



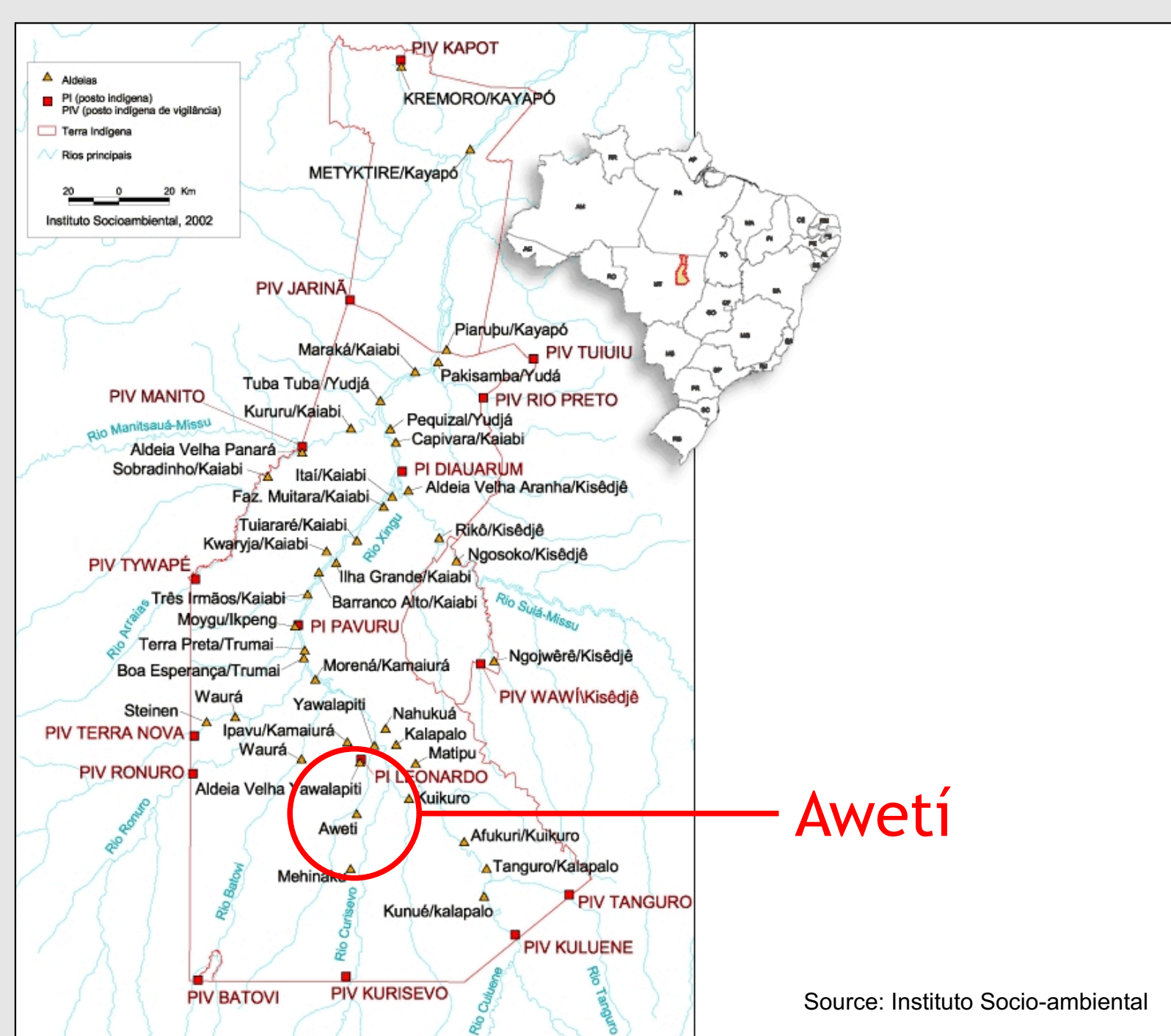
Awetí Language Documentation Project



The Awetí Project participated in the DOBES-programme from November 2000 until March 2006, under the responsibility of Prof. Dr. Hans-Heinrich Lieb and with Sebastian Drude as its principal researcher and Sabine Reiter as a second researcher. The project worked in close co-operation with two other DOBES-projects on Kuikuro and Trumai, situated in the same geographic and cultural area.

Although all the children born in the major villages still acquire Awetí as their first language, young people, which form the lion's share of the population, are particularly open-minded to Brazilian language and culture which is reinforced by mass media (television / radio broadcast). One consequence of this is the decreasing interest in learning the rich oral literature, still known by very few older specifically trained story-tellers. Preservation of the latter precious tradition is most important to the Awetí, thus much data of this type was collected.

During the project, the researchers have started to train the two Awetí teachers and other young men and women to record audio and video data in order to continue the documentation on their own in the future, and to write their language, also to help with the time consuming data-transcription.



Awetí is a Tupian language spoken by an indigenous community in the multilingual environment of the Upper Xingú, Central Brazil. The Awetí currently (2012) are approximately 190 persons, the vast majority of which lived in a single village until 2002/03. Now, there are three villages of 95, 65 and 30 persons, respectively.

Language / dialect	villages	population (source: ISA, 2011)	language family
• Trumai	4 (2002)	97 (<20 spks)	(isolated)
• Waurá	2 (2012)	409	Aruak (Maipure central)
• Mehinako	2 (2006)	254	
• Kustenuau	0	+	
• Yawalapiti	1 (2010)	156 (<10 spks)	
• Kuikuro	3 (2004)	522	Karib (South/Xingú)
• Kalapalo	2 (2002)	385	
• Matipu	2 (2003)	149	
• Nahukwa	1 (2002)	126	
• Kamayurá	2 (2006)	467	Tupi-Guarani
• Awetí	3 (2012)	195	Awetí (Tupi)

Lines without separating dashes refer to dialects or closely related languages. Those marked with a ● are documented directly (and those with a ○, indirectly) within DOBES

After the first contact (end of 19th century), the Awetí estimated population of 400 was reduced by several epidemic diseases to less than 25 members in the 1950s. Despite the late considerable increase, their overall number is still low, and now there are other factors contributing to the endangerment of their language and culture.



Sebastian Drude with the father of the chief and main informant, recording narratives (2002)

A fair and respectful cooperation between the Awetí and the documentation team was of utmost importance for the project. Thus, in 2003 the chief Yakumin was invited to Europe in order to provide a better understanding of the project's institutional background and working procedures.

In addition to the work directly related to the documentation itself, the Awetí project has made some valuable methodological and technical contributions within the context of DOBES as one of the pioneer attempts to unify language documentation on a larger scale. The maximal linguistic annotation system of Advanced Glossing has been proposed by Lieb and Drude in 2001, and Drude has developed a Shoebox-Set-up, based on Advanced Glossing, for interlinearising and annotating texts with linguistic data of a broad range of systematic types, from phonetic/phonological, morphological and syntactic to semantic.

There have been several follow-ups growing out of the DOBES project on Awetí. Sabine Reiter has completed a PhD thesis about ideophones in Awetí, and Sebastian Drude has held a Diltthey-Fellowship with support from the VolkswagenStiftung, also dedicated to the study of Awetí. He is now head of The Language Archive, at the Max-Planck-Institute for Psycholinguistics in Nijmegen, which has as one of its roots the technical work for DOBES. Most importantly, a local Archive on Amazonian Languages has been established at the Museu Goeldi in Belém, Pará, which was also largely motivated by the several DOBES project based at Museu Goeldi. Sebastian Drude was involved in this, and also in the establishing of another Archive at the Museu do Índio in Rio de Janeiro.

Glossing Tables

Syntactic Glossing Table

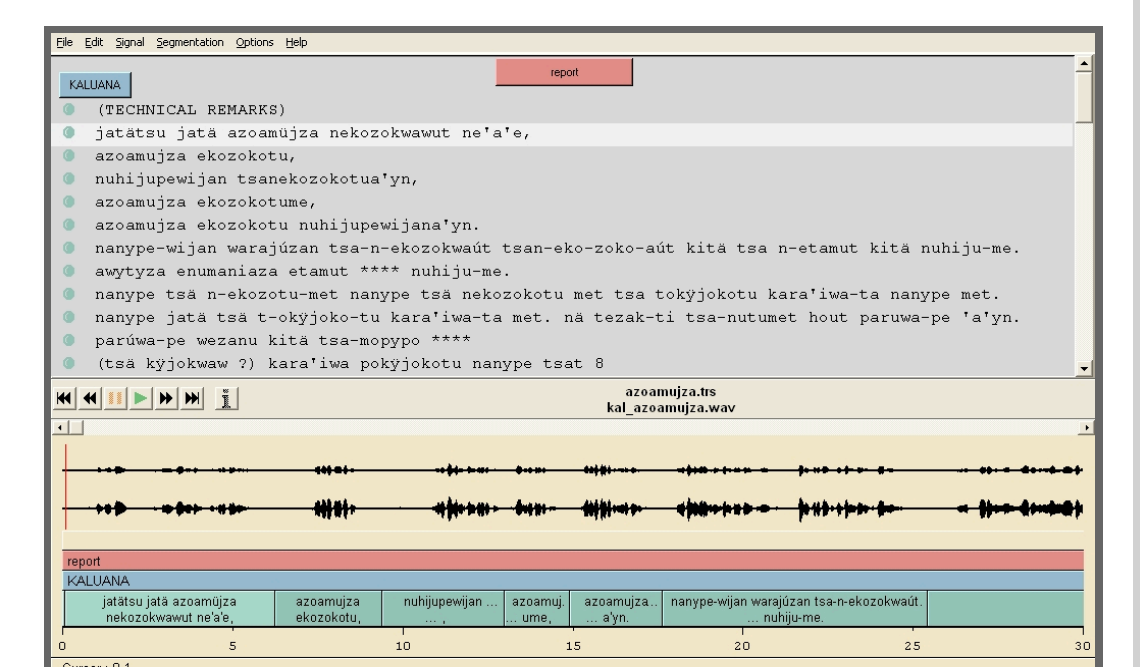
I	1	2	3
II	[di.]	ʔonéy.bD.zx.tʔ(ñ)	pro.blec.mə]
III	L	H ₁ M ₁ L H H H	H H ₁ L ₁
IV	/ di. /	/ ʔon.ʔyy.ba.t-zix.t-ixɔ.n /	/ pro.blec.mə /
V	L	H L L H H H	H H ₁ L
VI	die	unübersichtlichen	probleme
VII	DefArt	Adj	Sub Neut
VIII	Nom Pl Umm ₂ Str	Umm ₂ Pl Umm ₂ Wk	Nom Pl Umm ₂
IX	Nom Pl Def	'involved'	'problem'
X	1,3: NF	2: NF	1,2,3: NGr
XI	mod: 2 1,3		
XII	Die unübersichtlichen Probleme.		
XIII	E: The involved problems.		

Morphological Glossing Table

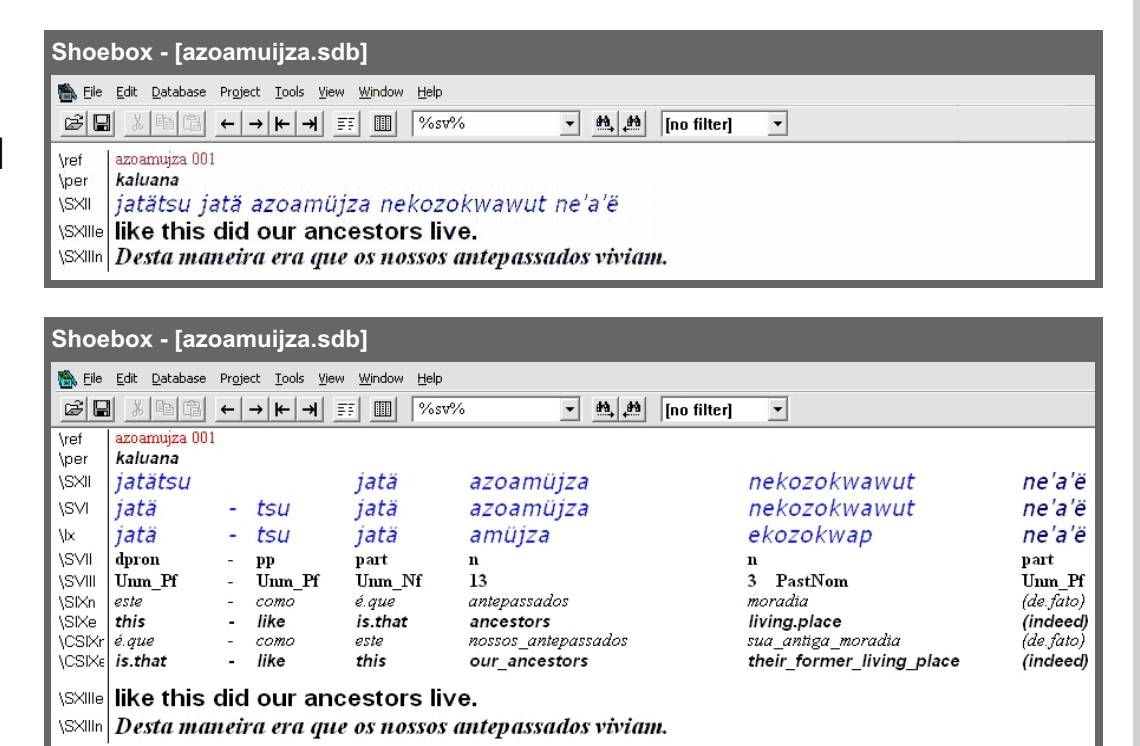
I	1	2	3	4	5
II	/ʔon.	ʔyy.ba.r.	zix.t-	ix	ɔ.n /
III	H	H.L L	H.L	H.L	L
IV	/ʔon. /	/ ʔyy.ba.t /	/ zix.t /	/ ix /	/ ɔ.n /
V	H	H.L L	H.L	H.L	L
VI	un	über	sicht	lich	en
VII	Pref ₁	PrepSt ₁	SubSt ₁	SubSt ₁ /AdjSt ₁	AdjFlex ₁
VIII	—	—	—	—	—
IX	not	'over'	'view'	suitable-for	Umm ₂ Pl Umm ₂ Wk
X	1: Af	2: Stf	3: Stf	4: Af	5: Af
XI	m-mod: 2 3	m-qual: 4 2 3 3	m-mod: 1 2 3 4	m-qual: 5 1 2 3 4	1 2 3 4 5: StGr
XII	unübersichtlichen				
XIII	'involved'				

Shoebbox

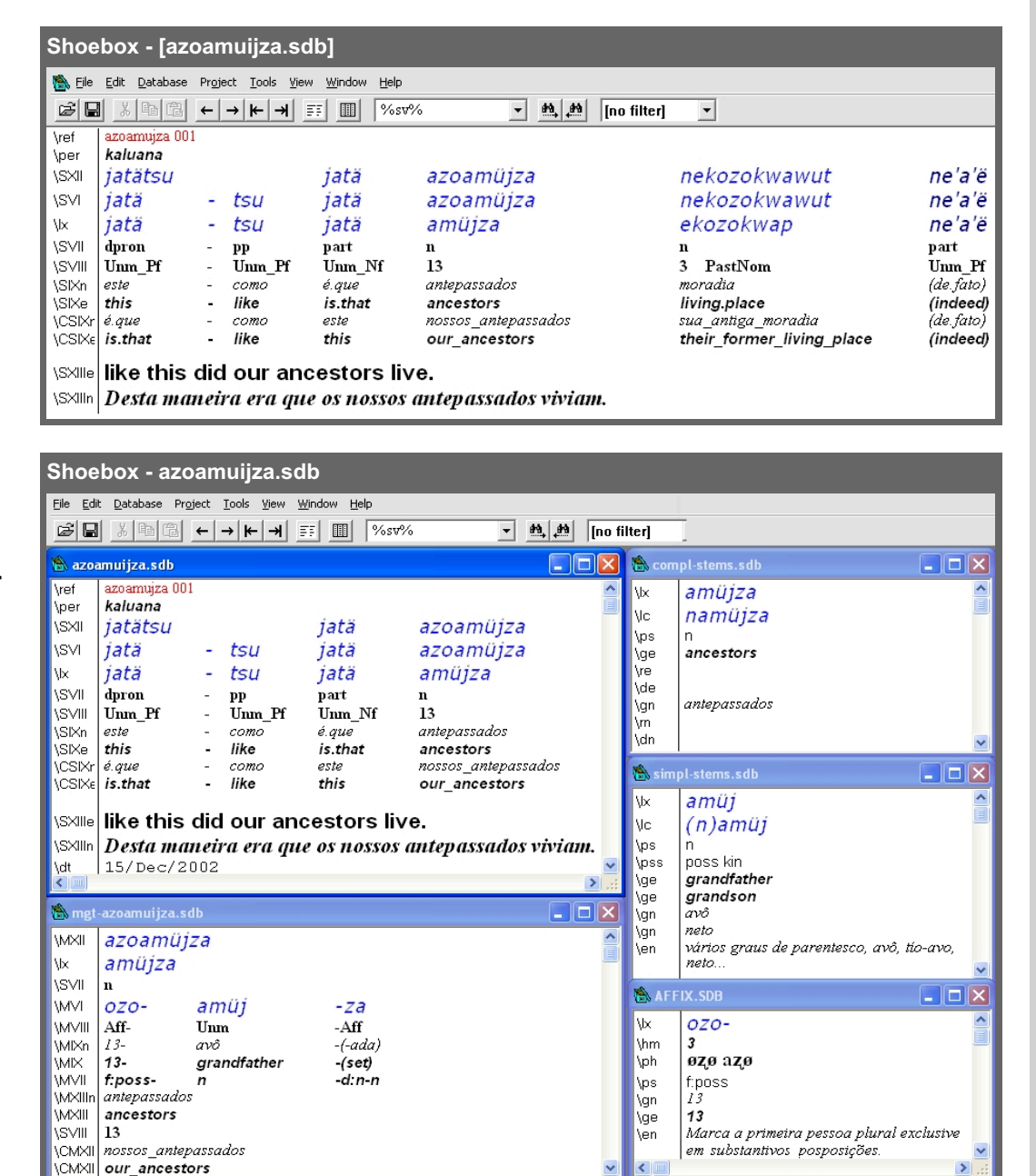
A: Transcription in Transcriber



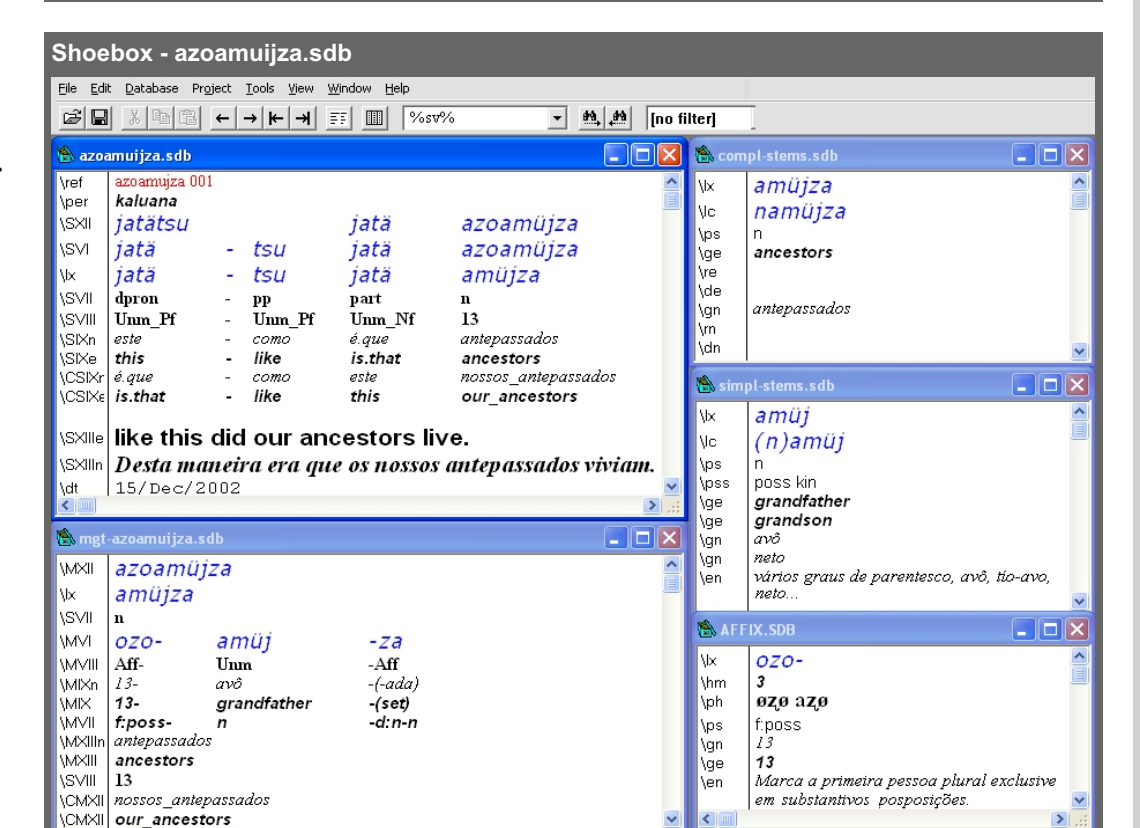
B: Converted to Shoebbox, translations added



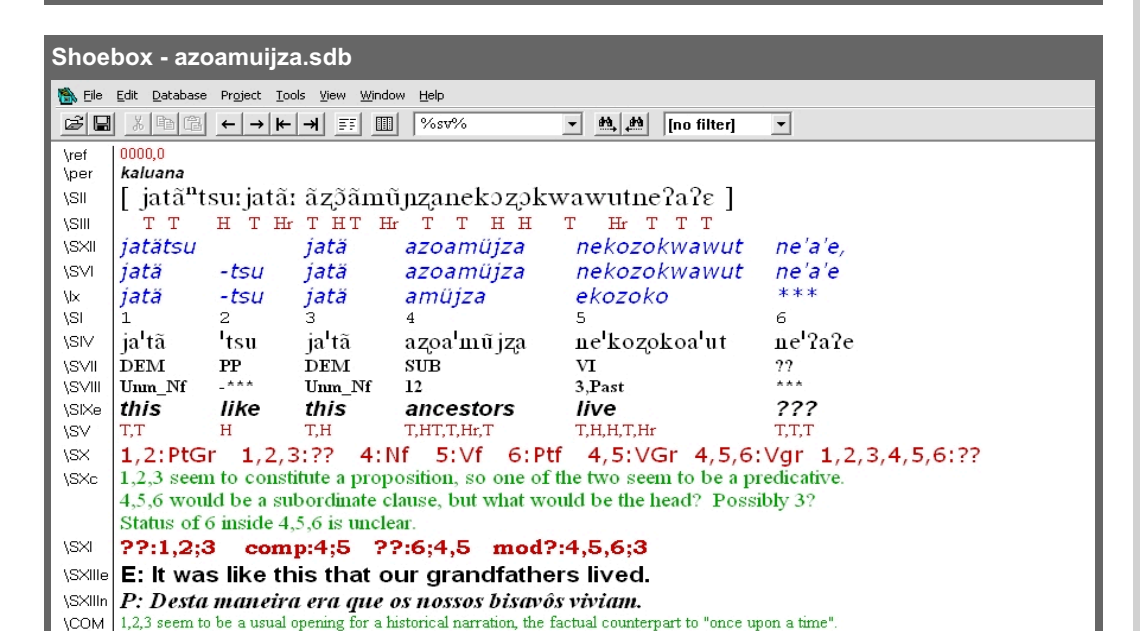
C: Wordwise interlinearized



D: The Shoebbox-Setup (syntactic and morphological glossing tables, lexical databank for complex and simple words and affixes)



E: A complete syntactic glossing table in Shoebbox



Project data obtained and processed

Category	sessions	audio files #	audio (h:mm)	video files #	video (h:mm)	transcript (-h:mm)	translation (-h:mm)
A) linguistic data: sum	131	385	97:43	112	24:51	3990	26:30
1. dialogue	11	18	3:57	9	2:52	200	0:45
2. elicitation task	9	22	5:38	2	0:33	0:45	0:45
3. word list	7	17	3:52	27	3:49	230	2:30
4. monolingual text: sum	104	338	84:18	74	17:37	33:45	22:30
a. biographic	5	12	1:27	13	1:31	1:00	0:45
b. description	10	21	2:10	13	2:20	2:00	1:45
c. explanation	9	14	2:11	12	4:34	1:30	1:15
d. narratives	74	288	77:03	21	7:23	27:15	18:15
e. personal	8	15	2:25	15	1:50	0:45	0:30
B) cultural data: sum	99	52	14:52	227	40:25	0:30	0:15
1. artefact+houses	14	0	0:00	35	6:02	0:00	0:00
2. culture contact	15	0	0:00	27	3:43	0:00	0:00
3. food	13	0	0:00	48	9:53	0:00	0:00
4. miscellaneous	27	1	0:19	57	4:33	0:00	0:00
5. special events: sum	20	51	14:33	60	17:13	0:30	0:15
a. jwan2002	21	26	8:08	37	9:15	0:30	0:15
a. jwan2002	9	25	6:27	23	7:59	0:00	0:00
TOTAL	230	437	112:35	339	65:16	39:30	28:45

Obs.: In most cases there is an audio and a video file of the same session. Due to better sound quality transcriptions were made from audio files. In general, there will not be any transcription of cultural data, except for ritualized speech at certain recurrent events.

CONTACT ADDRESS

Sebastian Drude
The Language Archive, MPI
P.O. Box 310
6500 AH Nijmegen - The Netherlands
Sebastian.Drude@mpi.nl