Practical remarks about the interoperability of the computer programmes Folker, ELAN and Praat for transcription and multimodal linguistic annotation from the user's point of view

Silvia Bonacchi, Mariusz Mela

University of Warsaw
Institute of Specialised and Intercultural Communication,
ul. Szturmowa 4, 02-678, Warsaw, Poland

Email(s):

s.bonacchi@uw.edu.pl, mariusz.mela@gmail.com

Abstract

The paper describes the experience of the MCCA research group with regards to the interoperability of Folker, ELAN and Praat computer programmes for multimodal linguistic annotation, describing the reasons for choosing them instead of other available software. Furthermore, from the point of view of users, the authors indicate the possible (technical) solutions that could facilitate the work of linguistic annotators of multilingual data.

Keywords: interoperability, spoken language, culturological analysis, suprasegmental analysis, multimodal analysis

1. Introduction

In the present paper we would like to present our experiences with the use and the interoperability of the computer programmes Praat, Folker and ELAN for transcription and multimodal linguistic annotation as part of the MCCA project ("Multimodal Communication: Culturological Analysis", **www.mcca.uw.edu.pl**, in full: "Culturological and Suprasegmental Analysis of Communicative Interactions Marked by (Im)Politeness", financed by the Polish National Centre for Science). It is a project based on collaboration between Polish and German scientific units (specifically the Institute of Specialised and Intercultural Communication at the University of Warsaw, and the Institute of Slavic Studies and the Institute of Computational Linguistics and Phonetics at Saarland University) for suprasegmental, multimodal and culturological analyses (see Müller 1998; Ogden 2006; Poggi 2007; McNeill 2005; Schmitt 2005; Bonacchi 2013; Bonacchi, Karpiński 2014), i.e. the analysis of vocal, verbal and kinetic displays (according to Sager 2004: 123ff.) of (im)polite behaviour as relevant communicative behaviour in several cultural settings.

The aim of the project is not only to transfer specialist (culturological and phonetic) knowledge and produce new knowledge about intra- and intercultural dialogue and mechanisms that disturb effective face-to-face communication, but also to develop standards of linguistic annotation for the Polish language that would be compatible with international tools for the description and analysis of speech data. Even though the primary aim of the project is to investigate suprasegmental cues and the culturological characteristics of interactions marked by polite and impolite behaviour (according to the second-order framework in Bonacchi 2013) in a corpus of digitalised Polish and German audio and video recordings (dyadic communication units), the annotation layers in ELAN have also been successfully used for the multimodal analysis of friendly and aggressive interactions in further pilot studies (see Bonacchi, Mela 2015).

2. ELAN, Folker and Praat in speech analysis

We have used the Praat (Boersma, Weenink 2015), Folker (Schmidt, Schütte, Hartung 2010) and ELAN (Sloetjes 2015) computer programmes in tandem for the annotation of speech data. We chose to use EL AN for several reasons. Firstly, it offers the possibility of integrating suprasegmental and verbal analysis with multimodal analysis thanks to its technical characteristics and its high degree of interoperability with other speech analysis programmes – for example it is interoperable to a high degree with EXMARaLDA (www.exmaralda.org), another programme we considered using. An important reason for preferring ELAN as an "umbrella-tool" (for the creation of complex annotations for video and audio resources) over other programmes was that it offers a very high degree of flexibility when it comes to defining the tier-structures (annotation tracks or layers) and thus defining the levels of linguistic analysis. Secondly, we wanted to work with an open system which could also be used both by students and for teaching purposes. These possibilities were also available with ANVIL (Annotation of Video and Language Data, www.anvil-software.org), another programme we considered using. ANVIL interfaces well with Praat because it allows the pitch contour and a waveform to be displayed. On the homepage of the programme (http://www.anvil-software.org/, last view: 20.6.2015) it has been announced that the forthcoming version of ANVIL will be compatible with ELAN files, permitting the user to switch between the programmes for optimised linguistic annotation and analysis. The main reason for choosing ELAN for our project was that it was already being used in other Polish research centres (for example by the Centre for Speech and Language Processing in Poznań for the DiaGest Corpus, http://cslp.wa.amu.edu.pl, a scientific unit which we work closely with, see Jarmołowicz-Novikow, Karpiński 2011).

While working on the project we have encountered various problems with the interoperability of ELAN, Folker and Praat, which we will describe in this paper**1**. At the same time we will indicate from the point of view of users the possible (technical) solutions that could facilitate the work of linguistic annotators of multilingual data.

2.1. Transcription: GTA2 conventions and Folker

A crucial moment in the annotation work turned out to be the transcription of verbal display and its connection with tiers related to the description of vocal display, for example annotation levels related to the use of voice, turn-taking, paraverbal features like backchannel and hesitation signals etc. Even though transcriptions in other Polish research groups (for example the Pelcra-research group at the University of Lodz, *Spokes*-corpus, s. Pęzik 2012) are carried out directly in ELAN, we have found it necessary to do the transcription as a separate step.

In order to have the transcription as a separate step, we have carried it out using Folker, a programme developed at the Institute for the German Language in Mannheim, Germany. The programme makes it possible to modify the transcription at any moment during further elaboration and to read it again as a tier in ELAN without compromising time-alignment.

At the moment Polish does not have a national standardised convention system for transcription like the GTA2 (Gesprächsanalytisches Transkriptionssystem, Selting et al.

[1] A shorter version of this paper was presented as a poster at the CLARIN–meeting CAC2014 in Soesterberg (Holland), 23-25.10.2014 (www.clarin.eu/sites/default/files/cac2014_submission_32_0.pdf).

2009) for German. The programmes for transcription developed by Polish research group are, despite being very promising, still in a phase of verification. Some proposals of transcription conventions for Polish for conversational analysis (e. g. Ranczew-Sikora 2007 and Frei 2013) are based on British/American ethnographic methods (ethnomethodology), which often turn out not to be suitable for wide multimodal analysis. Some research groups conduct their transcription in non-specific programmes, which are exportable in many text-formats, or, as mentioned, directly in ELAN. From Elan it is possible to export the text in many formats, for example in Excel-files (fig. 1).

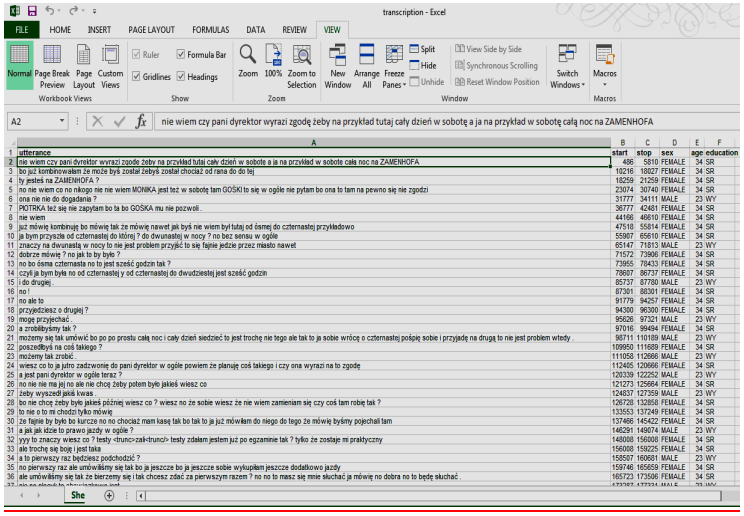

Figure 1: Transcription of a Polish dialogue, originally carried out in ELAN and exported as an Excel file (courtesy of Piotr Pęzik, Pelcra-Project, University of Łódź, http://clarin.pelcra.pl/Spokes/)

A very good example of software for the transcription of verbal data is Annotation Pro, developed at the University of Poznań (version 2.2.0.4., http://annotationpro.org/, see Klessa, Karpiński, Wagner, 2013). Annotation Pro is an innovative programme designed for the precise time-aligned transcription and annotation of audio recordings. It can work with multiple annotation tiers and offers functions that support fast manual segmentation and transcription. Besides graphic representations of speech signal (spectrogram, waveform view), it features various signal modes (fig. 2).

Moreover, Annotation Pro can host plug-ins for automatic or semi-automatic segmentation and transcription. It has a unique function that allows quasi-continuous variables to be used in annotations. Their values can be selected from pre-designed or user-designed graphical representations of one- or two-dimensional spaces (e.g. perceived pitch or heightened emotion in expressions). Additionally, it offers an experiment mode in which one can design and carry out perception tests. It can also provide basic statistics for annotations. Annotation Pro exports and imports annotation data from Praat as well as other popular systems. It has already been equipped with plug-ins and extensions for annotation analysis and processing. Finally, it can work with a user-defined workspace

which facilitates dealing with large corpora. As mentioned by Klessa and Karpiński (2012), it is planned to further develop Annotation Pro in order to include a video annotation mode with a host of necessary functions. Since Annotation Pro was in a phase of development when we started our project, it was only possible to run preliminary tests with the use of this tool. Based on these, we see the software as being potentially useful for multimodal annotation tasks in the future, provided that interoperability conditions are met.

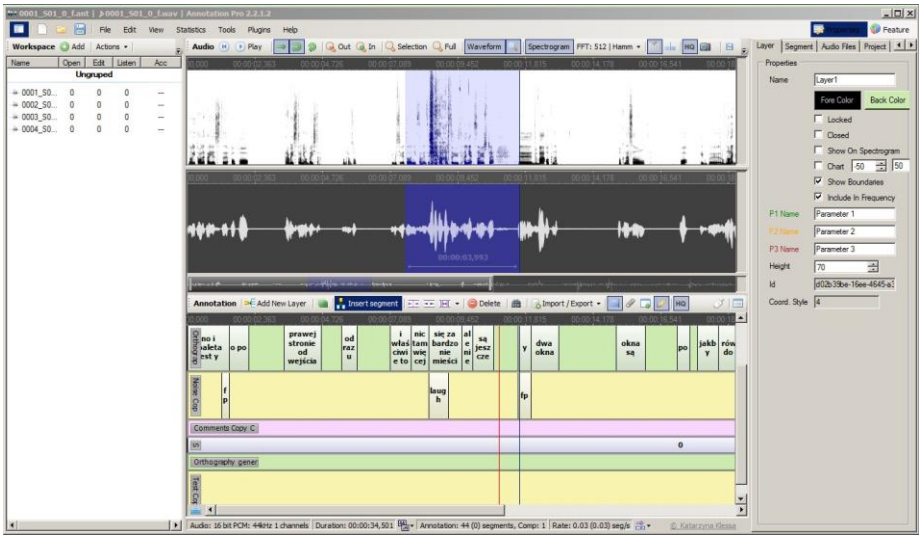

Figure 2: Transcription of a Polish dialogue using Annotation Pro
(courtesy of Maciej Karpiński, Adam Mickiewicz University in Poznań).

2.2. Folker to ELAN

For the transcription of the verbal display we used Folker (Fig. 3), developed by Schmidt Thomas, Schütte Wilfried, and Hartung Martin, which is particularly suitable for the GAT2-transcription conventions.

For our transcription of speech data we have used a system based mainly on GAT2 standard conventions for Basic Transcription, which has proved to be very suitable not only for German, but also for the Polish language. GAT2 conventions make it possible to refine the transcription by indicating prosodic and acoustic phenomena like speech pace, loudness, changes in intonation, pauses etc., and also turn taking dynamics, such as:

Simultaneous utterances (overlaps), e.g.:

B: <<*all*> what sense of [responsibility do you…]>

A: <<*f*> [quiet!]>

Latching:

= no interval between the end of the prior turn and the start of the next turn, e.g.

A: My eyes started tearing up =
= I started crying

Intervals between and within utterances:

(.) an estimated micropause of less than 0.2 seconds

(-) an estimated short pause of 0.2-0.5 seconds

Intonation contours at turn completion:

? clearly rising intonation

‘ rising intonation

↑ a mid turn sharp rise in intonation

↓ a mid turn sharp fall in intonation

<h> high tone of voice

Characteristics of speech delivery:

: a colon indicates extension of the preceding sound or syllable, e. g. *ba:d*

QUIET capital letters indicate focus stress and increased loudness

Dynamics of speech delivery:

<all> a fast manner of speaking

<acc> a speaker starts speaking faster

<f> a loud manner of speaking

<ff> a very loud manner of speaking

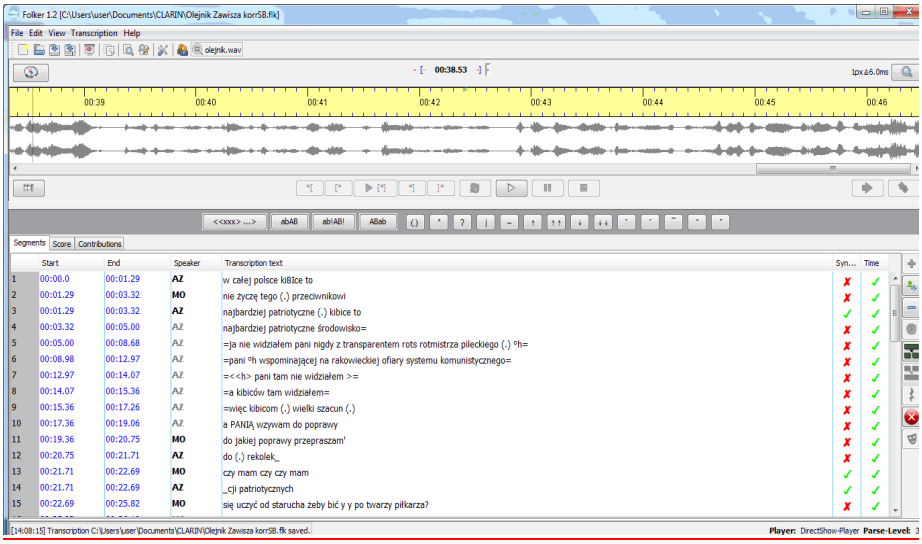

Figure 3: Folker-transcription of Polish data (a TV debate between Artur Zawisza and Monika Olejnik about an Independence Day march in Poland in 2012, http://www.tvn24.pl/kropka-nad-i,3,m/niesiolowski-zule-bili-policje-zawisza-pan-powiedzial-juz-wszystko,288294.html), segment-view

We have found the GAT2 conventions to be well-suited to transcribing Polish. Just a few small adjustments to Folker would facilitate transcription work with Polish speech data:

1. Polish diacritics (ą, ę, ń, ś, ł) direct on the Folker keyboard;

2. the introduction of special characters to mark sentence or word disruptions (anacoluthons) which are not always marked by a repair initiation;
3. special mark-up for repair initiation;
4. special mark-up "_ _" for interruptions due to overlaps, as shown in the following Folker output as a segment list:

{00:19} **0011 MO** do jakiej poprawy przepraszam'

{00:21} **0012 AZ** do (.) rekolek_

{00:22} **0013** [_cji patriotycznych]

{00:22} **0014 MO** [czy mam czy czy mam]

We have encountered some problems while adapting the GAT2 conventions to Polish, for example:

- different intonation patterns, which make it difficult to evaluate default (not marked) values for utterances;
- the lack of a consistent inventory of so-called 'filled pauses' (hesitation signals) and backchannel signals, which are different in various languages (e.g. the German 'ähm' vs. the Polish 'hmm' or the English 'umm');
- the annotation of non-lexical units, quasi-lexical units, and paralinguistic sounds (like clicks and percussives, see Karpiński 2012).

Folker makes it possible to show moments where the contributions of speakers overlap, which helps a conversation analyst to notice, among other things, regularities and interdependencies between speakers (e.g. which of the speakers − and when/how often − interrupts his/her interlocutor), and so it offers a preview of the linguistic phenomena that have to be annotated in the final tier-structures of ELAN.

The program also offers the possibility of producing an output of the transcription in *.html-format as a segment list or score, as a compact score with audio player file (fig. 4), as a contribution list, as a contribution list with audio player (fig. 5), as a GAT basic transcript (fig. 6) and as a "quantification" (quantitative data) (fig. 7).

Figure 4: Output as a compact score with audioplayer

```
Audiodatei: D:\KONGRESSE UND WORKSHOPS\CLARIN\MATERIALS AND FILES\olejnik.wav
Bedienung: [Einfacher Klick] - Audio an der betreffenden Stelle abspielen / [Doppelklick] - Audio stoppen / [Mouseover] - Über
[00:00] 001 AZ w całej polsce kibice to[najbardziej patriotyczne (.) kibice
            to] najbardziej patriotyczne środowisko==ja nie widziałem pani nigdy z
            transparentem rots rotmistrza pileckiego (.) °h=
[00:01] 002 MO [nie życzę tego (.) przeciwnikowi]
[00:03] 003 AZ =pani °h wspominającej na rakowieckiej ofiary systemu komunistycznego=
            =<<h> pani tam nie widziałem >==a kibiców tam widziałem==więc kibicom
            (.) wielki szacun (.)
[00:17] 004 AZ a PAnią (.) wzywam do poprawy
[00:19] 005 MO do jakiej poprawy przepraszam'
[00:21] 006 AZ do (.) rekolek_[_cji patriotycznych]
[00:22] 007 MO [czy mam czy czy mam]się uczyć od starucha żeby bić y y po twarzy
            piłkarza[tak?][mam się]mam się[bić tak?]niech pan mnie lepiej nie
            poucza dobrze'=[°niech pan niech pan poucza] [swoich kolegów z o en
            eru][żeby nie mówili]żeby nie[mówili]
[00:24] 008 AZ [nie] [niech się pa]
[00:27] 009 AZ [niech się]
[00:29] 010 AZ [niech się pani uczy][niech się pani uczy][od huSA:RII]
[00:34] 011 AZ [która]rozpoczynała[marsz niepodległości][od grup]rekonstruk[cyjnych]
[00:34] 012 MO [niech pan nie przesadza (.)] [naprawdę]
[00:38] 013 MO [a pan]niech się uczy[też]
[00:38] 014 AZ [en es zet]od motocyklistów rajdu katyńskiego niech PANI od nich się
            uczy
[00:43] 015 AZ naprawdę na naukę nigdy nie jest za późno
[00:46] 016 MO dla pana też nie jest yy za późno na naukę
```

Figure 5: Output as contribution list with audioplayer

```
[00:00] 0001 AZ w całej polsce kibice to
[00:01] 0002    [najbardziej patriotyczne (.) kibice to ]
[00:01] 0003 MO [nie życzę tego (.) przeciwnikowi]
[00:03] 0004 AZ najbardziej patriotyczne środowisko=
[00:05] 0005    =ja nie widziałem pani nigdy z transparentem rots rotmistrza pileckiego (.) °h=
[00:09] 0006    =pani °h wspominającej na rakowieckiej ofiary systemu komunistycznego=
[00:13] 0007    =<<h> pani tam nie widziałem >=
[00:14] 0008    =a kibiców tam widziałem=
[00:15] 0009    =więc kibicom (.) wielki szacun (.)
[00:17] 0010    a PAnią (.) wzywam do poprawy
[00:19] 0011 MO do jakiej poprawy przepraszam'
[00:21] 0012 AZ do (.) rekolek_
[00:22] 0013    [_cji patriotycznych ]
[00:22] 0014 MO [czy mam czy czy mam]
[00:23] 0015    się uczyć od starucha żeby bić y y po twarzy piłkarza
[00:26] 0016    [tak?]
[00:26] 0017 AZ [nie ]
[00:26] 0018 MO [mam się]
[00:26] 0019 AZ [niech się pa]
[00:27] 0020 MO mam się
```

Figure 6: Output as GAT basic transcript

	MO	AZ	Without	Total
Contributions (number)	6	10	0	16
Contributions (length)	22.21	36.25	0	58.46
Words (tokens)	0	0	0	0
Words (types)	0	0		0
Micro pauses	0	0	0	0
Non-phonological	0	0	0	0
Breathing	0	0	0	0
Measured pauses (number)	0	0	0	0
Measured pauses (length)	0	0	0	0

0 hours, 0 minutes, 46.81 seconds total transcribed time. **16** contributions, of which **16** with syntax errors and **0** with time errors.

Figure 7: Output as quantification

At the end of the transcription work with Folker, a final transcription format (*.flk) is obtained which can be exported to different formats compatible with the current programmes for further annotation: either as EXMARaLDA Basic Transcription (*.exb, *.xml), as an ELAN annotation file (*.eaf), as PRAAT TextGrid (*.textGrid), as F4 Transcript (*.rtf, *.txt), as an Audacity label file (*.txt), as a TEI file (*.xml) − which afterwards permits the user to mark up the text syntactically at any level of granularity −, or as plain text subtitles (*.txt).

After importing the Folker-transcription in eaf.format into ELAN, in which the scores of the speakers are displayed in separate tiers, we defined further annotation layers related to further levels of analysis (speech acts, vocabulary, types of sentences, PoS, voice, gestures, facial movements, etc.). To standardise analysis work within the research team we have created templates (MCCA-StandardTemplates) for linguistic annotation. In the following example (fig. 8) we have provided templates for the following description layers: event (description, not aligned with the signal), comments (description, not aligned with the signal), transcription (whole transcription, not aligned with the signal), verbal utterances of S(peakers) (words_S1, words_S2, words_S3, aligned to the signal), close transcription (close phonetic transcription, for example for backchannel-signals or hesitation signals, for lengthening cases etc., aligned with the signal), semantics (particular meaning of words, for example pejorative or meliorative forms, aligned with the signal), morphology (morphological information, aligned with the signal), translation (translation in

English, not aligned with the signal), suprasegmental features (aligned with the signals, for further analysis with Praat), intonation (aligned with the signals, for further analysis with Praat), smile (smile-voice and laughter, aligned with the signal), accents (for information structure, aligned with the signal), motion (body movements of the Speakers, aligned with the signal), gaze direction (facial movements and eye movements, aligned with the signal), axial direction (axial movements of the speakers, aligned with the signal), and gesture phases (gesture phases of the Speakers, aligned with the signal). Of course, the tier structure is the result of a work convention and can be adapted to different research aims.

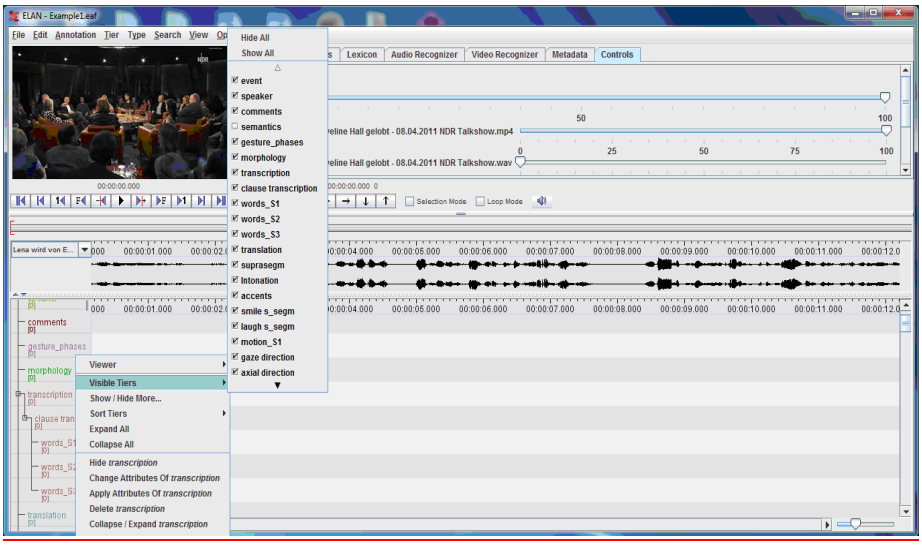

Figure 8: ELAN-view including MCCA-StandardTemplates

2.3. Folker and ELAN to Praat

Folker and ELAN files can also be exported as Praat-TextGrid files for further annotation (for example for the annotation of pitch, duration, intensity, intonation contours, characteristics of filled pauses, laughter, reductions, corrections, turn-taking etc.). For a graphic visualisation of the vocal display in the form of spectrograms and additional graphs based on the extraction of a range of values from the acoustic signal, we used Praat, which makes possible the creation of Praat images and the extraction of values related to the vocal performance.

Furthermore, thanks to the availability of a spectographic display of the speech signal, Praat makes a fine phonetic transcription possible − fundamental for such parameters as duration, lengthenings and hesitation phenomena, which are very important in (im)polite communication.

Unfortunately, it is not possible to read visual representations produced by Praat directly in ELAN. An integration of the functionalities of Praat and ELAN would permit the simultaneous use of the visual/graphical representations and the options supporting multi-modal annotation (multi-tier annotation, user-defined templates, and video display) without the need to import/export data between various tools. The importance of developing in the future such functionality for creating visual representations of the signal in ELAN or in other software for multimodal annotation, like Annotation Pro, could permit users to analyse directly in them, for example, the stress on the word "panią" in the Polish utterance "a PAnią (.) wzywam do poprawy" (English: "and I call on YOU (.) to change your behaviour!") (see fig. 9), which is very important for revealing the aggressive intention of Speaker.

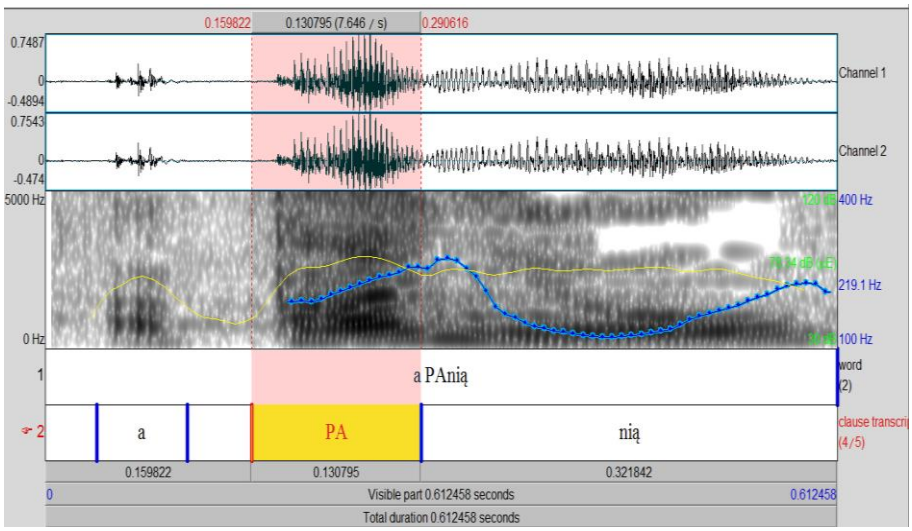

Figure 9: ELAN-files with the Polish word "PAnią" (English: "YOU" with a focus accent) exported to Praat for further investigation.

For further multimodal analysis of this conflict interaction we have defined the following tiers: words, translation, conflict initiator, argumentative structure, vocabulary, man's signals, woman's signals, gestures, face, voice, and type of situation (see fig. 10). The list is, however, not final or universal. Our ultimate aim is to develop a tier-structure for multimodal analysis which would permit the preparation of compatible data in several European languages (see Bonacchi, Mela, 2015).

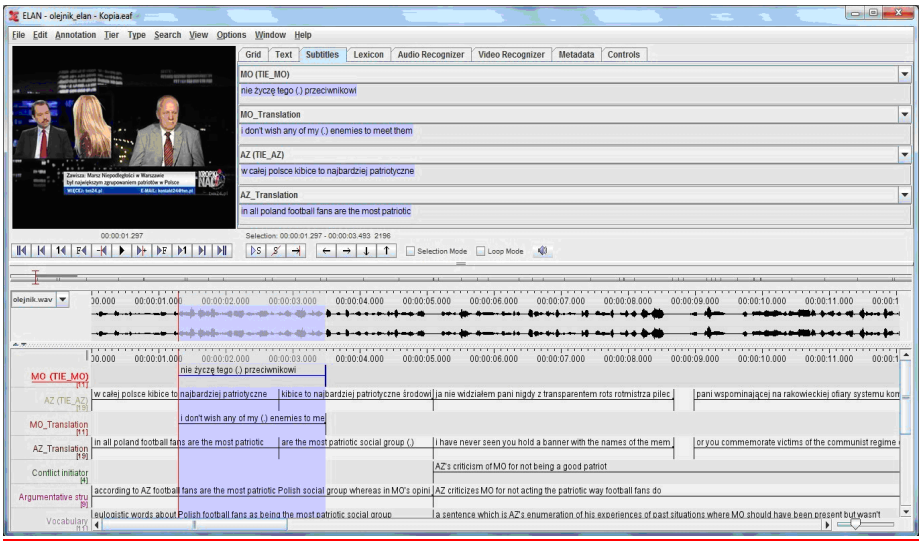

Figure 10: An example of a multimodal annotation in ELAN for the analysis of conflictual communicative behaviour.

3. Concluding remarks

The implementation of the tools described above would make a thorough analysis of face-to-face-interactions possible. The authors of the paper would like to draw the attention of software developers to the following aspects:
the introduction of a user-friendly (not only for scientific aims, but also for didactic purposes) unitary standard of transcription for European languages based on the Latin alphabet. For this aim the creation of widely flexible conventions which cover the whole range of linguistic phenomena in various spoken languages is necessary;

- the advantages of more interoperable IT tools which can permit a deep analysis of communicative displays, from verbal ones through the prosodic aspects to nonverbal communication;

- the benefits of investigating multilingual corpora in order to develop second-order frameworks for the annotation, comparison and explication of communicative phenomena.

Acknowledgements

MCCA is a project supported by a grant from the Polish National Science Centre (Narodowe Centrum Nauki, UMO-2012/04/ M/HS2/00551), which permitted the completion of the present paper.

References

Bonacchi, S. *(Un)Höflichkeit. Eine kulturologische Analyse Deutsch-Italienisch-Polnisch.* Frankfurt et al.: Peter Lang, 2013.

Bonacchi S, Karpiński M. "Remarks about the use of the term 'multimodality." In *Journal of Multimodal Communication Studies,* no. 1 (2014): 1-7.

Bonacchi S, Mela M. "Multimodal Analysis of Conflict: A proposal of a Dynamic Model." In *Conflict and Multimodal Communication,* by Francesca D'Errico, Isabella Poggi, Alessandro Vinciarelli, Laura Vincze, Berlin: Springer, 2015: 267-294.

Boersma P, Weenink D. *Praat* (version 5834_win64, www.fon.hum.uva.nl/praat), 2015.

Boersma P, Weenink D. *Praat: Doing phonetics by computer* [Computer program]. Version 5.3.51, retrieved 20 April 2015 from http://www.praat.org.

Demenko G, Klessa K, Szymański M, Breuer S, Hess W. "Polish unit selection speech synthesis with BOSS: Extensions and speech corpora." In *International Journal of Speech Technology,* 13(2). (http://link.springer.com/article/10.1007%2Fs10772-010-9071-3) 2010: 85-99.

Frei, R. "Analiza konwersacyjna – zarys metody." In *Via Communicandi*, edited by Beata Sierocka, Atut 2013, 35-51

Jarmołowicz-Nowikow E, Karpiński M. "Communicative intentions behind pointing gestures in task-oriented dialogues." In *Proceedings of GESPIN 2011 Conference: Gesture and Speech in Interaction Conference,* by Petra Wagner, Zofia Malisz, Caro Kirchhof. Bielefeld, 2011.

Karpiński M, Kleśta J. "The Project of Intonational Database for the Polish Language." In Prosody 2000, by Stanisław Puppel, Grażyna Demenko. Poznań: Faculty of Modern Languages and Literature UAM, 2001.

Karpiński, Maciej. "The Boundaries of Language: Dealing with Paralinguistic Features". In *Lingua Posnaniensis,* vol. LIV 2 (2012): 37-54.

Klessa K, Karpiński M, Wagner A. "Annotation Pro – a new software tool for annotation of linguistic and paralinguistic features". In *Proceedings of the Tools and Resources for the Analysis of Speech Prosody (TRASP) Workshop*, by Daniel Hirst, Brigitte Bigi. Aix en Provence, 2013: 51-54.

Klessa K, Karpiński M. "Annotating paralinguistic features in quasi-spontaneous speech. Adding the 'vision' component?" In *Proceedings of Workshop on Vision and Language, December 13th and 14th, 2012.* University of Sheffield UK, 2015.

McNeill, D. *Gesture and Thought.* Chicago: University of Chicago Press, 2005.

Müller, C. *Redebegleitende Gesten. Kultur – Theorie – Sprachvergleich.* Berlin: Arno Spitz, 1998.

Ogden, R. "Phonetics and social actions in agreements and disagreements." In *Journal of Pragmatics*, no. 38 (2006): 1752-1775.

Pęzik, P. "Język mówiony w NKJP." In *Narodowy Korpus Języka Polskiego,* by Adam Przepiórkowski, Mirosław Bańko, Rafał Górski, Barbara Lewandowska-Tomaszczyk. Warszawa, 2012: 37-47.

Poggi, I. *Mind, Hands, Face and Body: A Goal and Belief View of Multimodal Communication.* Berlin: Weidler Buchverlag, 2007.

Rancew-Sikora, D. *Analiza konwersacyjna jako metoda badania rozmów codziennych*. Warszawa, Trio, 2007

Sager, Sven F. *Kommunikationsanalyse und Verhaltensforschung. Grundlage einer Gesprächsethologie.* Tübingen, 2004.

Schmidt Ts, Schütte W, Hartung M. *Folker* (version 1.2., agd.ids-mannheim.de/folker.shtml). Mannheim: 2015.

Schmitt, R. "Zur multimodalen Struktur von turn-taking," In *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 6 (2005): 17-71. Accessed June 15, 2015.

Selting M, Auer P, Barth-Weingarten D, Bergmann J, Bergmann P, Birkner K et al. "Gesprächsanalytisches Transkriptionssystem 2 (GAT2)." In *Gesprächsforschung-Online-Zeitschrift zur verbalen Interaktion* 10 (2009): 353-402. Accessed June 15, 2015.

Sloetjes, H. *ELAN* (Eudico Linguistic Annotator, version 4.7.3, http://www.lat-mpi.eu/tools/elan).

Wells, J C. "SAMPA computer readable phonetic alphabet". In *Handbook of Standards and Resources for Spoken Language Systems,* Part IV, section B. (http://www.phon.ucl.ac.uk/home/sampa/, Polish: http://www.phon.ucl.ac.uk/home/sampa/polish.htm), by Dafydd Gibbon, Roger Moore, Richard Winski. Berlin and New York: Mouton de Gruyter, 1997.