãe
- - ʔp/b á/ə - - - - (ʔ)pə, pá, bə, bá 84 mão
- - ʔ é - - - - ʔé 518 dizer
Count of Mundurukú 4
Total Count of Mundurukú 4
Total Count of Mundurukú 4
Total Count of Mundurukú 4
Ø - - - - tʃ/dʒ ó Ø - - - dʒó, tʃó 103 ver
Count of Mundurukú 1
Total Count of Mundurukú 1
Total Count of Mundurukú 1
Total Count of Mundurukú 1
C - - - - n õ ŋ - - - nõŋ-ʔá 20 pulga, bicho-de-pé
- - t/n ə̃/ã j - - - tə̃j, nə̃j, tãj, nãj 25 dente
- - t/d a p - - - táp, dáp, tap, dap 29 cabelo
ʃ i dʒ á p - - - ʃidʒáp 40 tapiri, tenda, abrigo
- ą dʒ ó k - - - ądʒók 43 banhar-se
- - ʔ í/i p - - - ʔip, ʔíp 113 árvore, madeira, pau
- ą b í k - - - ąbík 115 sentar-se, sentado
- - t/d i ŋ - - - tiŋ, diŋ 116 fumaça
- - ʔ á/a t - - - ʔát, ʔat 122 cair
- - p ə w - - - pəw 247 soprar, assoprar
- - s ə̃ t - - - sə̃sə̃n (sə̃t) 410 ter vergonha, envergonhar-se
Count of Mundurukú 13
Total Count of Mundurukú 13
Total Count of Mundurukú 13
Ø Ø - - - t/d o ps Ø Ø ą topsą, dopsą 229 ovo
Count of Mundurukú 1
PivotTable View

N	=2
ON	=6
ONC	=14
ONON	=9
NON	=1
NONC	=1
ONONC	=5



Issues

- Difficult to track the overall distribution of segments. For example, to check the distribution or environment of all the vowels in a language, there is at present to straightforward way to collect all of the vowels across any and all Segment fields.
- Sometimes querying an environment is necessary (i.e. Boolean): it is tedious to find all the occurrences of, for example, CVC, or anything more specific.
- Some management/stability/memory issues.
- Sometimes difficult to interpret results.
- Although all of the language data itself is stored in text files to preserve its declarative format, MS Access itself is not cheap.

Conclusion

- BDEC-T presents a practical methodology that is easily and quickly implementable, and that makes use of a function that many people already have with their database or spreadsheet.

i. **Efficient:** data analysis is quick – pivot table and chart queries can be stored and easily printed.

ii. **Precise:** the user ultimately has control over the correspondences through the segmentation interface. Potential mergers and splits can be tracked.

iii. **Functional/Flexible:** any arrangement of Segment columns, features, glosses, words etc. can be related simultaneously across any number of languages.

iv. **Expandable:** new features can be added or old ones rearranged to suit evolving needs. Can be mounted on the internet and accommodate several different users.

Thank you!

Tyler Peterson
tylerrp@interchange.ubc.ca

Gessiane Picanço
picanco.g@hotmail.com