Berlin Database of Emotional Speech

General information

As a part of the DFG funded research project SE462/3-1 in 1997 and 1999 we recorded a database of emotional utterances spoken by actors. The recordings took place in the anechoic chamber of the Technical University Berlin, department of Technical Acoustics. Director of the project was Prof. Dr. W. Sendlmeier, Technical University of Berlin, Institute of Speech and Communication, department of communication science. Members of the project were mainly Felix Burkhardt, Miriam Kienast, Astrid Paeschke and Benjamin Weiss.

More information about the Berlin Emotional Speech Database and the analysis results you will find in the several publications mentioned here.

For details about the usage of this database read section Emo-DB.

Elements of the navigation frame

- Home
- Emo-DB
- Download
- Documentation
- Results of Analyses and Perception Test
- Contact

Home

With a click on this link you come back to the very beginning. At this site you can choose between normal (1024x768 pixel) and high resolution (1280x1024 pixel) in order to get the best viewing results for your screen size.

Emo-DB

This section explains how to use this website to see what is in the database and how to configure the layout of the presentation of the various kinds of information included in this speech database.
In short there are the sound files itself, the label files (syllable label files and phone label files), information about the results of different perception tests (including the recognition of emotions, the evaluation of naturalness, the syllable stress and the strength of the displayed emotions) as well as some results of the measurements of fundamental frequency, energy, loudness, duration, stress and rhythm.

How to start:

After a click on Emo-DB (in the navigation frame on top) you will see a new frame here. There you have the possibility to choose speaker, text and emotion. The database will show you all the utterances available fulfilling your request. Since the database contains more than 500 utterances be careful not to choose all speakers, all texts and all emotions if you do not want to wait too long for a response.
The result of your request will be shown after clicking on the button "...show!".

You will see a table containing the following information:

- 1. column: a number (continuous)
- 2.-6. column: the results from the perception tests:
    - percentage of correct recognition of emotion
    - percentage of people who thought that the displayed emotion is performed reasonably convincing
    - the emotion recognition of every single participant in the perception experiment (especially interesting if you want to know with which other emotions the utterance has been confused)
    - emotional strength (1= very weak emotion, 7= very strong emotion)
    - standard deviation of the afore mentioned value
- 7. column: a pho-File containing the original F0- and duration values for synthesizing this specific file with MBROLA
- 8. column: the letters P - L - A - Y: you can click on each of them to hear ...
    - (P) the original sound
    - (L) the MBROLA resynthesized version (with the original F_0 contour)
    - (A) a version with stylized F_0 contour resynthesized with MBROLA according to the stylization algorithm by D'Alessandro & Mertens
    - (Y) another resynthesized audio file with a stylized F_0 contour according to an algorithm written by Sascha Fagel
- 9. column: a button to click to see the graphic of the utterance (according to the display configuration shown on the right side)

In the right frame you can configure the display of the chosen utterance according to your needs.

Note: After changing an option to see the effect you will have to reload the configuration AND click again on the button with the name of the utterance.

Options for configuring the graphic display:

- Scaling:
    - real = time scale is the same for all utterances
    - perc = timeline is scaled so that the whole utterance fits in the window size
- Label:
    - none = no label files are shown

- |||| = only borders between syllables are shown
- |a|b|c| = borders and syllable labels are shown
- a,b,c = borders are only shown between syllable labels, not over the complete window height (try and look - you will see)

- Stress:
  - you can only choose if you want to see it or not
  - the yellow bars indicate the stress level of each syllable (average value from the perception test);
    units are from 0 to 3 (according to unstressed, normal stressed, strongly stressed and emphatically stressed)
  - thin dark-red lines constitute the reticule with lines at the values 1 (lowest one), 2 (middle one) and 3 (highest one)

- F_0:
  - normal = F_0 values are shown as measured (no values at voiceless parts)
  - interpolated = F_0 values are shown with intermediate values (calculated by linear interpolation)
- Stylization:
  - none = no F_0 stylization will be displayed
  - spline = F_0 stylization with spline functions will be displayed (programmed by Sascha Fagel)
  - linear = linear F_0 stylization will be displayed (stylization method developed by D'Alessandro & Mertens)
- gliss. threshold:
  - differential glissando threshold is a parameter used for the stylization algorithm by D'Alessandro & Mertens), a higher value results in less chunks, a lower value produces a more detailed stylization and therefore more chunks
- diff. gliss. threshold:
  - differential glissando threshold (another parameter used for the stylization algorithm by D'Alessandro & Mertens), lower value = more chunks

- Trend:
  - linear global trend (=regression line) and the slope of it will be displayed

- Histograms:
  - F_0 histogram of the utterance will be displayed
    if you are especially interested in the histograms you can get a better view (with units marked) under "results - histograms"

- Energy:
  - a dark blue curve representing the measured energy values is shown

- Loudness:
  - a blue curve representing the calculated loudness values is shown, the algorithm used for loudness calculation has been developed by Zwicker (see Zwicker, Fastl (1990): "Psychoacoustics")

- Rhythm Events:
  - if you choose one of the letters A-H red dots (mostly on maxima of the loudness curve) representing the calculated rhythm events are shown
  - the algorithm used for calculation of the rhythm events is also based on a method developed by Zwicker (see the aformentioned book, chapter "Rhythm", page 245f.)
  - letters A to H represent different reference values for the loudness maximum which is needed for calculation
    A - reference value equals the loudness maximum of every single utterance
    B - reference value equals the loudness maximum of a specific text and a specific emotion (all speakers)
    C - reference value equals the loudness maximum of the respective text (all speaker, all emotions)
    D - reference value equals the loudness maximum of the respective emotion (all texts, all speakers)
    E - reference value equals the loudness maximum of all utterances (max. of whole database)
    F - reference value equals the loudness maximum of a specific text and a specific speaker (all emotions)
    G - reference value equals the loudness maximum of the respective speaker (all texts, all emotions)
    H - reference value equals the loudness maximum of a specific emotion and a specific speaker (all texts)

Download

There you will find the possibility (not now but soon!) to download the audio and label files of this database. You can use it for your own analyses as long as you point out the origin of the data correctly. However you will not be able to download and install this graphical web interface for local installation.

Documentation

Its the page you are reading now. If you want to know more let me know by email.

Results of Analyses and Perception Test

There you will find a few of our analysis results regarding measurements of fundamental frequency, duration and stress. You can see histograms of the fundamental frequency of one or more utterances at the same time in different scales (linear scale in Hz and a logarithmic scale in semi tones). Furthermore you can see the results of one of the perception tests which included the recognition of emotions and the naturalness of the utterances.

I hope this is self-explaining - if not, ask me for explanation.

Contact

On this site you will find a contact address. If you would like to know more, if you have questions or want to make comments do not hesitate to ring or mail us.

Additional Information

Every utterance is named according to the same scheme:

- Positions 1-2: number of speaker
- Positions 3-5: code for text
- Position 6: emotion (sorry, letter stands for german emotion word)
- Position 7: if there are more than two versions these are numbered a, b, c ....

Example: 03a01Fa.wav is the audio file from Speaker 03 speaking text a01 with the emotion "Freude" (Happiness).

Information about the speakers

- 03 - male, 31 years old
- 08 - female, 34 years
- 09 - female, 21 years
- 10 - male, 32 years
- 11 - male, 26 years
- 12 - male, 30 years
- 13 - female, 32 years
- 14 - female, 35 years
- 15 - male, 25 years
- 16 - female, 31 years

Code of texts

code	text (german)	try of an english translation
a01	Der Lappen liegt auf dem Eisschrank.	The tablecloth is lying on the frigde.
a02	Das will sie am Mittwoch abgeben.	She will hand it in on Wednesday.
a04	Heute abend könnte ich es ihm sagen.	Tonight I could tell him.
a05	Das schwarze Stück Papier befindet sich da oben neben dem Holzstück.	The black sheet of paper is located up there besides the piece of timber.
a07	In sieben Stunden wird es soweit sein.	In seven hours it will be.
b01	Was sind denn das für Tüten, die da unter dem Tisch stehen?	What about the bags standing there under the table?
b02	Sie haben es gerade hochgetragen und jetzt gehen sie wieder runter.	They just carried it upstairs and now they are going down again.
b03	An den Wochenenden bin ich jetzt immer nach Hause gefahren und habe Agnes besucht.	Currently at the weekends I always went home and saw Agnes.
b09	Ich will das eben wegbringen und dann mit Karl was trinken gehen.	I will just discard this and then go for a drink with Karl.
b10	Die wird auf dem Platz sein, wo wir sie immer hinlegen.	It will be in the place where we always store it.

Code of emotions:

letter	emotion (english)	letter	emotion (german)
A	anger	W	Ärger (Wut)
B	boredom	L	Langeweile
D	disgust	E	Ekel
F	anxiety/fear	A	Angst
H	happiness	F	Freude
S	sadness	T	Trauer
N = neutral version			