## CLARIN-Menü



Über das CLARIN-Menü können Web-Services, die im Rahmen der Forschungsinfrastruktur *CLARIN* bereitgestellt werden, ausgeführt werden, um die Datei, die der derzeit im Partitur-Editor geöffneten Transkription zugrunde liegt, auf eine jeweils „Web-Service“-spezifische Weise zu verändern.

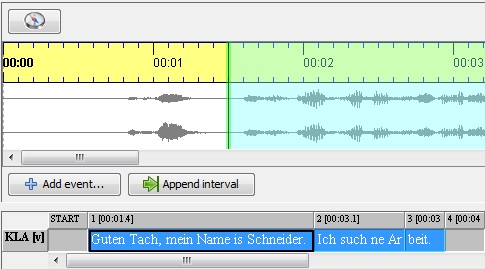
*CLARIN* ist eine Forschungsinfrastruktur, an deren Aufbau verschiedene Forschungszentren beteiligt sind und in deren Rahmen Linguistische Daten, Werkzeuge und Dienste für die Fachdisziplinen der Geistes- und Sozialwissenschaften hervorgebracht und bereitgestellt werden sollen.[[1]](#footnote-1)

Im Allgemeinen sind Web-Services Softwareanwendungen, die über ein Netzwerk bereitgestellt werden und die verwendet werden können, um Dateien automatisch auf eine jeweils „Web-Service“-spezifische Weise zu verändern. Im Besonderen können mithilfe der Web-Services, die über das CLARIN-Menü ausgeführt werden können, die Dateien, die den mit dem Partitur-Editor erstellten Transkriptionen zugrunde liegen, automatisch auf eine jeweils „Web-Service“-spezifische Weise verändert werden.

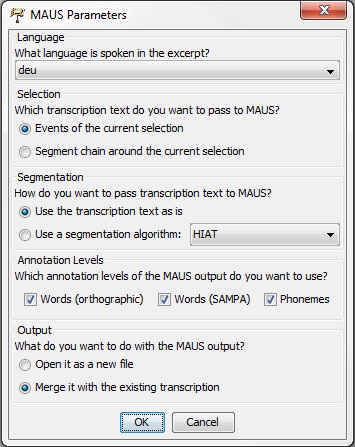
### Clarin > WebMaus…

Öffnet einen Dialog, über den die Ausführung des Web-Services *WebMaus* eingeleitet werden kann. *WebMaus* ist ein Webservice, der vom *CLARIN*-Zentrum BAS (Bavarian Archive for Speech Signals) entwickelt und bereitgestellt wird.[[2]](#footnote-2) Mithilfe von diesem Web-Service können Sie eine ausgewählte Segmentkette[[3]](#footnote-3) der im Partitur-Editor geöffneten Transkription automatisch in einzelne Ereignisse zerlegen und diese dann alignieren lassen. Dabei kann zwischen verschiedenen Graden der Feinheit bei der Zerlegung und Alignierung gewählt werden: die Segmentkette kann in alignierte Ereignisse, in denen jeweils die Artikulation einer Wortes beschrieben wird, oder in alignierte Ereignisse, in denen jeweils die Artikulation eines Phons beschrieben wird, zerlegt werden. Die Ereignisse, in die die Segmentkette zerlegt wird, werden in einer Annotationsspur erstellt. Diese wird bei der Anwendung von *WebMaus* automatisch erstellt und automatisch der Spur zugeordnet, in der sich die zerlegte Segmentkette befindet.

Die Ausführung dieses Web-Services soll nun veranschaulicht werden. Betrachten wir dazu die unten abgebildete Segmentkette.



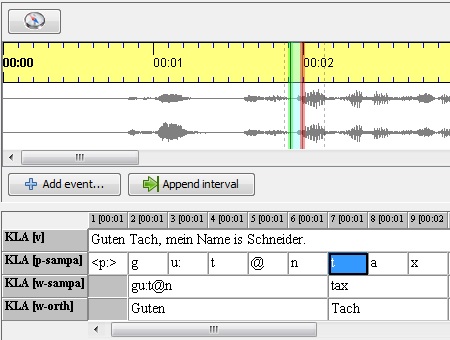
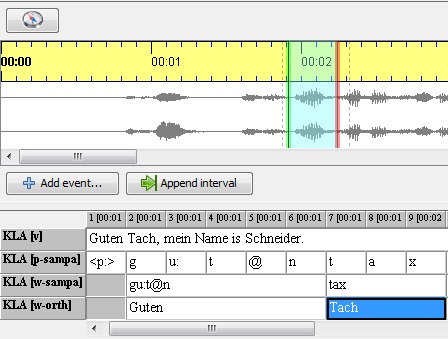
Diese Segmentkette besteht aus drei alignierten Ereignissen. Die drei Ereignisse unterscheiden sich im Hinblick darauf, wie umfangreich die jeweils mit ihnen beschriebenen verbalen Handlungen sind: mit dem dritten Ereignis wird die Artikulation einer Silbe beschrieben; mit den ersten beiden Ereignissen wird jeweils die Artikulation von Folgen von Silben beschrieben. Mithilfe des Web-Services *WebMaus* kann diese Segmentkette nun zerlegt werden in alignierte Ereignisse, in denen jeweils die Artikulation eines einzelnen Phons oder die Artikulation eines einzelnen Wortes beschrieben wird.

Wenn die Nutzung des Web-Services *WebMaus* eingeleitet wird, dann öffnet sich der unten abgebildete Dialog.

In diesem können sie folgende Parameter für die Verarbeitung der Transkription durch *WebMaus* festlegen:

* 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* zerlegt werden sollen.
  + Events of the current selection: Wenn dieser Parameterwert ausgewählt wird, dann werden nur die markierten Ereignisse durch *WebMaus* zerlegt.
  + Segment chain around the current selection: Wenn dieser Parameterwert ausgewählt wird, dann werden nicht nur die markierten Ereignisse durch *WebMaus* zerlegt, sondern gegebenenfalls all die Ereignisse, die mit den markierten Ereignissen eine nicht-unterbrochene Ereigniskette bilden.
* Segmentation: Der Web-Service *WebMaus* kann nur Texte verarbeiten, in denen Zeichen(folgen) vorkommen, die mit den Standard-Zeichen(folgen) der Schriftsprache übereinstimmen[[4]](#footnote-4). In den Texten, die den Ereignissen einer Transkription zugeordnet sind, werden aber manchmal Zeichen(folgen) verwendet, die von den Standard-Zeichen(folgen) der Schriftsprache abweichen. Die Festlegung dieses Parameters beeinflusst, ob und auf welche Weise der Text, der den zu zerlegenden Ereignissen zugeordnet ist, für die Verarbeitung durch *WebMaus* vorbereitet wird: Je nach Wahl des Parameterwertes, wird er oder wird er nicht von bestimmten Zeichen(folgen), die von den Standard-Zeichen(folgen) der Schriftsprache abweichen, bereinigt.
  + Use the transcription as is: Wird dieser Parameterwert gewählt, werden alle Zeichen(folgen), aus denen der Text besteht, zur Verarbeitung an *WebMaus* weitergeleitet. Dies führt dann zu einer vom Gewünschten abweichenden Zerlegung der zu zerlegenden Ereignisse, wenn der ihnen zugeordnete Text Zeichen(folgen) enthält, die von den Standard-Zeichen(folgen) der Schriftsprache abweichen.
  + Use segmentation algorithm: Wird dieser Parameterwert gewählt, werden die Zeichen(folgen), aus denen der den Ereignissen zugeordnete Text besteht, 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: Zum einen beeinflusst die Festlegung dieses Parameters, wie umfangreich die alignierten Ereignisse sein sollen, in die die zu verarbeitenden Ereignisse durch *WebMaus* zerlegt werden. Zum anderen beeinflusst sie, wie die verbalen Handlungen, die jeweils in einem der durch die Zerlegung hervorgebrachten Ereignisse beschrieben werden, beschrieben werden.
  + Words (orthographic): Wird dieser Parameterwert gewählt, so werden die zu zerlegenden Ereignisse in Ereignisse zerlegt, in denen jeweils die Artikulation eines Wortes beschrieben wird. Außerdem wird die Artikulation dieser Worte dann mithilfe der orthographischen Schreibweise beschrieben.
  + Words (SAMPA): Wird dieser Parameterwert gewählt, so werden die zu zerlegenden Ereignisse in Ereignisse zerlegt, in denen jeweils die Artikulation eines Wortes beschrieben wird. Außerdem wird die Artikulation dieser Worte dann mithilfe des phonetischen Alphabets *SAMPA* beschrieben.
  + Phonems: Wird dieser Parameterwert gewählt, so werden die zu zerlegenden Ereignisse in Ereignisse zerlegt, in denen jeweils die Artikulation eines Phons beschrieben wird. Außerdem wird die Artikulation dieser Phone dann 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 Annotationsspuren und den in ihnen angelegten Ereignissen, die bei der Anwendung dieser Funktion automatisch erstellt und angelegt wurden. 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: Wird dieser Parameter, so werden die bei der Anwendung der Funktion automatisch erstellten Annotationsspuren zur bestehenden Transkription hinzugefügt.

Die folgenden beiden Abbildungen zeigen das Ergebnis einer Ausführung des Web-Services *WebMaus* mit den oben abgebildeten Einstellungen. Die oberste Spur ist die Spur, aus der die Transkription vor der Anwendung des Web-Services bestand. Die unteren drei Spuren sind die Spuren, die durch die Anwendung des Web-Services erstellt und zur Transkription hinzugefügt worden sind.

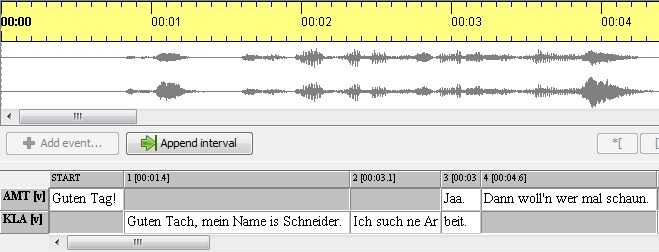


### Clarin > WebLicht…

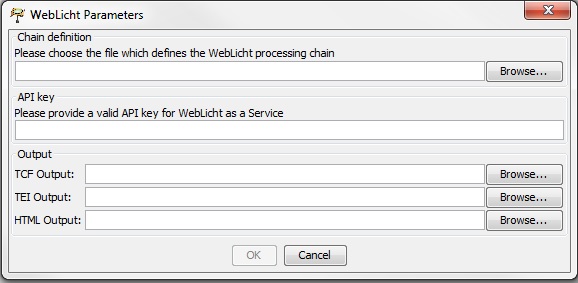
Öffnet einen Dialog, über den die Nutzung des Web-Services *WebLicht as a Service* (kurz: *WaaS*) eingeleitet werden kann*. WaaS* ist ein Web-Service, der von verschiedenen *CLARIN*-Zentren entwickelt und betrieben wird. Im Allgemeinen können mithilfe von *WaaS* die Dateien, aus denen ein digitales Textkorpus besteht, automatisch auf verschiedenen Annotationsebenen mit Informationen angereichert werden. Dem Web-Service *WaaS* liegen weitere Web-Services zugrunde. Den zugrundeliegenden Web-Services kommt während des Anreicherungsvorgangs die Rolle zu, die aufzubereitende Datei jeweils auf eine „Web-Service“-spezifische Weise zu verändern; dem Web-Service *WaaS* hingegen kommt während des Anreicherungsvorgangs die Rolle zu, den Einsatz der zugrundeliegenden Web-Services zu koordinieren.[[5]](#footnote-5) Mit welchen der grundlegenden Web-Services eine aufzubereitende Datei mit Informationen angereichert werden soll, legt ein Benutzer durch die Definition einer sogenannten „processing chain“[[6]](#footnote-6) fest.

Im Besonderen können mithilfe von *WaaS* die Dateien, die den mit dem Partitur-Editor gemäß den HIAT-Konventionen erstellten Transkriptionen zugrunde liegen, automatisch auf verschiedenen Annotationsebenen mit Informationen angereichert werden. Um dies tun zu können, wird neben einer geeigneten Transkriptions-Datei auch eine Datei, in der die auszuführende „processing chain“ definiert wird, benötigt.

Nun wird die Anwendung der Funktion „Weblicht…“ im Partitur-Editor mithilfe eines Beispiels veranschaulicht: Die Datei, die zu dem unten abgebildeten Abschnitt einer Transkription gehört, wird gemäß einer definierten „processing chain“ mithilfe von *WaaS* mit Informationen angereichert werden.

**

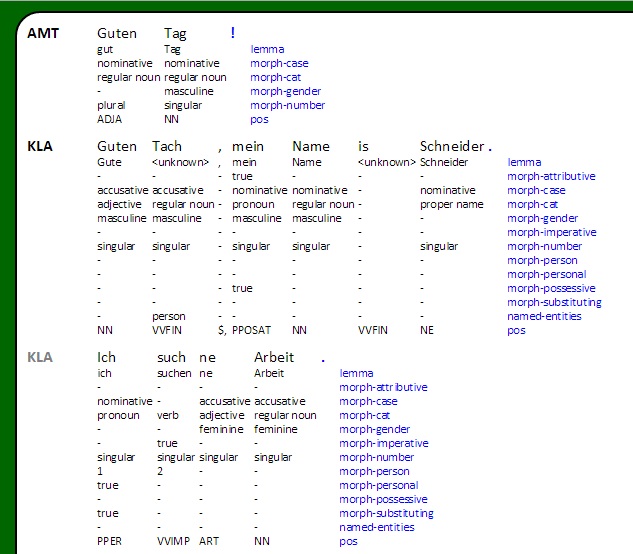
Wenn die Ausführung des Web-Services *WebLicht as a Service* eingeleitet wird, dann öffnet sich der unten abgebildete Dialog.

**

In diesem können Sie folgende Parameter für die Aufbereitung der Datei festlegen:

* Chain definition: An dieser Stelle kann festgelegt werden gemäß welcher „processing chain“ die Datei verändert werden soll. Dies kann festgelegt werden, indem der Speicherort einer Datei angegeben wird, in der die gewünschte „processing chain“ definiert wurde.
* API key[[7]](#footnote-7):
* TCF Output:
* TEI Output:
* HTML Output: An dieser Stelle kann festgelegt werden, ob durch die Ausführung der Funktion „WebLicht…“ eine HTML-Datei erstellt und an welcher Stelle diese Datei gespeichert werden soll.

Auf der folgenden Abbildung ist ein Ausschnitt aus der HTML-Datei zu sehen, die durch die Ausführung der Funktion „WebLicht…“ mit den oben beschriebenen Einstellungen erstellt wurde.

**

1. Einen umfassenden Überblick über das Forschungsprojekt erhält man auf der Projekt-Homepage:

   <<http://www.clarin-d.de/de/>>. [↑](#footnote-ref-1)
2. Einen umfassenden Überblick über die vom *CARIN*-Zentrum BAS bereitgestellten Softwareanwendungen erhält man auf der BAS-Homepage:

   <<http://clarin.phonetik.uni-muenchen.de/BASWebServices/#/services>>. [↑](#footnote-ref-2)
3. Eine Segmentkette ist eine zeitlich ununterbrochene Folge von Ereignissen in einer Spur der Transkription. Im Abschnitt „Was wird segmentiert“ im Anhang B des hier vorliegenden Handbuches wird erläutert, was eine Segmentkette ist. [↑](#footnote-ref-3)
4. Die Funktionsweise von *WebMaus* wird in diesem Video beschrieben: <<https://www.youtube.com/watch?v=7lI-gOShtFA>>. [↑](#footnote-ref-4)
5. Neben dem Web-Service *WebLicht as a Service* ist auch die Web-Anwendung *WebLicht* entwickelt worden. Sowohl mit der Web-Anwendung als auch mit dem Web-Service können die Dateien, aus denen ein digitales Text-Korpus besteht, auf die beschriebene Weise verändert werden. Allerdings unterscheidet sich die Web-Anwendung von dem Web-Service dadurch, wie sie bedient werden kann: Die Web-Anwendung verfügt über eine graphische Benutzeroberfläche, während der Web-Service nicht über eine solche verfügt. Eine Beschreibung der Web-Anwendung, die mit wenigen Einschränkungen auch auf den Web-Service zutrifft, ist an folgender Stelle zu finden:

   <<http://de.clarin.eu/index.php/de/sprachressourcen/weblicht>>.

   Eine Beschreibung der Besonderheiten des Web-Services ist an folgender Stelle zu finden: <<https://weblicht.sfs.uni-tuebingen.de/WaaS/>>. [↑](#footnote-ref-5)
6. Eine sogenannte „processing chain“ ist eine Folge von Web-Services, die so nacheinander ausgeführt werden, dass die einzelnen Web-Services jeweils die Datei verändern, die durch die Ausführung des zuvor ausgeführten Web-Services hervorgebracht worden ist. Eine „processing chain“ wird durch die Erstellung einer besonderen Datei definiert. Eine solche Datei kann mithilfe der Web-Anwendung *WebLicht* erstellt werden. Dabei wird der Benutzer bei der Definition der „processing chain“ dadurch unterstützt, dass er nur die Web-Services an einer bestimmten Stelle zu der „processing chain“ hinzufügen kann, die an dieser Stelle auch ausgeführt werden können. Eine Anleitung zum Erstellen einer „processing chain“ mithilfe der Web-Anwendung *WebLicht* ist an folgender Stelle zu finden:

   <<http://weblicht-clarin.esc.rzg.mpg.de/weblichtwiki/index.php/Getting_Started#Processing_Chains>>. [↑](#footnote-ref-6)
7. Ein API-Key ist . Informationen über die Verwendung von API-Key im Rahmen der Forschungsinfrastruktur *CLARIN* sind an folgender Stelle zu finden: <>. [↑](#footnote-ref-7)