

## Awetí Language Documentation Project



The Awetí Project has participated in the DOBES-programme since November 2000, under the responsibility of Prof. Dr. Hans-Heinrich Lieb and with Sebastian Drude as its principal researcher. In April 2002 (main phase), Sabine Reiter permanently joined the team as a second researcher. The project works in close co-operation with two other DOBES-projects on Kuikuro and Trumai, situated in the same geographic and cultural area.



Awetí is a Tupian language spoken by an indigenous community in the multilingual environment of the Upper Xingú, Central Brazil. The Awetí currently are approximately 130 persons, the vast majority of which lived in a single village until 2002/03, when a new village of about 40 members was founded.

Peoples and languages of the Upper Xingú			
Language / dialect	villages (2004)	population (source: ISA, 2002)	language family
● Trumai	2	120 (<30 spk)	(isolated)
● Waurá	1	321	
○ Mehinako	1	199	
○ Kustenuá	0	†	
○ Yawalapiti	1	208 (<20 spk)	
● Kuikuro	3	415	Karib
○ Kalapalo	3	417	
○ Matipu	1	119	
○ Nahukwa	1	105	
● Kamayurá	2	355	Tupi-Guarani
● Awetí	2	138	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.

Probably most pressing is the growing importance of Portuguese which is used not only with people from outside the reservation (government officials, NGOs' members and researchers, trading partners in Brazilian towns, etc.) but nowadays also with other ethnic communities within the park.

Although all the children born in the village 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 discourse tradition (myths and other narratives, representing the memory of the Awetí people and foundation of their culture), still known by very few older specifically trained story-tellers.

In the last years 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 data-transcription which is the main task with the remaining time of the project. For the cultural and linguistic material obtained so far see the diagram on corpus structure. Linguistic data is chosen in agreement with the two other Xingú projects in order to obtain analogue corpora for comparative studies, including a broad variety of texts / discourse types, ranging from informal and formal conversation, ritualized speech, task and place descriptions, autobiographic narratives to myths and historical narratives. Preservation of the latter precious tradition is most important to the Awetí, thus more data of this type will be collected.



Awetí has markedly distinct male and female varieties, differing mainly in the use of pronouns and discourse particles. Accounting for that, it has proven most favourable that S. Drude works primarily with male and S. Reiter with female informants.

A fair and respectful cooperation between the Awetí and the documentation team is 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.

### Glossing Tables

#### Syntactic Glossing Table

I	1	2	3
II	idi.	ʔandy bCLzng; tɔʔfɔ	pro. blec.mo 1
III	L	H <sub>1</sub> M <sub>1</sub> L H H H	H H <sub>1</sub> L <sub>1</sub>
IV	/dii. /	/ʔam.ʔyɔ.ɓa.r.ʔɔʔɔ.ʔɔʔɔ.n /	/pɔn.blec.mo /
V	L	H L L H H H	H H <sub>1</sub> L
VI	die	unübersichtlichen	probleme
VII	DefArt	Adj	Sub Neut
VIII	Nom Pl Urmg, Str	Urmg, Pl Urmg, Wk	Nom Pl Urmg
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	/ʔam.	ʔyɔ.ɓa.r.	ʔɔʔɔ.	ɔɔ.	ɔ.n /
III	H	H <sub>1</sub> L	H <sub>1</sub> L	H <sub>1</sub> L	L
IV	/ʔam. /	/ʔyɔ.ɓa.r. /	/ʔɔʔɔ. /	/ɔɔ. /	/ɔ.n /
V	H	H <sub>1</sub> L	H <sub>1</sub> L	H <sub>1</sub> L	L
VI	usr	über	sicht	lich	en
VII	Pref.	ProgSt	SubSt	SubSt/AffSt	AdjFlex
VIII	-				
IX	not	'over'	'view'	'suitable-for'	Urmg, Pl Urmg, Wk
X	1: Af	2: Sff	3: Sff	4: Af	5: Af
XI	m-mod: 2 3	m-qual: 4 2 3	m-mod: 1 2 3 4	m-qual: 5 1 2 3 4	Sff 1,2,3,4,5: SffGr
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

Linguistic Data	Sessions	Audio files	Duration (h:m)	Transcribed	Translated	Video files	Duration (h:m)
Elicitation	9 (6)	13 (12)	0:01	8 word lists	722 lex. items	9	0:18
Conversations/Dialogues	0	0	0:00	0:00	0:00	0	0:00
Biographies	4	11	01:27	01:15	01:15	10	01:25
Narratives	37	162	03:33	03:10	02:35	20	03:02
Procedural descriptions	7	16	01:20	00:29	00:29	12	01:21
Explanations	11	23	01:55	00:27	01:37	11	01:47
Interviews	1	1	00:05	-	-	-	00:05
				12:23	09:00		

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

Museu Paraense  
Emílio Goeldi Linguística, CCH  
Caixa Postal 399  
66040-170 Belém - PA - BRAZIL  
aweti@zedat.fu-berlin.de