

## Multimodal Transcription Software Programmes

- Anvil
- ChronoViz
- CLAN
- ELAN
- EXMARaLDA
- Praat
- Transana

### ANVIL

ANVIL describes itself as a 'video annotation tool'. It allows for information to be coded along a series of different level 'tracks' which can also be colour-coded as per the user's preference. ANVIL also provides a timeline of the multimodal text. It can play videos within the software, and where there is audio data can present this information as a waveform or pitch contour. ANVIL can also highlight position within the transcribed text track as the video plays.

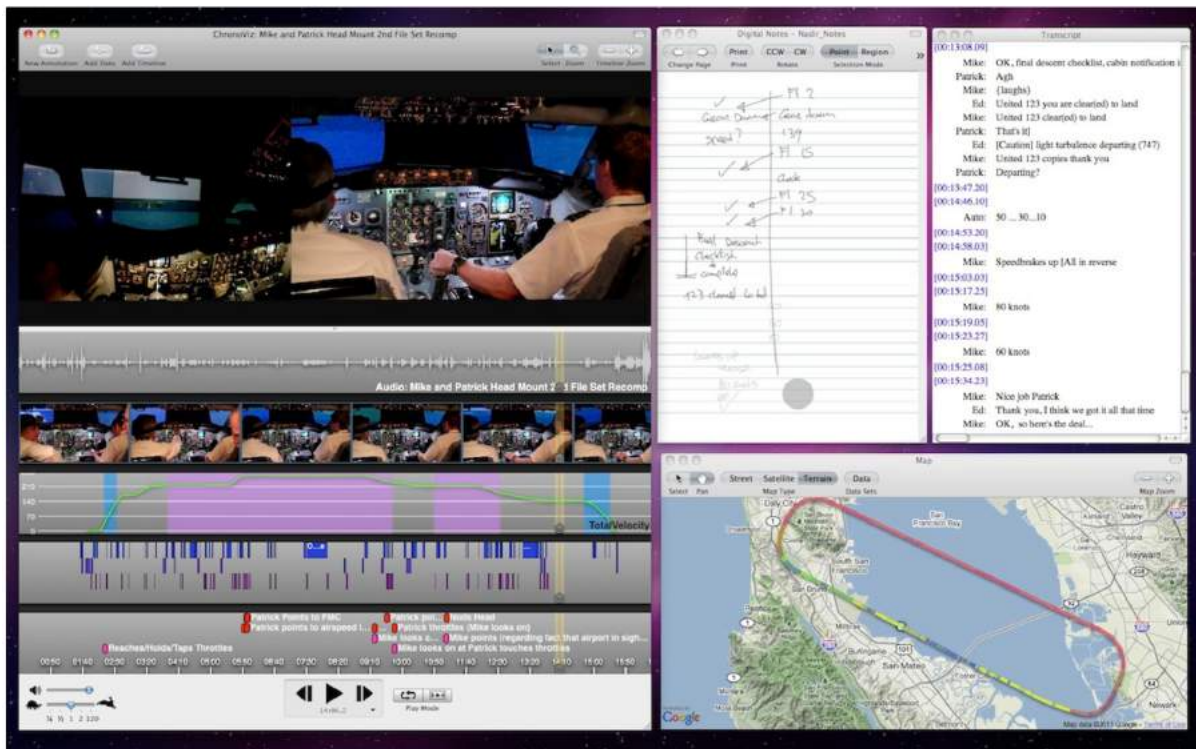


**Figure 1.1.** Initial description: *No initial description.*

Whilst one can transcribe speech manually through ANVIL's text tools, it is also possible to import from more specialised software. For example, ANVIL can import from PRAAT and ELAN if initial software work has been conducted through these programs.

## ChronoViz

ChronoViz is described as a 'tool to aid visualization and analysis of multimodal sets of time-coded information', and as such it excels at presenting multiple different pieces multi-modal data simultaneously, and showing how they progress through time. It also has tools for adjusting the speed of playback for each of the data sets being displayed.



**Figure 2.1.** Initial description: *Integration of digital notes and transcript data with multiple videos.*

Figure 2.1, for example, is a collection of recordings of two men landing an aeroplane. There is video data from their perspective, and from a camera set behind them. There is an audio waveform, a graph showing velocity (presumably of the plane), a map of their flight, a reproduction of digital notes taken and a speech transcript. ChronoViz excels at displaying many different types of information whilst maintaining the accessibility of each data set.

Unfortunately, ChronoViz does not have a cohesive export option. Parts of the information can be exported, such as annotations of data, parts of video clips and screenshots of timelines; but information coded in ChronoViz cannot be automatically exported as a completed transcript.

## CLAN

CLAN was developed in tandem with the CHILDES project in 1984, but continues to be updated (most recently in February of 2017). CLAN is a multi-functional software dealing with aspects both of transcription, and corpus creation, analysis and data sharing. Similarly to ChronoViz, CLAN allows simultaneous portrayal of original media and the transcript being produced.

The image shows the CLAN software interface. On the left is a transcript with phonetic annotations. On the right is a video player window titled "Movie - Sound" showing a video of an elderly woman. The transcript is as follows:

```

5 *K: go'dag:,
7 *FM: g'f'dag?
8 *M: l(g')dag:, *
9 Ps: (3.5) *
10 %com: K taler lavmælt til M)
11 *FM: hvā' skull' (.) det vær'? *
12 %com: pros.contour *
13 Ps: (0.3) *
14 *K: det ska vi li::' (.) sfe. *
15 *M: lnja:::,
16 Ps: (3) *
17 *K: de::'n lāks det der.= *
18 *FM: =ja::?
19 Ps: (.)
20 *K: hvā' koster den der.= *
21 *FM: =æ::: så:d'n=n: (.) pæ:n skive koster en:: omkring: f
22 krøner skiven, *
23 *K: ja::: *
24 Ps: (2.8) *
25 *FM: der ligger ogs'='n (.) pæn hale de:r? = *
26 *K: ja:::
27 Ps: (2.1) *
28 *K: °naj°.=sku' vi ikk' vent' med det syns du. *
29 *M: °jō- la' os ba: venfte°,
30 *K: lskal viđ (.) skryberen.= *
```

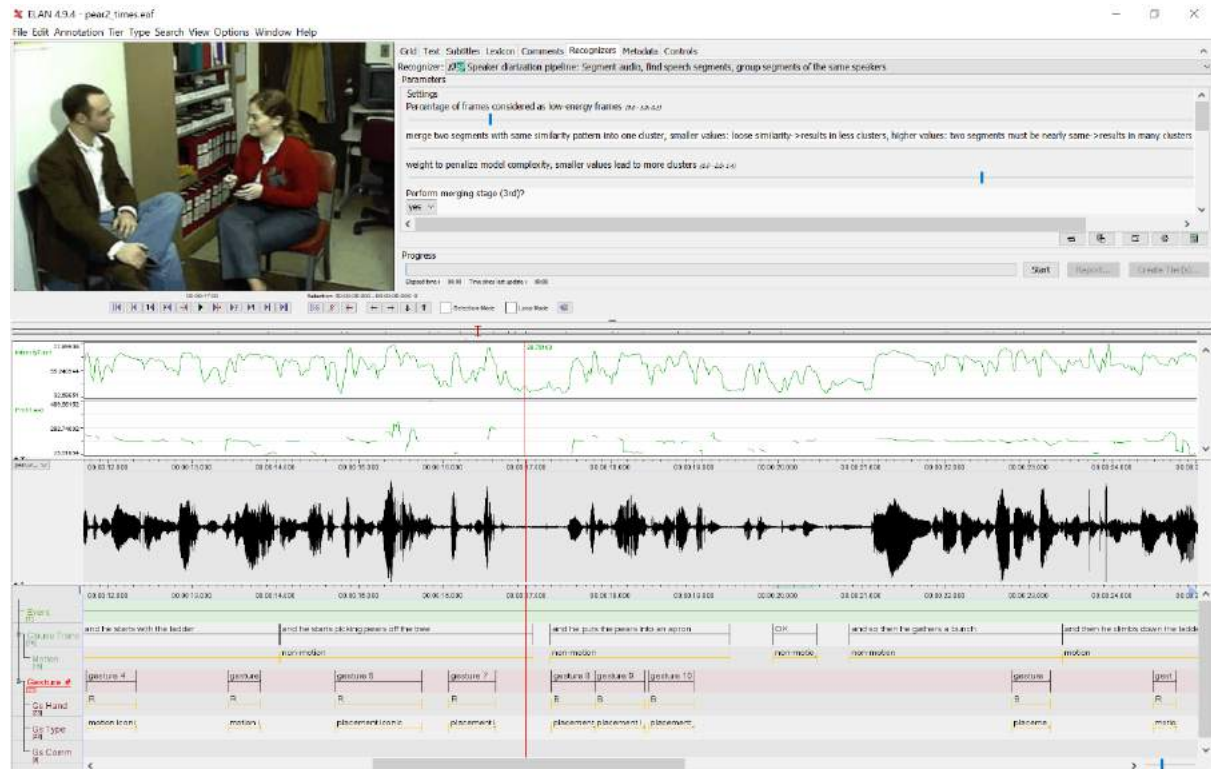
The video player window shows a video of an elderly woman with a white hat and a brown coat. The player has a progress bar at 4533, a "Save" button, volume controls, a zoom of 0.51, a "Repeat" button, and a "fisk.mov" file name.

**Figure 3.1.** Initial description: *Video transcription*

CLAN also links transcribed data to the original media temporally. Thus, one can play the original media and relevant material in the text transcript will be highlighted. Much like previously discussed software, CLAN also has resources to display audio in waveform or pitch contour format. CLAN is designed to work alongside PRAAT which is a more specialised phonetics software. An important function of CLAN is ease of sharing via TalkBank, allowing for the development of corpuses.

## ELAN

ELAN is a software that allows for extensive annotation of multimodal data. ELAN presents this information in a timeline format, and transcription can be separated into various modalities. ELAN has all the standard functionality of a transcription software including multimodal data playback, and audio representation as waveform or pitch contour. Presentation within ELAN is much like timeline multimodal transcription, thus duration and coincidence of actions is easy to recognise.

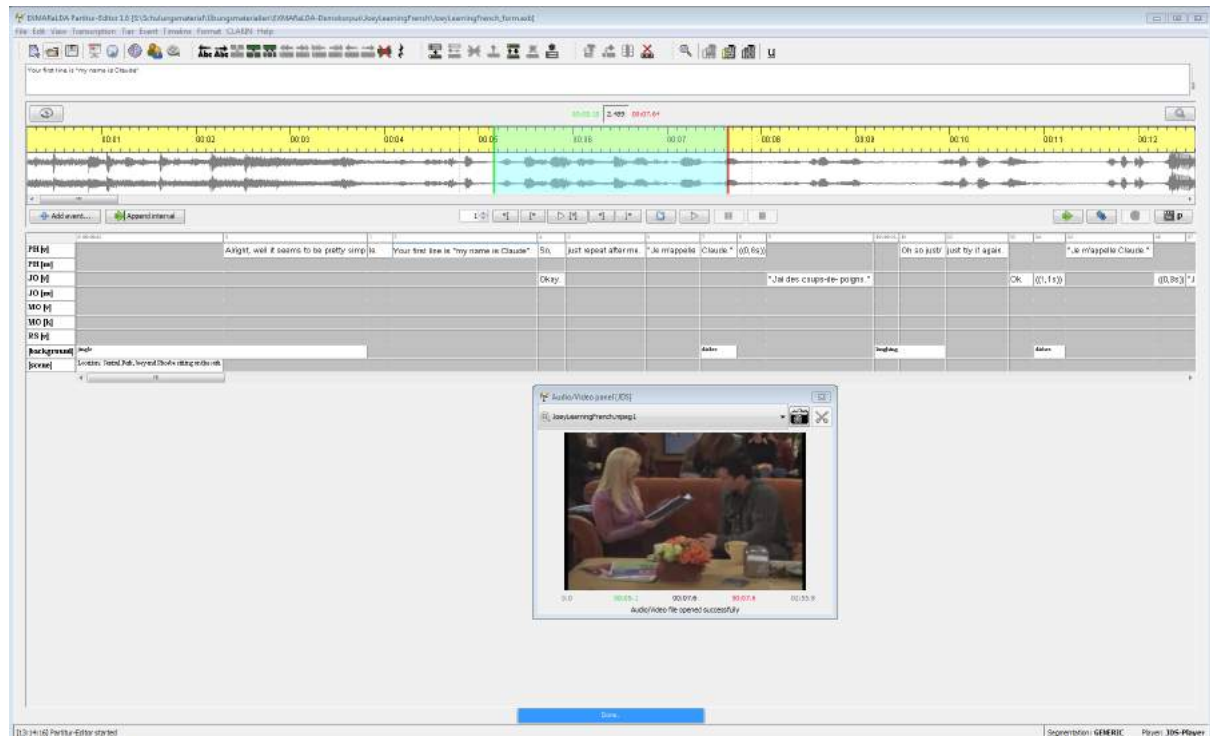


**Figure 4.1.** Initial description: *No initial description.*

ELAN aims to be compatible with other multimodal analysis software, and thus can export to a number of other programs. ELAN has one of the best layouts for developing extensive transcription, as it's timeline style layout is conducive to a finalised transcription.

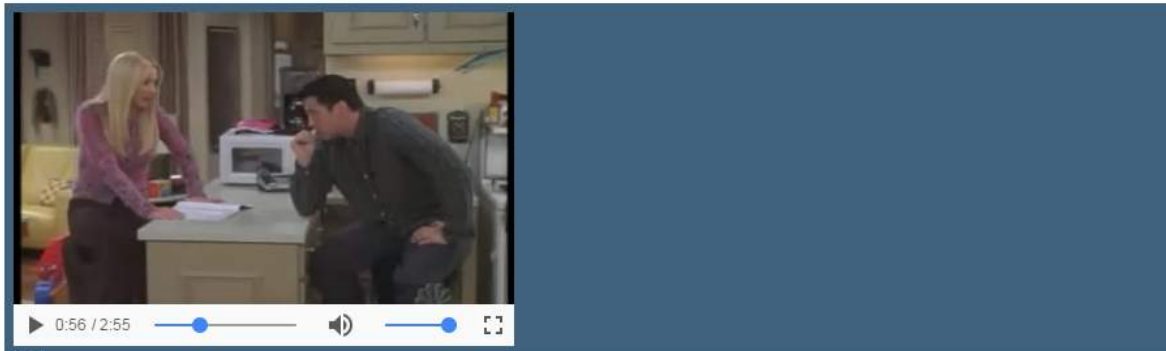
## EXMARaLDA Partitur Editor

EXMARaLDA is 'a tool for transcribing and annotating of digital and audio files'. EXMARaLDA's layout is most similar to ELAN, although the timelines for each modality are contained within a table. Thus, in terms of multimodal transcription style, it is a hybrid of tabular and timeline.



**Figure 5.1.** Initial description: *Interface with Audio/Video Panel*

Its output is also flexible. It can output in various layouts like line for line and musical score, and formats such as HTML or MS Word. Moreover, EXMARaLDA aims to maintain ease of data exchange with other major programs such as PRAAT and ELAN. EXMARaLDA works with a number of transcription conventions, some for German only, but also the widely accepted CHAT and IPA. Below is an example of one output format.

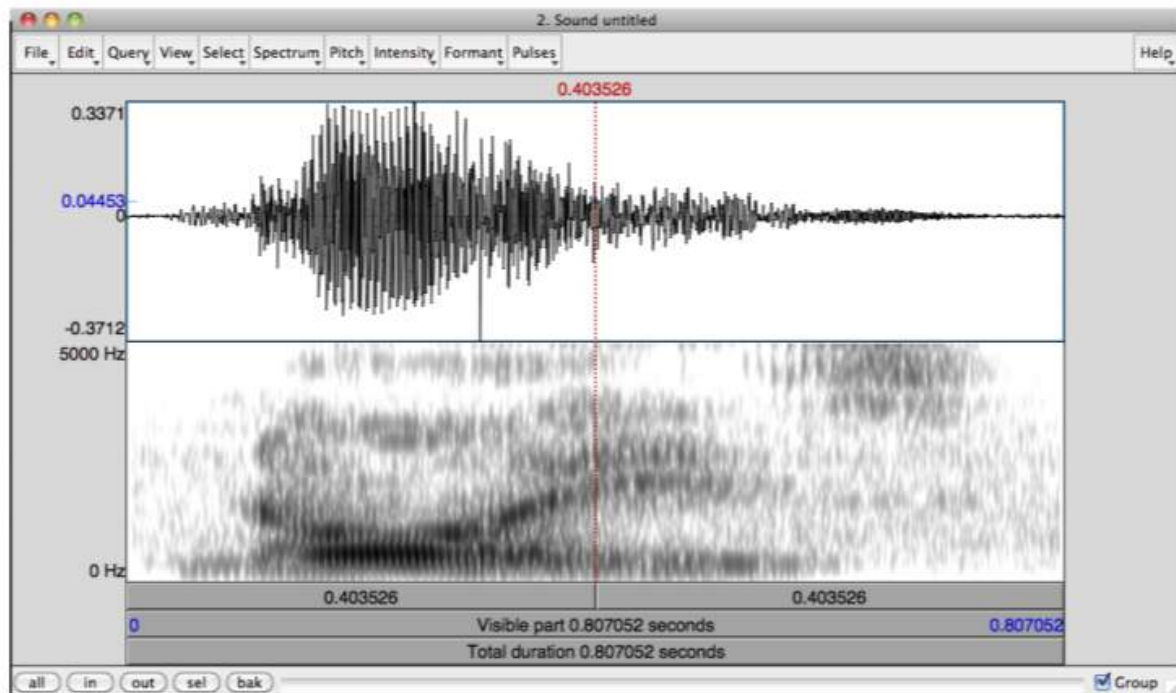


[2]	..* 4* 5* 6* 7* 8* 9* 10* 11*	PH [v] Claude". So, just repeat after me. "Je m'appelle Claude." ((0,6s))
		JO [v] Okay. "J'ai des coups-de- poigns."
		[background] dishes laughing
[3]	11* 12* 13* 14* 15* 16* 17* 18* 19* 20* 21* 22*	PH [v] Oh so just/ just try it again. "Je m'appelle Claude." Mhh. ((0,4s))
		JO [v] Ok. ((1,1s)) ((0,8s)) "Je t'appli m bl/ u."
		[background] dishes dishes laughing
[4]	22* 23* 24* 25 26* 27* 28* 29* 30* 31* 32*	PH [v] It's not ((0,9s)) quite what I'm saying. It do/oes
		JO [v] Really. ((0,7s)) Sounds exactly the same to me.
		[background] dishes laughing laughing
[5]	..* 33* 34* 35* 36* 37* 38* 39* 40* 41*	PH [v] really? Ah well let's just try it again. Really listen. Okay. ((0,5s)) "Je m'appelle
		JO [v] Naha. ((1,3s)) Okay. Got it.
[6]	..* 42* 43* 44* 45* 46* 47* 48* 49* 50* *	PH [v] Claude." Oh mon dieu.
		JO [v] ((0,4s)) "Je te pflu p/ fli." Oh des fouffes.
		MO [v] Hey you guys.
		MO [k] Enters the room.
		[background] laughing jingle
		[scene] Next day, Joey's appartement
[7]	51* 52* 53* 54* 55* 56* 57* 58*	PH [v] "Je m'appe lle Claude." No. Okay maybe we just break it down. ((0,8s)) Okay let'
		PH [nv] screaming
		JO [v] "J'ai des coins mblue."
		[background] laughing

Figure 5.2. Initial description: Output "HTML Partitur + HTML 5 Audio/Video Browser"

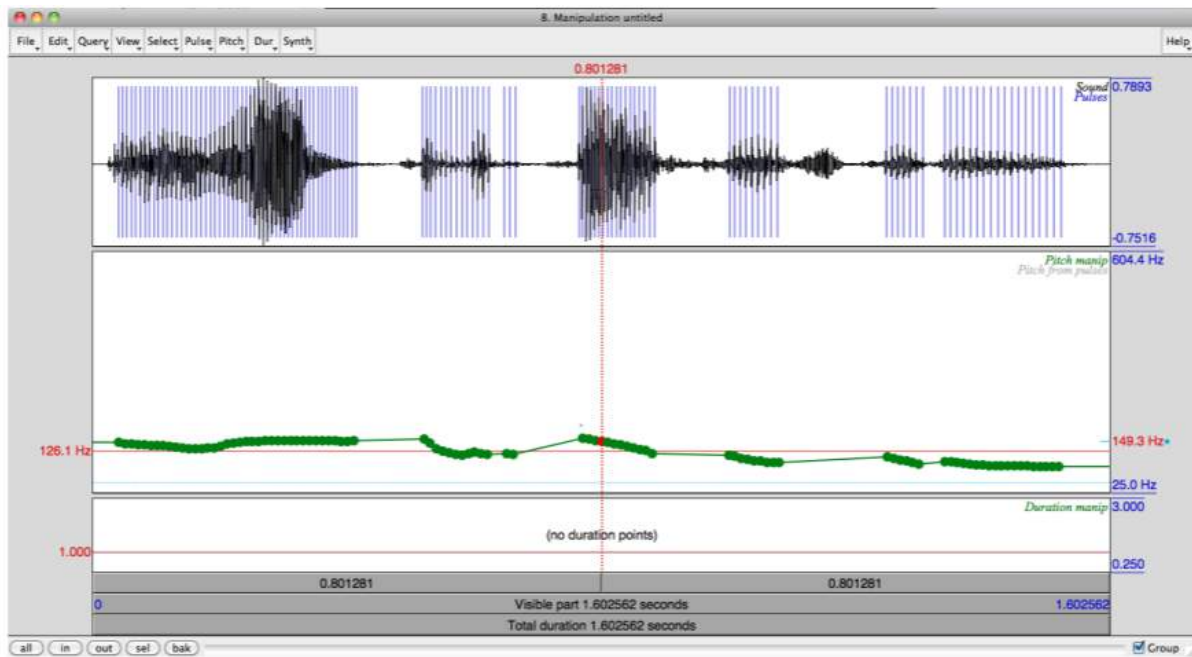
## Praat

Praat is a software tool with many resources to aid in speech analysis. Audio files can either be created by recording through Praat, or imported into the software. Praat can present this audio in a number of formats, including waveform but also more specialised forms such as energy spectrograms.



**Figure 6.1.** Initial description: *The Praat Editor Window*

Other windows within the Praat software can help analyse this data, such as resources for pitch tracking, formants, nasality, amplitude and more. Praat aims to be a comprehensive tool for audio analysis, and as such has more extensive tools than most multimodal transcription software. Praat can also help modify recorded audio.



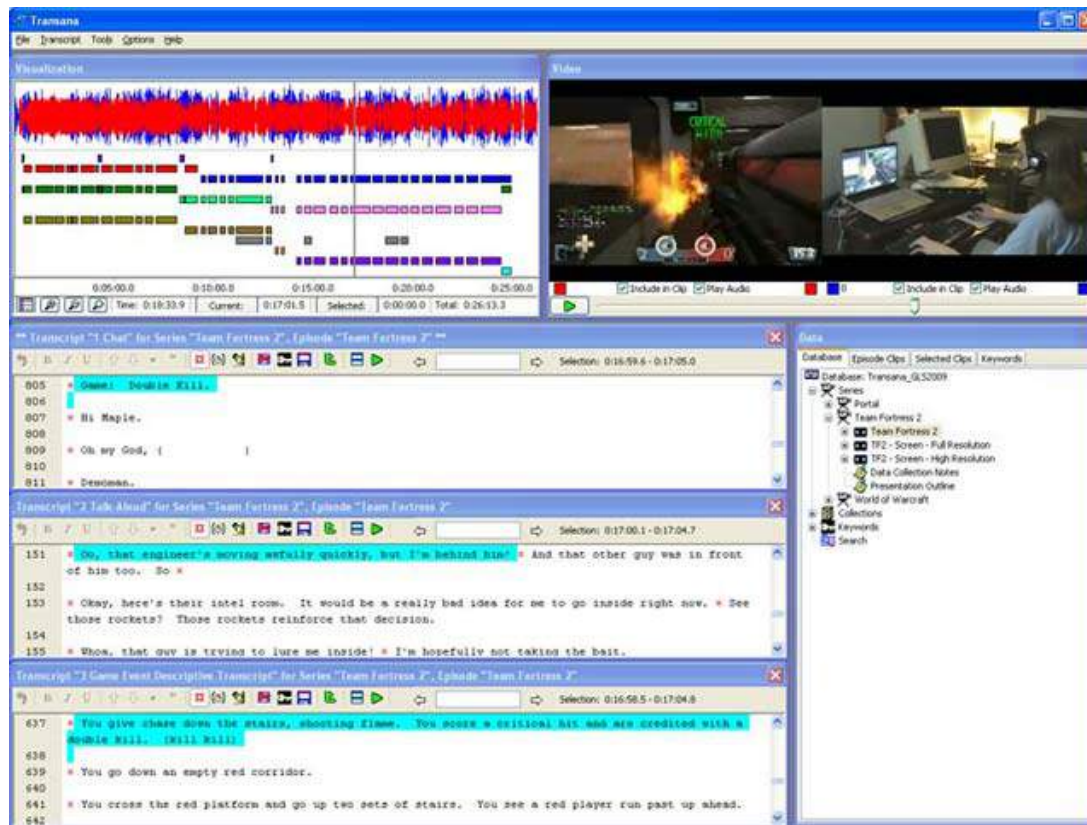
**Figure 6.2.** Initial description: *The Praat Manipulation Window*

This text does not aim to give a full overview of the potential uses of the Praat software. Data analysed within Praat can be exported into many other transcription programs which have limited audio analysis tools of their own. Praat, therefore, may be useful in conjunction with other programs which allow for video analysis.



## Transana

Transana is a multimodal analysis software capable of dealing with text, still-frame images, video and audio within a single analytical frame. Much of Transana's focus lies with coding and categorising data so as to make it easier to manage. However, like other software, it can display data in the most common formats: video player, waveform audio, transcribed speech and so on.



**Figure 7.1.** Initial description: *No initial description.*

A progressive element of Transana is the capacity to share a project between multiple researchers. Different users can work on the same data simultaneously from different locations, with changes happening in real time. Combined with a conference call, this could be effective for collaborative work. Transana also has search tools that can isolate data within a larger field, though this is not a unique feature.

## Conclusion

Whilst it has been useful to gauge the range of transcription and multimodal analysis software available, on a functional level many of these programs are very similar. It has been challenging to find finalised documents featuring products of these programs, thus it is hard to determine which outputs in the most appropriate layout. On the other hand, it is enlightening that most of these programs are open source, and aim to promote compatibility between them. The most prominent programs discussed here seem to be CLAN, ELAN and Praat. Most of the other software aims to be export-friendly to these programs, even if the same is not true in reverse. Praat is slightly different in that it is a specialised audio data analysis tool rather than a comprehensive multimodal software. Considering the high degree of compatibility between software, it seems that a researcher would not be amiss to use the software they are most comfortable or proficient with, so long as they keep in mind they may need to export to a different format when sharing their work.