

Annotation Pro



Anotacja i eksploracja anotacji nagrań mowy
Prezentacja i demonstracja działania programu

Katarzyna Klessa

Uniwersytet im. Adama Mickiewicza w Poznaniu

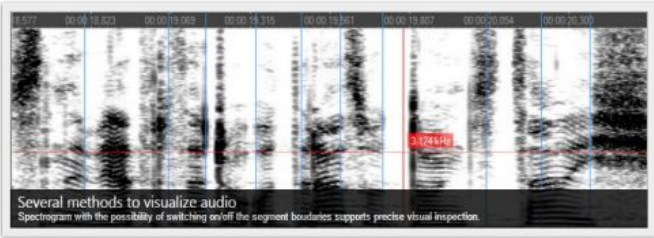
Instytut Lingwistyki Stosowanej

Zakład Komunikacji Multimodalnej

ANNOTATION PRO Let's Annotate Together


Home What's New Features **Downloads** Plugins Help Publications & Cooperation Contact

My Solution in Annotation




Several methods to visualize audio
Spectrogram with the possibility of switching on/off the segment boundaries supports precise visual inspection.

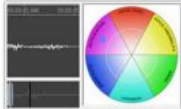
Annotation Pro main features



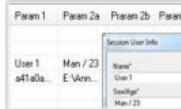
Annotation



Workspace



Graphical Feature Space



Perceptual Tests

ANNOTATION PRO Let's Annotate Together

Home What's New Features Downloads Plugins Help Publications & Cooperation Contact

Downloads

Windows

Current Version: 2.5.2.0
(in case if you have an older version than 2.3.2.4 you need to manually uninstall it and install the current one)

Install With Microsoft ClickOnce Deployment

Installation with OneClick Deployment. Automatically checks for updates at every launch of the programme.

Download Classic Setup

Standard installation. Does not automatically check for updates & new versions. You need to check for them yourself by visiting this website. If update needed in classic setup, uninstall old version first, then install new version.

System requirements:

- Windows 7 and above
- Windows Installer 4.5;
- Microsoft .NET Framework 4.0

Instalacja i pobieranie programu: annotationpro.org

Motywacja i inspiracje



[ELAN](#), [Praat](#), [Wavesurfer](#)

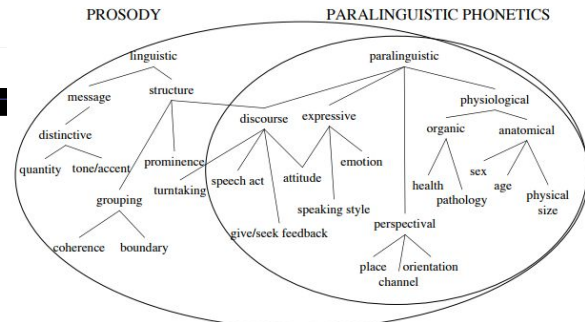


Figure 1. Some functions of prosody and paralinguistic phonetics.

[Schötz, S. \(2002\). Linguistic & paralinguistic phonetic variation in speaker recognition & text-to-speech synthesis.](#)

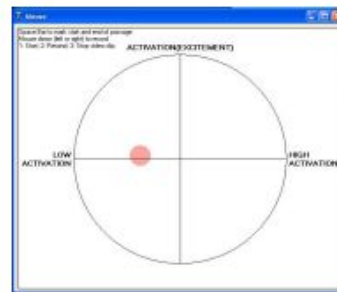


Fig. 1. Screenshot of the modified Feeltrace interface.
Metallinou, A., & Narayanan, S. (2013); zob. też Cowie & Cornelius, 2003

Materiału projektu:
borderland.amu.edu.pl



Cechy językowe i parajęzykowe

Cechy:

- **językowe**: stanowią kod, są specyficzne dla języka (leksyka, składnia, morfologia, inwentarze fonemów, intonemów...)
- **parajęzykowe**: w pewnym stopniu także mogą być specyficzne dla języka, ale mniej ściśle, są też mniej skodyfikowane (“wypełniacze”, “jęki namysłu”, ekspresja emocji, gesty, środki regulowania konwersacji, powtarzanie jednostek/struktur użytych przez rozmówcę, gesty, mimika, język ciała....)
- **pozajęzykowe** (np. fizjologiczne, zdrowie, płeć).

Abercrombie'68: the initial justification for using the term 'paralanguage' was to distinguish purely linguistic features from a highly diversified and heterogeneous group of other types of features that play an important role in human communication

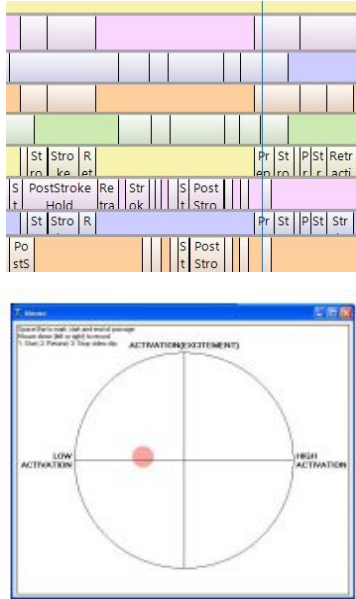
Wielowarstwowość opisu, multimodalność

- Komunikację międzyludzką możemy zdefiniować jako proces złożony z wielu następujących jednocześnie i/lub sekwencyjnie zjawisk, zob. np. <http://www2.ims.uni-stuttgart.de/EGG/>



- Wielowarstwowe, uporządkowane w czasie opisy (anotacje) nagrań zdarzeń komunikacyjnych są naturalną konsekwencją prób odzwierciedlenia tej wielowymiarowości.
- Dla zrozumienia istoty badanych zjawisk istotny może być wpływ cech związanych z szerszym kontekstem aktu komunikacji, sytuacją nagrania -> metadane.

Zróżnicowanie cech – skale oceny



- Przykład: **opis emocji w głosie**. Etykiety, kategorie bardziej przydatne do opisu tzw. emocji podstawowych typu: złość, radość, wstręt, strach, oczekiwanie (np. Ekman; Plutchik). Drobniejsza kategoryzacja / skale ciągłe przydatne w opisie emocji złożonych, o mniej wyrazistym charakterze.
- Ciągłe skale oceny - bywają przydatne gdy trudno jednoznacznie określić kategorie cech; możliwość wykorzystania wartości wynikowych w statystykach opartych o zmienne ciągłe.

Fig. 1. Screenshot of the modified Feeltrace interface.
Metallinou, A., & Narayanan, S. (2013); zob. też Cowie & Cornelius, 2003

Motywacja

i inspiracje

☐ Specyfikacja

- Anotacja cech językowych i para/pozajęzykowych w mowie
- Możliwość zastosowania ciągłych, dyskretnych oraz mieszanych skal oceny percepcyjnej
- Możliwość przeprowadzenia prostych testów percepcyjnych i włączenia ich wyników w warstwy anotacji
- Eksploracja anotacji nagrań mowy, anotacji gestu i mimiki
- Możliwość wykorzystania wyników uzyskanych z użyciem innych narzędzi (np. Inne programy do anotacji, programy automatycznie przetwarzających dane)
- Anotacja tekstu (bez multimediów, por. [Typecraft](#))

Dostępność, Ergonomia

Annotation Pro: obszary wykorzystania

Anotacja

The screenshot displays the software's interface for audio annotation. At the top, there are playback controls and a menu bar. Below, two audio waveforms are visible. The main area is an annotation table for the sentence "pretim guoja celiniks". The table has columns for "Initial", "Phoneme", "Syllable", "Word", and "Comment". The "Phoneme" column shows individual phonetic segments like "p r ' e t i m g u ' o j a t ' s e l i n i k s". The "Word" column shows the words "pretim", "guoja", and "tsel'niks". The "Initial" column shows the first letters "p", "g", "t".

Eksploracja i zarządzanie

Form for exploring and managing annotations:

- P1 Name:
- P2 Name:
- P3 Name:
- Height:
- Font Size:
- Id:
- Coord. Style:



Testy percepcyjne

Options dialog box with tabs: General, Sessions, Appearance.

General tab settings:

- User Must Authorize
- Hide File Name
- Allow Edit Existing Layers
- Allow One Edit

Listen Count (-1 = unlimited):

The diagram is a circle divided into four quadrants by a vertical axis (Active at top, Passive at bottom) and a horizontal axis (Negative on left, Positive on right). The quadrants are labeled with emotions: ANGER (top-left), EXCITEMENT (top-right), FEAR (bottom-left), and JOY (bottom-right). The bottom-left quadrant is also labeled SADNESS, and the bottom-right is labeled SERENITY. The bottom-most part of the circle is labeled BOREDOM. A "Password" field is on the left and a "Cancel" button is on the right.

Anotacja (1): Paralingua

The screenshot displays the Praat software interface. At the top, there's a menu bar (File, Edit, View, Analysis, Plots, Tools, Plugins, Help) and a toolbar. The main window is divided into several panels:

- Workspace:** A table with columns 'N.', 'Open', 'Edit', 'Listen', and 'Acc.'. It lists several audio files with their durations and status.
- Audio:** A spectrogram and a waveform view of the audio file.
- Annotation:** A timeline showing segments. A segment labeled '-18,-7' is highlighted in blue. Below it, a pink bar represents the 'Distance Result'.
- Text:** A transcription of the audio content: 'dzisiaj jest poniedziałek'. Below the transcription, there's a phonetic breakdown: 'd'z'1' s'aj j'est po me d'z'a wek'.
- Feature:** A circular color wheel with segments for different features like 'WYKAZA', 'SMUTEK', 'ROZDROZENIE', etc.
- Properties:** A panel on the right with fields for 'Label', 'Speaker', 'Name', 'Par: 1', 'Par: 2', 'Par: 3', and 'R Script'.

[\(Klessa et al., 2013\)](#)

- Dialogi zadaniowe (rozmowy telefoniczne)
 - z ograniczeniem czasowym
 - bez ograniczeń czasowych
- Nagrania w studio
- Portrety emocjonalne (por. Bänziger & Scherer 2007; aktorzy: opis sytuacji, mającej wywołać określoną emocję - wypowiedź: leksykalnie neutralna - względnie!)

Anotacja (2): Pogranicze borderland.amu.edu.pl

- Kontekst interkulturowy: język polski, niemiecki i angielski;
- Mówcy 12-14 lat;
- Rejestracja fonii i wizji w warunkach “studia terenowego”;
- Dialogi zadaniowe o charakterze
 - “konkurencyjnym” (prezent urodzinowy)
 - “kolaboracyjnym” (budowa wieży);
- Emblematy (zadanie indywidualne, por. prace E. Jarmołowicz-Nowikow).



Fot. Archiwum Projektu Borderland, CMU, Poznań



Borderland

Language of Boundaries and Boundaries of Language

The aims of the project are the documentation and interdisciplinary analysis of phenomena related to interpersonal communication. The main research focus is on paralinguistic phenomena (prosody, gestures) and accommodation within the communication process for fieldwork analyses of an exceptionally close co-existence of cultures, languages, and social groups connected by the Oder. The project's goals include: studies of linguistic and cultural awareness, the sense of identity, the awareness of language and the dynamics of the processes of mutual accommodation (communicative convergence) of paralinguistic features in task-oriented gestures used by Polish and German speakers.



NARODOWY PROGRAM
ROZWOJU HUMANISTYKI

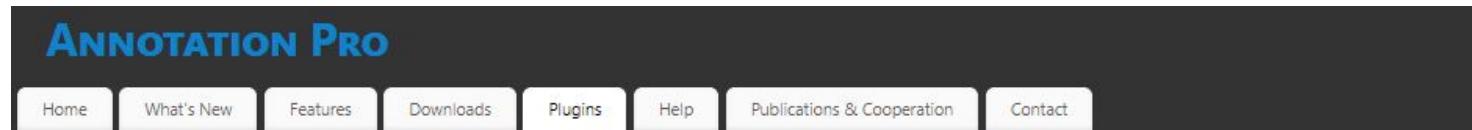
Narodowy Program Rozwoju Humanistyki powstał w oparciu o konsultacje z ekspertami, a w konkursowym trybie wybrano zaproszeń i dofinansowanie spośród wniosków zgłoszonych w ramach Narodowego Programu Rozwoju Humanistyki o zasadniczym znaczeniu dla zachowania polskiej tożsamości

Rozszerzenia: mowa.clarin-pl.eu/tools/

<http://mowa.clarin-pl.eu/tools/annotationpro>



Rozszerzenia: <http://annotationpro.org/plugins/>



Plugins

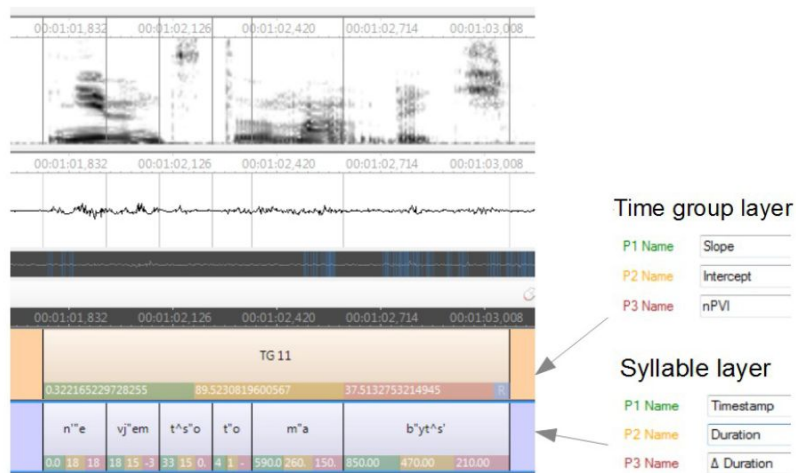
Annotation Pro enables easy and efficient extensions of the built-in functionality by means of plugins. In this section we will publish some of the plugins. Further below you can also find some help on how to create a plugin of your own, and a link to Annotation Pro API.



In order to use any of the Annotation Pro plugins it is sufficient to save the plugin to the 'Plugins' folder in the 'Annotation Pro' folder on your disk (the plugins below are zipped, so after saving them you also need to unzip them). Then the plugin's name becomes visible in the Plugins menu in your Annotation Pro programme (you might need to press 'Refresh' in order to update the display of the menu list). The plugins can be used for a single annotation file or for a number of files selected in the programme's workspace (choose "Workspace Mode" for using a plugin with multiple files selected in your workspace).

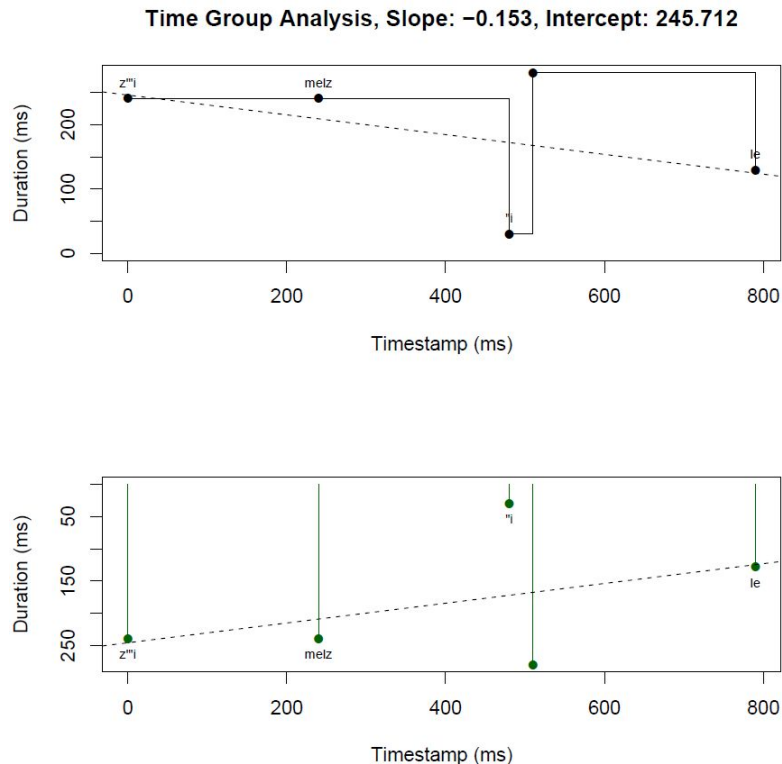
Plugin name	Description	Details	Download
Statistics – Annotation Pro + TGA	Plugin based on the TGA idea and on-line tool solution by Dafydd Gibbon. Update: enables file collections processing (Workspace Mode)	Read more...	Download
Statistics – Segment Rate Moving Average (SRMA)	Calculates segment rate using moving average method	Read more...	Download
Statistics – nPVI moving average (nPVIMA)	Computes nPVI (Normalized Pairwise Variability Index)	Read more...	Download

Rozszerzenia: analiza zmienności czasowej TGA

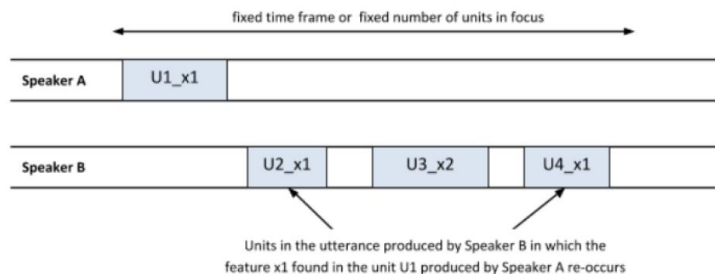


Example TGA results as displayed in the *Annotation Pro* interface for a Polish utterance: *Nie wiem co to ma być*. The obtained values are visible as segment parameters (P1, P2, P3) with coloured background at the bottom of each segment (a time group segment or a syllable segment respectively).

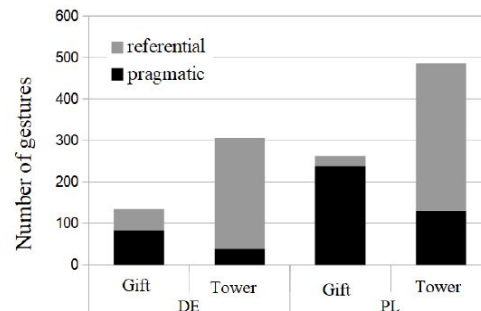
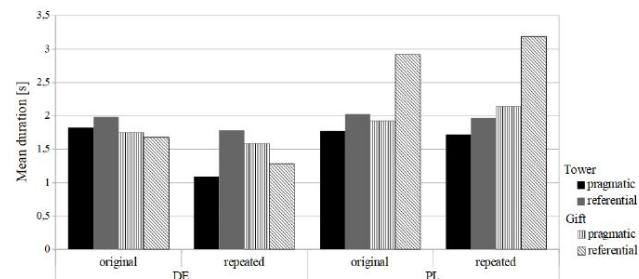
Gibbon & Jue, 2015; Klessa & Gibbon, 2014



Rozszerzenia: analiza akomodacji w dialogu



- Liczba gestów większa u mówców polskich (Karpiński i in., 2018).
- Wpływ typu zadania - obserwowany w obu grupach (więcej referencyjnych w zadaniu *Tower*, a pragmatycznych w *Gift*).
- Gesty powtarzane przez Niemców średnio krótsze niż oryginalne; u Polaków - bez istotnej różnicy.



Tworzenie materiału do badań odsłuchowych

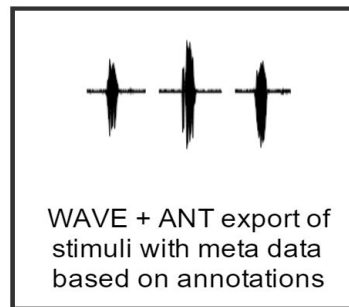
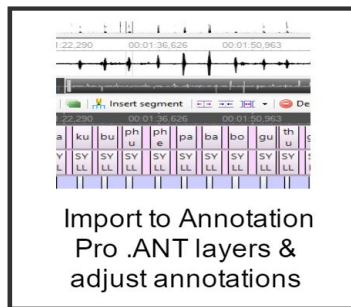
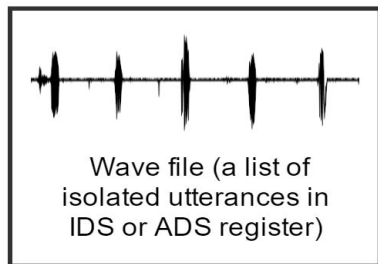
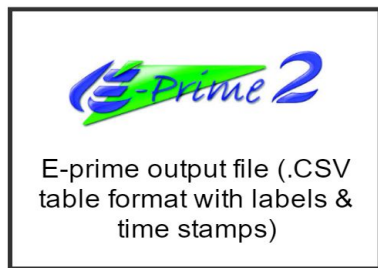
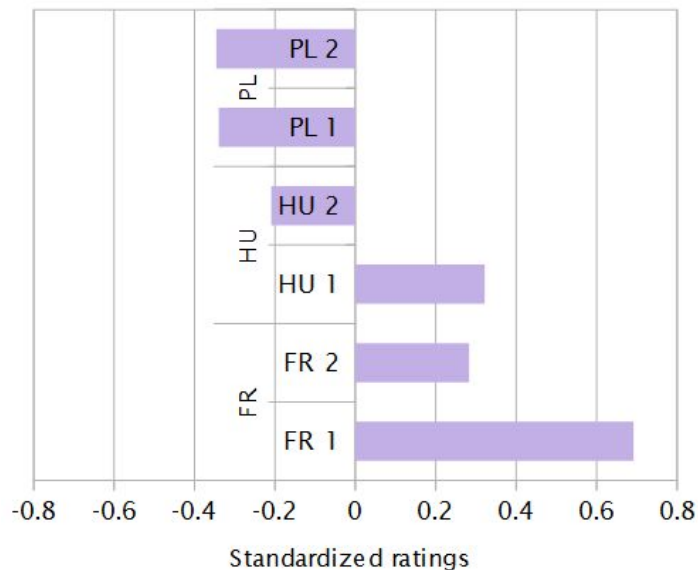


Photo: courtesy of Magdalena Oleškowicz-Popiel



NARODOWE CENTRUM NAUKI

Testy percepcyjne



Klessa, Karpiński & Czoska (2015); Klessa (2016).

- Sylaby zestaw FR: /fe/, /fɛ/, /ne/, /nɛ/; zestaw HU: /fø/, /fu/, /nø/, /nu/, zestaw PL: /fe/, /fi/, /fu/, /ne/, /ni/, /nu/;
- 23 listeners (all aged 20-23, 15 f, 8 m);
- Each sample was repeated 4x (56 stimuli altogether).

Demonstracja działania

- **Pytanie badawcze** (por. A. Ericsson et al.): jakie są wyznaczniki akcentu leksykalnego w języku polskim w zależności od stylu wypowiedzi?
- **Nagrania:** Korpus dialogów zadaniowych i wypowiedzi izolowanych *Axio* (Klessa & Niećko-Bukowska, 2017)

Cel działania: przygotowanie zestawów fraz i wyrazów, wraz z:

- 1) transkrypcją ortograficzną;
- 2) segmentacją i transkrypcją fonetyczną na poziomie głosek.

Metoda i materiał

Sesja 1. i sesja 2.

The Acoustics of Word Stress in English as a Function of Stress Level and Speaking Style

Anders Eriksson, Mattias Heldner

The Acoustics of Lexical Stress in Italian as a Function of Stress Level and Speaking Style

Anders Eriksson¹, Pier Marco Bertinetto², Mattias Heldner¹, Rosalba Nodari², Giovanna Lenoci²





ADAM MICKIEWICZ
UNIVERSITY
POZNAŃ

Axiio Speech Corpus

design, structure, and preliminary analyses

Katarzyna Klessa & Bożena Niecko-Bukowska
(klessa, bniecko)@amu.edu.pl

Abstract

We present a newly created speech corpus in Polish (Axio Corpus) along with the results of preliminary evaluation and analysis of the corpus content.

Design assumptions & Goals

According to the assumptions underlying the corpus design, its first application should enable task-based or content-based investigation of the properties of semi-spontaneous evaluative utterances, especially from the perspective of the notion of responsibility (cf. Isaacs, 1979/1996). Other aspects of phonetic-acoustic correlates of word stress in Polish semi-spontaneous oral utterances.

By aiming at these two different research goals within the design of one common corpus we wished to extend the potential of the resource usability in line with the contemporary priorities for language resource design and development (cf. g. Vraini et al., 2008).

Speakers, recording procedures & speech material

The Axio Speech Corpus includes recordings of 52 Polish native speakers (adults, 20-60). The recordings were collected through a language production.

Stage 1: Three task-oriented dialogues aiming at eliciting semi-spontaneous conversational. The dialogue tasks involved discussing photographs of everyday life situations potentially encouraging reflections concerning responsibility in its various meanings. Additionally, as part of the final task, the speakers cooperatively designed a semi-structured narrative of their discussion.

Stage 2: Speakers recorded task up to four weeks after the first stage. Here, the speakers were requested to read a phrase list designed based on the materials used during the first stage. The phrase lists were divided into three different sets: **read and semi-spontaneous speaking style** (Eriksson, A., 2016), **read** and **semi-spontaneous speaking style** (Eriksson, A., 2016), **read** and **semi-spontaneous speaking style** (Eriksson, A., 2016).

Responsibility questions

Scenario 1: Responsibility as a goal.
Scenario 2: Mutual responsibility + responsibility for people.
Scenario 3: Responsibility as a task.

Task oriented dialogues, conversational speaking style

Scenario 1



Initial setting: at the beginning only Speaker 1 can see the picture.
Task: Speaker 1 describes and comments on the picture. Speaker 2 makes questions and related details. Then both participants look at the picture and express their opinions.

Scenario 2



Initial setting: at the beginning only Speaker 1 can see the picture.
Task: The same as in Scenario 1 (participants exchange roles).

Scenario 3



Initial setting: both participants can see the picture from the very beginning.
Task 1: Discuss details of the picture. Together try to figure out what might have happened. Speaker 2 asks additional task questions.
Task 2: Design a brief press note respecting on the situation.
Task 3: Read your note (both speakers).

Scenario 4



Initial setting: both participants can see the picture from the very beginning.
Task: Participants discuss the picture together for a few seconds and then disappear. Please say the phrase and then press space bar to see the next phrase.

Dialogi jednolite (do 1:10)

Procedure outline

```

    graph LR
      A["Stage 1: Scenarios 1-3"] --> B["Phase data selection  
2-4 weeks interval"]
      B --> C["Stage 2: Scenario 4"]
    
```

Phrase extraction, annotation & first insights

Annotation: Praetorius et al. (2017)
Multilayer annotation scheme

So far, a subset of the data (i.e. the candidate utterances for the phrase list) annotated on the orthographic level & transcribed in IPA/ARF (Grimm) using segmentation with open & PoS/parse embedded in Annotation Pro, and time alignments with Annotation Pro.

Phase lists & target word lists exported to separate files with Annotation Pro.

Session duration in Stage 1: 20-30 minutes per pair.
The number of phrases, list in total, had larger word lists (i.e. 10-20 phrases / speaker; 15-20 target words / speaker).

References

Eriksson, A., Heldner, M. (2016). The acoustics of word stress in English as a function of stress level and speaking style. Proceedings of INTERSPEECH 2016, Dresden, 6-11.

Isaacs, H. (1979) Das Prinzip Verantwortung. Versuch einer Ethik für den technologischen Zeitalter. Frankfurt am Main, (1986) Deutsche Übertragung: Die Ethik der Verantwortung. (Herausgegeben von Hans Jonas, M. Salmela, S. Wigger, J. Wigger, M. A. Kambhampati, K. (2008). E-Learning. Cambridge Language Resources and Technology Infrastructure. Proceedings of the 4th International Conference on Language Resources and Evaluation, Manchester.

Annotation Pro – okno programu

The screenshot shows the main interface of Annotation Pro. On the left, a 'Workspace (.antw)' panel lists files with columns for Name, Lines, and Abs. The main window displays a spectrogram at the top, a waveform below it, and a segmented transcription at the bottom. The transcription is divided into segments with different background colors and contains text in Polish: 'Proszę pana ja zgłaszałam przed chwilą proszę szybko przejechać tutaj na przystanku autobusowy'. Below the text, a phonetic transcription is visible: 'e - a j a z b a s a e - o - o s i e s y k o p y a t u t j a s y s a e a w t o u s o v y'. On the right, a 'Feature' window shows a circular color-coded feature space and a 'Layer, segment properties. Segment list display' panel with fields for Label, Separator, Name, Par 1, Par 2, Par 3, and a list of segments with their start and end times.

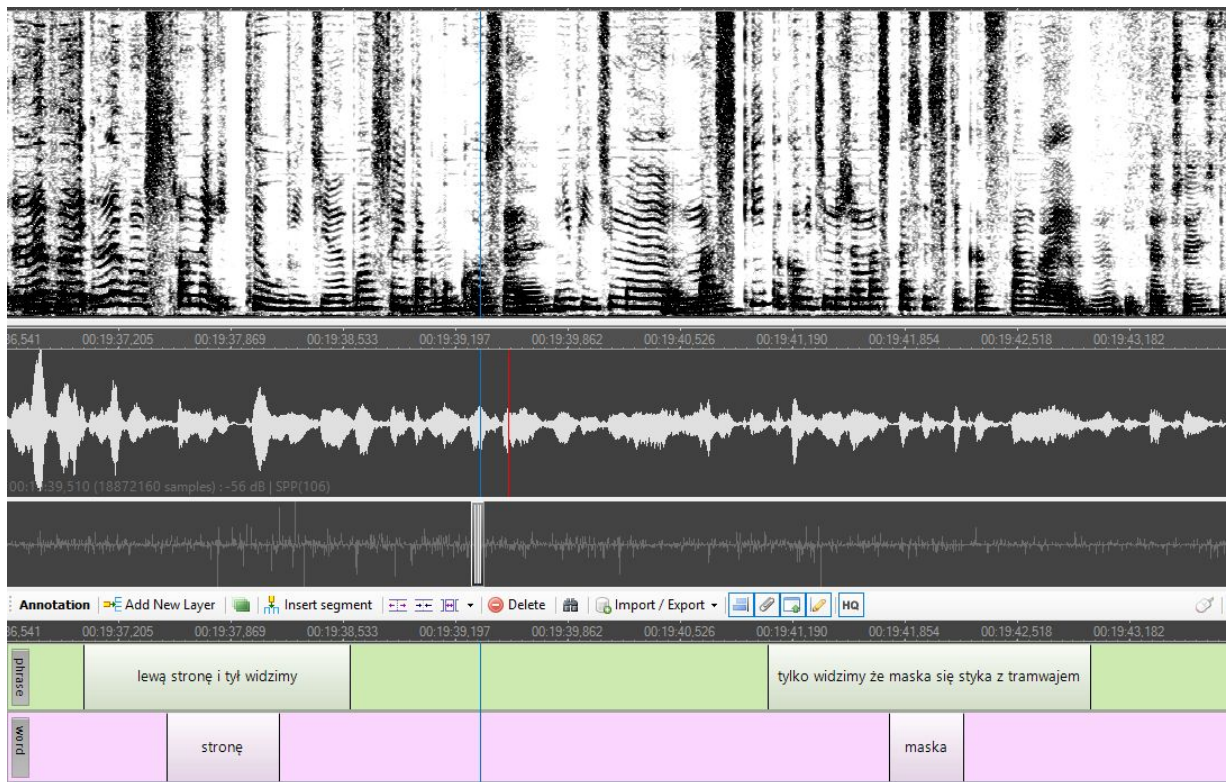
**File collection
Workspace
(.antw)**

**Configurable speech
signal display (.wav)
& annotations (.ant) for
a single recording**

**Graphical
representation
of the feature
space**

**Layer,
segment
properties.
Segment list
display.**

Anotacja wybranych fragmentów w dialogach



Ekspert anotacji i stworzenie listy bodźców

Import / Export

- Import CSV File...
- Import Layers From CSV (Label, Start, Stop)...
- Import Layers From Text Files...
- Import Layers From BLF Files...
- Import Layers From TRS Files...
- Import Layers From TextGrid Files...
- Import Layers From XRA Files...
- Import Layers From EAF Files...
- Import Praat F0 File...
- Import Praat Formant File...
- Export Layer To CSV File...**
- Export Layer To Text File...
- Export Layer To Text & Audio Files...
- Export To CSV...
- Export To TextGrid...
- Export To XRA...
- Export To EAF...
- Export To ANT & Audio Files...

Layer Segment Audio Video Meta Segment List

Label

to jakim cudem on mi zasłania schody
dopiero dalej jak patrzysz to widać dachy
dokładnie widoczne jak te bliżej
bardzo mały blok
na krańcu tej fotografii
już się kilka ustawiło w kolejce
i sięgają aż do końca ulicy
jeszcze się styka ten samochód
taki bardzo ciemny kolor
my widzimy go tak ukosem
i my tego tramwaju nie widzimy w całości
kierowcy w środku już nie ma
lewą stroną i tył widzimy
tylko widzimy że maska się styka z tramwajem
jest tyłem do tego samochodu
właśnie tylko ten bordowy
między czwartym i piątym stoją samochody
nie widzę tu żadnych otarć tam gdzie jest ten samochód
ale nie ma też jakiegoś dużego tłumu ludzi
ale jest to niski krawężnik nie jest wysoki



stimuli - loop
Repeatedly runs another item

Run trial_sequence Break if never

Repeat each cycle 3,00 x Evaluate Resume

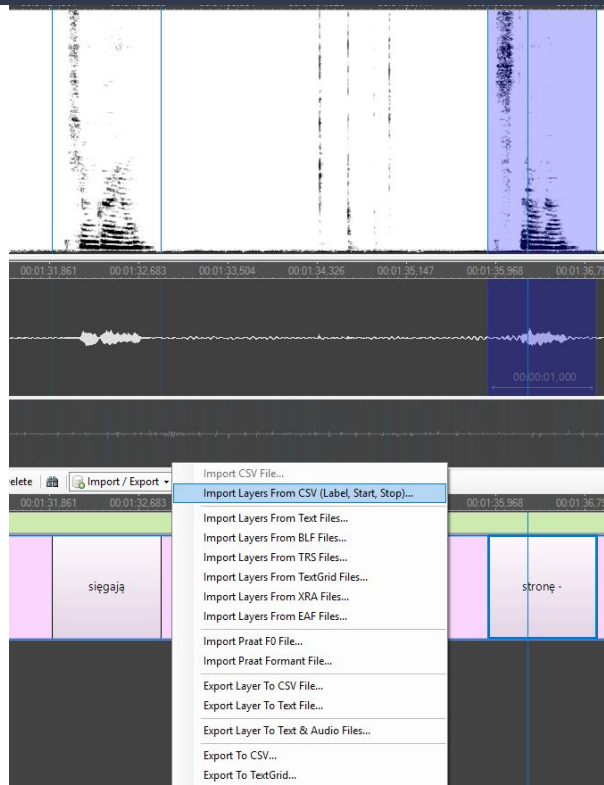
Order random

Source table

Summary: trial_sequence will be called 60 times in random order. The number of times it is called is determined by the number of items in the table.

	stimulus	type
1	to jakim cudem on mi zasłania schody	phrase
2	dopiero dalej jak patrzysz to widać dachy	phrase
3	dokładnie widoczne jak te bliżej	phrase
4	bardzo mały blok	phrase
5	na krańcu tej fotografii	phrase
6	już się kilka ustawiło w kolejce	phrase
7	i sięgają aż do końca ulicy	phrase
8	jeszcze się styka ten samochód	phrase
9	taki bardzo ciemny kolor	phrase
10	my widzimy go tak ukosem	phrase
11	i my tego tramwaju nie widzimy w całości	phrase
12	kierowcy w środku już nie ma	phrase
13	lewą stroną i tył widzimy	phrase
14	tylko widzimy że maska się styka z tramwajem	phrase

Nagranie sesji 2. i import wyniku z Open Sesame



```
44 stopnie →177.667 →178.667
45 leży →181.232 →182.232
46 schodach →184.86 →185.86
47 tramwaju →188.476 →189.476
48 uszkodzony →191.432 →192.432
49 który →202.148 →203.148
50 widać →206.399 →207.399
51 przy schodach →211.47 →212.47
52 części →215.515 →216.515
53 kawalków →219.946 →220.946
54 rozbity →225.324 →226.324
55 przy schodach →230.93 →231.93
56 teraz →236.548 →237.548
57 teraz →239.739 →240.739
58 stanie →244.781 →245.781
59 wyraźne →248.948 →249.948
60 w stanie →256.634 →257.634
61 było →261.306 →262.306
62 długim →264.988 →265.988
63 teraz →268.568 →269.568
64 który →272.355 →273.355
65 błuzka →275.996 →276.996
66 mocno →279.47 →280.47
67 błuzka →282.804 →283.804
68 strony →286.328 →287.328
69 uszkodzony →289.635 →290.635
70 kawalków →292.587 →293.587
71 kawalek →296.259 →297.259
72 błuzka →299.771 →300.771
```

	A	B	C	D
1	stimulus	start	stop	time s
2	części	0,8	1,8	30,
3	tramwaju	5,826	6,826	35,
4	siedzi	10,701	11,701	40,
5	uszkodzony	14,621	15,621	44,
6	mocno	22,284	23,284	52,
7	jedno	26,004	27,004	55,
8	siedzi	30,312	31,312	60,
9	który	33,875	34,875	63,
10	strony	37,203	38,203	67,
11	mocno	42,578	43,578	72,
12	wyraźne	45,941	46,941	75,
13	tego	49,89	50,89	79,
14	siedzi	53,346	54,346	83,
15	części	56,819	57,819	86,
16	stanie	60,796	61,796	90,
17	rozbity	65,028	66,028	94,
18	wyraźne	68,767	69,767	98,
19	długim	72,327	73,327	102,
20	tego	75,989	76,989	105,
21	rozmyślał	79,427	80,427	109,
22	rozmyślał	82,859	83,859	112,
23	było	87,289	88,289	117,
24	mogło	91,001	92,001	120,
25	widać	94,637	95,637	124,
26	wstanie	98,236	99,236	128,
27	tego	101,78	102,78	131,
28	schodach	105,315	106,315	135,
29	leży	108,885	109,885	138,
30	oparty	113,961	114,961	143,
31	schodach	121,155	122,155	151,
32	rozmyślał	124,876	125,876	154,
33	było	128,627	129,627	158,
34	stanie	132,163	133,163	162,

Korekta położenia granic i podział na głoski

The screenshot displays the ANNPRO software interface. At the top, a spectrogram shows frequency components over time, with a red box highlighting a 7.810 kHz component. Below it, a waveform shows the amplitude of the audio signal. A circular diagram on the right represents a sentiment model with axes for Active/Passive and Negative/Positive, containing labels for Anger, Excitement, Joy, Fear, Neutral, Sadness, Boredom, and Serenity. The bottom section shows an annotation table for the phrase "ma ciemne włosy" with columns for word and phone. An "Auto Segmentation" dialog box is open, showing the "CLARIN Align 190306" module selected, and options for "Segment Separation" and "Selected Segments".

1,955 00:00:12,231 00:00:12,507 00:00:12,783 00:00:13,059 00:00:13,335 00:00:13,611

7.810 kHz

1,955 00:00:12,231 00:00:12,507 00:00:12,783 00:00:13,059 00:00:13,335 00:00:13,611

00:00:00,305

Layer Segment Audio Video Meta Segm

Annotation Add New Layer Insert segment

1,955 00:00:12,231 00:00:12,507 00:00:12,783

phrase	ma ciemne włosy				
word	ma	ciemne	włosy		
phone	m	a	ts'	e	m n e v w o

Auto Segmentation

Segmentation Module

CLARIN Align 190306 Refresh Open Folder

Show Console Window

Operation

Segment Separation Fill Gaps

Selection

Layer Selected Segments

Options

Words Syllables Phonemes

Resize Insert #Sp

Separate Stop

Stworzenie zestawu nagrań i anotacji

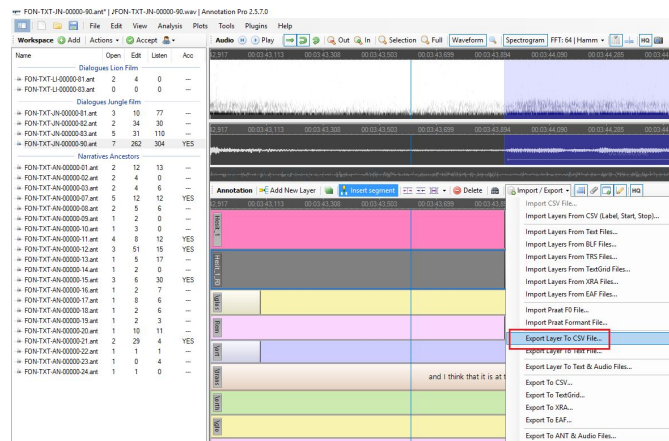
The screenshot displays the Audacity interface. On the left, a project pane lists files: KKJM1_fras.ant, KKJM1_ort.ant, WPMW2_fras.ant, WPMW2_ort.ant, and KKJM1_f_fras.ant. The main workspace shows two audio tracks with waveforms. The top track is labeled 'Wzrostka zwieta' and the bottom track is 'KKJM1_ort'. A menu is open over the 'Wzrostka zwieta' track, showing options like 'Import Layers From CSV (Label, Start, Stop)...', 'Export Layer To CSV...', and 'Export To ANT & Audio Files...'. The annotation track below shows segments with labels: 'ustawilo', 'krawę znak', 'stronę', and 'maly'.

Nazwa	Data modyfikacji	Typ
0001_bardzo.ant	24.05.2019 00:11	Annotation
0001_bardzo.wav	24.05.2019 00:11	WAV
0002_kilka.ant	24.05.2019 00:11	Annotation
0002_kilka.wav	24.05.2019 00:11	WAV
0003_ustawilo.ant	24.05.2019 00:11	Annotation
0003_ustawilo.wav	24.05.2019 00:11	WAV
0004_ciemny.ant	24.05.2019 00:11	Annotation
0004_ciemny.wav	24.05.2019 00:11	WAV
0005_stoja.ant	24.05.2019 00:11	Annotation
0005_stoja.wav	24.05.2019 00:11	WAV
0006_widoczne.ant	24.05.2019 00:11	Annotation
0006_widoczne.wav	24.05.2019 00:11	WAV
0007_niski.ant	24.05.2019 00:11	Annotation
0007_niski.wav	24.05.2019 00:11	WAV
0008_duzego.ant	24.05.2019 00:11	Annotation
0008_duzego.wav	24.05.2019 00:11	WAV
0009_w_środku.ant	24.05.2019 00:11	Annotation
0009_w_środku.wav	24.05.2019 00:11	WAV
0010_krańcu.ant	24.05.2019 00:11	Annotation
0010_krańcu.wav	24.05.2019 00:11	WAV
0011_maska.ant	24.05.2019 00:11	Annotation
0011_maska.wav	24.05.2019 00:11	WAV
0012_krawężnik.ant	24.05.2019 00:11	Annotation
0012_krawężnik.wav	24.05.2019 00:11	WAV
0013_zadnych.ant	24.05.2019 00:11	Annotation
0013_zadnych.wav	24.05.2019 00:11	WAV
0014_tylem.ant	24.05.2019 00:11	Annotation
0014_tylem.wav	24.05.2019 00:11	WAV
0015_ustawilo.ant	24.05.2019 00:11	Annotation

Eksport danych

Możliwości:

- Eksport pojedynczych plików do arkusza,
- Eksport pojedynczej warstwy,
- Wszystkie pliki w kolekcji do wspólnego arkusza lub do odrębnych arkuszy
- Poszczególne segmenty do kolekcji plików wave & audio (np. zestaw bodźców percepcyjnych).



Podsumowanie

- Annotation Pro jest dostępny nieodpłatnie do badań naukowych
- Możliwość wykorzystania dyskretnych i ciągłych skal oceny
- Dostępność podstawowych opcji testu percepcyjnego
- Wielowarstwowa anotacja różnego rodzaju cech w obrębie tej samej kolekcji warstw / plików
- Rozszerzenia funkcjonalności dzięki wtyczkom, w szczególności: automatyzacja prac
- Program może służyć jako konwerter formatów plików popularnych programów do anotacji (Praat, ELAN, Wavesurfer, Transcriber, pliki tekstowe)
- Pozytywna informacja zwrotna w zakresie wygody pracy, działania na kolekcjach.

Dziękuję za uwagę

Katarzyna Klessa

E-mail: klessa@amu.edu.pl

- Abercrombie, D. (1968). Paralanguage. *International Journal of Language & Communication Disorders*, 3(1), 55-59.
- Bertrand, R., Ferré, G., & Guardiola, M. (2013). French face-to-face interaction: repetition as a multimodal resource. *Coverbal Synchrony in Human-Machine Interaction*, 141.
- Bänziger, T., & Scherer, K. R. (2007). Using actor portrayals to systematically study multimodal emotion expression: The GEMEP corpus. *International conference on affective computing and intelligent interaction* (pp. 476-487). Springer, Berlin, Heidelberg.
- Bergmann, K. & Kopp, S. (2012). Gestural alignment in natural dialogue. In Cooper, R.P. & Peebles, D. & Miyake, N. (Eds.) *Proceedings of the 34th Annual Conference of the Cognitive Science Society (CogSci 2012)*, 1326–1331. Austin, TX: Cognitive Science Society.
- Bigi, B., & Bertrand, R. (2016). Laughter in French spontaneous conversational dialogs. *Proceedings of 10th International Conference on Language Resources and Evaluation (LREC)*. 23-28 May 2016, Portorož, Slovenia.
- Gibbon, D. & Yu, J. (2015). Time Group Analyzer: Methodology And Implementation. *The Phonetician 111/112*:9-34.
- Bird, S., & Harrington, J. (2001). Speech annotation and corpus tools. *Speech Communication*, 33(1), 1-4.
- Bonacchi, S., Jarmołowicz-Nowikow, E., Karpiński, M., Klessa, K., Wicherkiewicz, T. (2018). Borderland: A Polish-German Multimodal Communication Database, Accepted at: *X International Conference on Corpus Linguistics (CILC) 2018*, Caceres, Spain.
- Branigan, H. P., Pickering, M. P., Cleland, Alexandra, A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75, 13–25.
- Brennan, S. E. i Clark, H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22, 1482–1493.
- Czoska, A., Klessa, K., Karpiński, M., Nowikow-Jarmołowicz, E. (2015). Prosody and gesture in dialogue: Cross-modal interactions. *Proceedings of 4th Gesture and Speech in Interaction (GESPIN) Conference*, Nantes, France, s. 83-88.
- Karpiński, M., Klessa, K., Czoska, A. (2014). Local and global convergence in the temporal domain in Polish task-oriented dialogue, *Proceedings of the 7th Speech Prosody Conference*, 20-23 May 2014, Dublin, Ireland. ISSN: 2333-2042.

- Karpiński, M., Czoska, A., Jarmołowicz-Nowikow, E., Juszczyk, K., Klessa, K. (2018). Aspects of gestural; alignment in task-oriented dialogues *Cognitive Studies 2018 (18)*. DOI: 10.11649/cs.1640
- Klessa, K., Gibbon, D. (2014). Annotation Pro+ TGA: automation of speech timing analysis. *Proceedings of Language Resources and Evaluation Conference (LREC)*, 1499-1505, Reykjavik.
- Klessa, K., Karpiński, M., & Wagner, A. (2013). Annotation Pro-a new software tool for annotation of linguistic and paralinguistic features. *Proceedings of the Tools and Resources for the Analysis of Speech Prosody (TRASP) Workshop*, Aix en Provence, 51-54.
- Klessa, K. & Korżinek, D. (2019). Annotation Pro + CLARIN-PL Align: automatic segmentation and transcription module for desktop uses. *Proceedings of 9th Language & Technology Conference*, Poznań, 294-298.
- Klessa K., Niećko-Bukowska B. (2017). Axio Speech Corpus: design, structure, and preliminary analyses. *Proceedings of The 8th Language and Technology Conference (Short papers)*, 209.
- Klessa, K., Wagner, A., Oleśkiewicz-Popiel, M., & Karpiński, M. (2013). Paralingua—a new speech corpus for the studies of paralinguistic features. *Procedia-Social and Behavioral Sciences*, 95, 48-58.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press, p. 209.
- Llisterri, J., (1992). Speaking styles in speech research, *ELSNET/ESCA/SALT Workshop on Integrating Speech and Natural Language*, Dublin, Ireland, 15-17.
- Wharton, T. (2003). Interjections, language, and the ‘showing/saying’ continuum. *Pragmatics & Cognition*, 11(1), 39-91, <http://ftp.phon.ucl.ac.uk/home/PUB/WPL/00papers/wharton.pdf>.
- Wharton, T. (2009). *Pragmatics and non-verbal communication*. Cambridge University Press.

W prezentacji wykorzystano niektóre efekty prac wymienionych w ramach projektów finansowanych z niżej wymienionych źródeł (zob. też odwołania na poszczególnych slajdach).



**NARODOWY PROGRAM
ROZWOJU HUMANISTYKI**

Projekt: NPRH12H 13 0524 82. Praca naukowa finansowana w latach 2014-2016. Język pogranicza – pogranicze języka Parajęzykowe aspekty komunikacji interkulturowej.



NARODOWE CENTRUM NAUKI

NeuroPerKog: rozwój słuchu fonematycznego i pamięci roboczej u niemowląt i dzieci



Narodowe Centrum
Badań i Rozwoju

Projekt: O R00 0170 12 pn. Pozyskiwanie i przetwarzanie informacji słownych w militarnych systemach zapobiegania oraz zwalczania przestępczości i terroryzmu

Projekt nr: O R00 0170 12