## A. CLARIN-Menü



Über das CLARIN-Menü können die Web-Services, WebMaus und WebLicht aufgerufen werden.

## Clarin > WebMaus...

Die Anwendung von WebMaus ermöglicht eine vollautomatische Segmentierung und ein gleichzeitiges Labeln von Sprachaufnahmen in einzelne Wörter und/oder Phoneme.

**Schritt 1**: Wählen Sie das/die zu segmentierende(n) Ereignis(se) oder mehrere Segmentkettenketten in der entsprechenden Spur aus und öffnen Sie CLARIN > WebMaus.

		Abb. 1							
00:00	00:01	0	0:02		00:03				
		all a la l			(joint) and				
			(1), (1) why		(intel) and				
۲. III.		557	ala.	22 1					
🕂 Add event	Append interv	al							
START 1 [	00:01.4]		2 [00:03.1]	3 [00:03	4 [00:04				
KLA [v] G	uten Tach, mein Name	is Schneider.	Ich such ne A	ur beit.					
4	III								

Schritt 2: In diesem Dialogfenster können Sie nun folgende Einstellungen vornehmen:

- Language: An dieser Stelle wird festgelegt, welcher Sprache die in der Transkription beschriebenen verbalen Handlungen zugeordnet werden können.
- Selection: Die Festlegung dieses Parameters beeinflusst, welche Ereignisse in einer Transkriptionsspur durch WebMaus segmentiert werden sollen.
  - Events of the current selection: nur die markierten Ereignisse innerhalb einer Segmentkette werden segmentiert.
  - Segment chain around the current selection: alle Ereignisse innerhalb einer Segmentkette werden segmentiert.
- Segmentation: Der Web-Service WebMaus kann nur Texte verarbeiten, in denen Zeichen(folgen) vorkommen, die mit den Standard-Zeichen(folgen) der Schriftsprache übereinstimmen<sup>1</sup>. Je nach Wahl



des Parameterwertes, wird der Transkriptionstext von bestimmten Zeichen(folgen), die von den Standard-Zeichen(folgen) der Schriftsprache abweichen, bereinigt.

- Use the transcription as is: alle Zeichen(folgen), aus denen der Transkriptionsext besteht, werden an *WebMaus* weitergeleitet. Sollte der

<sup>&</sup>lt;sup>1</sup> Die Funktionsweise von *WebMaus* wird in diesem Video beschrieben: <<u>https://www.youtube.com/watch?v=7II-gOShtFA</u>>.

Text nicht den Standard-Zeichen(folgen) der Schriftsprache entsprechen, führt dies zu einer abweichenden Segmentierung Ereignisse.

- Use segmentation algorithm: Wird dieser Parameterwert gewählt, werden die Zeichen(folgen) vor der Weiterleitung an *WebMaus* zunächst von einigen der Zeichen(folgen) bereinigt, die von den Standard-Zeichen(folgen) der Schriftsprache abweichen. Der bereinigte Text wird dann zur Verarbeitung an *WebMaus* weitergeleitet.
- Annotation Levels: definiert ob Ereignisse als Wörter und/oder Phoneme annotiert werden sollen.
  - Words (orthographic): die Artikulation eines Wortes wird mithilfe der orthographischen Schreibweise beschrieben.
  - Words (SAMPA): die Artikulation eines Wortes wird mithilfe des phonetischen Alphabets *SAMPA* beschrieben.
  - Phonems: die Artikulation eines Phons wird mithilfe des phonetischen Alphabets *SAMPA* beschrieben.
- Output:
  - Open it as a new file: wird dieser Parameter gewählt, so wird eine neue Transkription erstellt. Diese besteht aus einer Spur, in der sich nur die ausgewählte Ereigniskette befindet, sowie aus den neu angelegten Annotationsspuren Außerdem wird eine neue Audio-Datei erstellt und mit der Transkription verknüpft. Diese Audio-Datei enthält nur den Abschnitt der ursprünglichen Audio-Datei, der der ausgewählten Ereigniskette zugeordnet ist.
  - Merge it with the existing transcription: bei der Anwendung der Funktion automatisch erstellten Annotationsspuren werden zur bestehenden Transkription hinzugefügt.

Image: Add event Image: Append interval Image: Test and test												*]	]*				11								
	START	1 [00:01	2 [00:01	3 [00:01	4 [00:01	5 [00:01	6 [00:01	7 [00:01	8 [00:01	9 [00:02	10 [00:0:	11 [00:0:	12 [00:0:	13 [00:0:	14 [00:0	15 [00:0:	16 [00:0	17 [00:0:	18 [00:0:	19 [00:0:	20 [00:0:	21 [00:0]	22 [00:0]	23 [00:0]	24 [00
AMT [v]	Guten Tag!																								
KLA [v]		Guten	ten Tach, mein Name is Schneider.																						
KLA (p-sampa)		<p:></p:>	g	u:	t	@	n	t	a	x	m	aI	n	n	е	Ι	m	Ι	S	S	n	aI	đ	6	Ι
KLA [w-sampa]			gu:t@n				tax			maIn			neIm				?Is		SnaIdó				?IC		
KLA [w-orth]			Guten					Tach			mein			Name				is		Schneider				Ich	

Beispiel folgt der Einstellung in Abb. 2