



**HAL**  
open science

# SPPAS - the automatic annotation and analysis of speech (version 4.7)

Brigitte Bigi

► **To cite this version:**

Brigitte Bigi. SPPAS - the automatic annotation and analysis of speech (version 4.7). 2022. hal-01414600v8

**HAL Id: hal-01414600**

**<https://hal.science/hal-01414600v8>**

Preprint submitted on 21 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain



# SPPAS

the automatic annotation  
and analysis of speech



**Brigitte Bigi**

*contact@sppas.org*

*Copyright © 2011-2022 – Brigitte Bigi  
Laboratoire Parole et Langage – France*

*Permission is granted to copy, distribute and/or modify this document  
under the terms of the GNU Free Documentation License, Version 1.3  
or any later version published by the Free Software Foundation;  
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.  
A copy of the license is included in the section entitled "GNU  
Free Documentation License".*

*Any and all constructive comments are welcome.*

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	About this documentation . . . . .	1
1.2	What is SPPAS? . . . . .	1
1.2.1	Overview . . . . .	1
1.2.2	User engagement . . . . .	2
1.2.3	Need help . . . . .	2
1.2.4	About the author . . . . .	2
1.2.5	Licenses . . . . .	3
1.2.6	Supports . . . . .	3
1.3	Getting and installing . . . . .	4
1.3.1	Websites . . . . .	4
1.3.2	Download and install SPPAS . . . . .	5
1.3.3	The package . . . . .	5
1.3.4	Update . . . . .	6
1.4	Features . . . . .	6
1.4.1	How to use SPPAS? . . . . .	6
1.4.2	What SPPAS can do? . . . . .	7
1.5	Main and important recommendations . . . . .	7
1.5.1	About files . . . . .	7
1.5.2	About automatic annotations . . . . .	7
1.5.3	About linguistic resources . . . . .	7
<b>2</b>	<b>User interfaces</b>	<b>9</b>
2.1	Introduction . . . . .	9
2.2	Workspaces . . . . .	9
2.2.1	Overview . . . . .	9
2.2.2	Create and save a workspace . . . . .	9

---

2.2.3	Files	9
2.2.4	References	11
2.2.5	Associate files and references	11
2.3	The Command-Line user Interface - CLI	11
2.3.1	Overview	11
2.3.2	Usage	12
2.3.3	Arguments for input/output	12
2.3.4	preinstall.py	13
2.4	The graphical user interface	13
2.4.1	Important	13
2.4.2	Launch the GUI	13
2.4.3	The GUI of SPPAS	14
2.4.4	The main frame	14
<b>3</b>	<b>Automatic Annotations</b>	<b>19</b>
3.1	Introduction	19
3.1.1	About this chapter	19
3.1.2	THE recommended corpus construction workflow	19
3.1.3	Recordings	20
3.1.4	File formats and tier names	21
3.2	Resources required to annotate	21
3.2.1	Download and install linguistic resources	22
3.2.2	New language support	22
3.3	Annotate with the GUI	22
3.4	Annotate with the CLI	23
3.5	The procedure outcome report	24
3.6	Orthographic Transcription	26
3.7	Search for Inter-Pausal Units (IPUs)	27
3.7.1	Overview	27
3.7.2	The parameters	27
3.7.3	Perform “Search for IPUs” with the GUI	28
3.7.4	Perform “Search for IPUs” with the CLI	28
3.8	Fill in Inter-Pausal Units (IPUs)	30
3.8.1	Overview	30
3.8.2	How does it work	30

---

3.8.3	Perform “Fill in IPUs” with the GUI . . . . .	31
3.8.4	Perform “Fill in IPUs” with the CLI . . . . .	31
3.9	Text normalization . . . . .	32
3.9.1	Overview . . . . .	32
3.9.2	Adapt Text normalization . . . . .	32
3.9.3	Support of a new language . . . . .	32
3.9.4	Perform Text Normalization with the GUI . . . . .	33
3.9.5	Perform Text Normalization with the CLI . . . . .	33
3.10	Phonetization . . . . .	35
3.10.1	Overview . . . . .	35
3.10.2	Adapt Phonetization . . . . .	36
3.10.3	Support of a new language . . . . .	36
3.10.4	Perform Phonetization with the GUI . . . . .	36
3.10.5	Perform Phonetization with the CLI . . . . .	37
3.11	Alignment . . . . .	38
3.11.1	Overview . . . . .	39
3.11.2	Adapt Alignment . . . . .	39
3.11.3	Support of a new language . . . . .	39
3.11.4	Perform Alignment with the GUI . . . . .	40
3.11.5	Perform Alignment with the CLI . . . . .	40
3.12	Activity . . . . .	42
3.12.1	Overview . . . . .	42
3.12.2	Perform Activity with the GUI . . . . .	42
3.12.3	Perform Alignment with the CLI . . . . .	42
3.13	RMS . . . . .	42
3.13.1	Overview . . . . .	42
3.13.2	Perform RMS with the GUI . . . . .	43
3.13.3	Perform RMS with the CLI . . . . .	43
3.14	Interval Values Analysis - IVA . . . . .	43
3.14.1	Overview . . . . .	43
3.14.2	Perform IVA with the GUI . . . . .	43
3.14.3	Perform IVA with the CLI . . . . .	44
3.15	Lexical Metric . . . . .	44
3.15.1	Overview . . . . .	44
3.15.2	Perform Lexical Metric with the GUI . . . . .	44

---

3.16	Syllabification	44
3.16.1	Overview	44
3.16.2	Adapt Syllabification	45
3.16.3	Support of a new language	45
3.16.4	Perform Syllabification with the GUI	46
3.16.5	Perform Syllabification with the CLI	46
3.17	TGA - Time Groups Analyzer	47
3.17.1	Overview	47
3.17.2	Result of TGA into SPPAS	48
3.17.3	Perform TAG with the GUI	49
3.17.4	Perform TGA with the CLI	49
3.18	Stop words	50
3.19	Self-Repetitions	50
3.19.1	Overview	50
3.19.2	Adapt to a new language	51
3.19.3	Perform Self-Repetitions with the GUI	51
3.19.4	Perform SelfRepetitions with the CLI	51
3.20	Other-Repetitions	52
3.20.1	Overview	52
3.20.2	Adapt to a language and support of a new one	52
3.20.3	Perform Other-Repetitions with the GUI	53
3.20.4	Perform Other-Repetitions with the CLI	53
3.21	Re-Occurrences	53
3.21.1	Perform Re-Occurrences with the GUI	54
3.21.2	Perform Re-Occurrences with the CLI	54
3.22	Momel (modelling melody)	54
3.22.1	Perform Momel with the GUI	55
3.22.2	Perform Momel with the CLI	55
3.23	INTSINT: Encoding of F0 anchor points	57
3.23.1	Perform INTSINT with the GUI	57
3.23.2	Perform INTSINT with the CLI	57
3.24	Hand&Pose	59
3.24.1	Overview	59
3.24.2	Perform annotation with the GUI	60
3.24.3	Perform with the CLI	60

---

3.25	Face Detection	61
3.25.1	Overview	61
3.25.2	Result of Face Detection	61
3.25.3	Perform Face Detection with the GUI	62
3.25.4	Perform Face Detection with the CLI	62
3.26	Face Identity	64
3.26.1	Overview	64
3.26.2	Perform annotation with the GUI	64
3.26.3	Perform with the CLI	64
3.27	Face Landmarks	65
3.27.1	Overview	65
3.27.2	Perform annotation with the GUI	65
3.27.3	Perform with the CLI	65
3.28	Cued speech - LfPC	66
3.28.1	Overview	66
3.28.2	Perform annotation with the GUI	66
3.28.3	Perform with the CLI	66
<b>4</b>	<b>Convert files</b>	<b>69</b>
4.1	Interoperability and compatibility: an introduction	69
4.2	SPPAS conversion method	69
4.3	Supported file formats	70
4.4	Convert with the UI of SPPAS	70
<b>5</b>	<b>Data Analyses</b>	<b>71</b>
5.1	Introduction	71
5.2	The page Analyze of the GUI	71
5.2.1	Displayed files content	71
5.2.2	The toolbar	72
<b>6</b>	<b>Scripting with Python and SPPAS</b>	<b>77</b>
6.1	Introduction	77
6.2	A gentle introduction to programming	77
6.2.1	Introduction	78
6.2.2	Variables: Assignment and Typing	78
6.2.3	Basic Operators	79

---

6.2.4	Data types	79
6.2.5	Conditions	80
6.2.6	Loops	81
6.2.7	Dictionaries	82
6.3	Scripting with Python	82
6.3.1	Comments and documentation	83
6.3.2	Getting started with scripting in Python	83
6.3.3	Blocks	83
6.3.4	Functions	84
6.3.5	Reading/Writing files	85
6.3.6	Python tutorials	87
6.3.7	Exercises to practice	87
6.4	anndata, an API to manage annotated data	87
6.4.1	Overview	87
6.4.2	Why developing a new API?	87
6.4.3	The “anndata” API class diagram	88
6.5	Creating scripts with anndata	89
6.5.1	Preparing the data	89
6.5.2	Read/Write annotated files	89
6.5.3	Manipulating a sppasTranscription object	90
6.5.4	Manipulating a sppasTier object	90
6.5.5	Manipulating a sppasAnnotation object	91
6.5.6	Search in annotations: Filters	91
6.6	More with SPPAS...	94
<b>7</b>	<b>References</b>	<b>97</b>
7.1	References	97
7.1.1	How to cite SPPAS?	97
7.1.2	SPPAS software description	97
7.1.3	About linguistic resources	98
7.1.4	About Search for IPU	98
7.1.5	About Text Normalization	98
7.1.6	About Phonetization	98
7.1.7	About Forced-Alignment	99
7.1.8	About Syllabification	99

---

7.1.9	About Repetitions . . . . .	100
7.1.10	About analyses tools . . . . .	100
7.1.11	About the API . . . . .	100
7.1.12	Related references . . . . .	100
<b>8</b>	<b>SPPAS Release notes</b>	<b>103</b>
8.1	The early versions . . . . .	103
8.1.1	Version 1.0 . . . . .	103
8.1.2	Version 1.1 . . . . .	103
8.1.3	Version 1.2 . . . . .	103
8.1.4	Version 1.3 . . . . .	103
8.2	The birth of SPPAS . . . . .	104
8.2.1	SPPAS 1.4.0 . . . . .	104
8.2.2	SPPAS 1.4.1 . . . . .	105
8.2.3	SPPAS 1.4.2 . . . . .	106
8.2.4	SPPAS 1.4.3 . . . . .	106
8.2.5	SPPAS 1.4.4 . . . . .	107
8.2.6	SPPAS 1.4.5 . . . . .	107
8.2.7	SPPAS 1.4.6 . . . . .	108
8.2.8	SPPAS 1.4.7 . . . . .	108
8.2.9	SPPAS 1.4.8 . . . . .	109
8.2.10	SPPAS 1.4.9 . . . . .	109
8.2.11	SPPAS 1.5.0 . . . . .	110
8.2.12	SPPAS 1.5.1 . . . . .	110
8.2.13	SPPAS 1.5.2 . . . . .	110
8.2.14	SPPAS 1.5.3 . . . . .	111
8.2.15	SPPAS 1.5.4 . . . . .	111
8.2.16	SPPAS 1.5.5 . . . . .	112
8.2.17	SPPAS 1.5.6 . . . . .	112
8.2.18	SPPAS 1.5.7 . . . . .	112
8.2.19	SPPAS 1.5.8 . . . . .	113
8.2.20	SPPAS 1.5.9 . . . . .	113
8.2.21	SPPAS 1.6.0 . . . . .	114
8.2.22	SPPAS 1.6.1 . . . . .	114
8.2.23	SPPAS 1.6.2 . . . . .	114

8.2.24	SPPAS 1.6.3	115
8.3	The beginnings of SPPAS	115
8.3.1	SPPAS 1.6.4	115
8.3.2	SPPAS 1.6.5	116
8.3.3	SPPAS 1.6.6	116
8.3.4	SPPAS-1.6.7	118
8.3.5	SPPAS-1.6.8	118
8.4	The development phase	119
8.4.1	SPPAS-1.6.9	119
8.4.2	SPPAS-1.7.0	120
8.4.3	SPPAS-1.7.1	120
8.4.4	SPPAS-1.7.2	121
8.4.5	SPPAS-1.7.3	121
8.4.6	SPPAS-1.7.4	122
8.4.7	SPPAS-1.7.5	122
8.4.8	SPPAS-1.7.6	123
8.4.9	SPPAS-1.7.7	123
8.4.10	SPPAS-1.7.8	124
8.4.11	SPPAS-1.7.9	124
8.5	The stabilization phase	125
8.5.1	SPPAS-1.8.0	125
8.5.2	SPPAS-1.8.1	125
8.5.3	SPPAS-1.8.2	126
8.5.4	SPPAS-1.8.3	126
8.5.5	SPPAS 1.8.4	126
8.5.6	SPPAS 1.8.5	127
8.5.7	SPPAS 1.8.6	127
8.5.8	SPPAS 1.9.0	127
8.5.9	SPPAS 1.9.1	128
8.5.10	SPPAS 1.9.2	129
8.5.11	SPPAS 1.9.3	129
8.5.12	SPPAS 1.9.4	129
8.5.13	SPPAS 1.9.5	130
8.5.14	SPPAS 1.9.6	130
8.5.15	SPPAS 1.9.7	130

8.5.16	SPPAS 1.9.8	131
8.5.17	SPPAS 1.9.9	132
8.5.18	SPPAS 2.0	133
8.5.19	SPPAS 2.1	133
8.5.20	SPPAS 2.2	134
8.5.21	SPPAS 2.3	134
8.5.22	SPPAS 2.4	135
8.5.23	SPPAS 2.5	135
8.5.24	SPPAS 2.6	136
8.5.25	SPPAS 2.7	136
8.5.26	SPPAS 2.8	137
8.5.27	SPPAS 2.9	137
8.6	Migrate to Python 3	138
8.6.1	SPPAS-3.0	138
8.6.2	SPPAS-3.1	138
8.6.3	SPPAS-3.2	139
8.6.4	SPPAS-3.3	140
8.6.5	SPPAS-3.4	140
8.6.6	SPPAS-3.5	141
8.6.7	SPPAS-3.6	142
8.6.8	SPPAS-3.7	142
8.6.9	SPPAS-3.8	143
8.6.10	SPPAS-3.9	144
8.7	Stable	145
8.7.1	SPPAS-4.0	145
8.7.2	SPPAS-4.1	146
8.7.3	SPPAS-4.2	147
8.7.4	SPPAS-4.3	149
8.7.5	SPPAS-4.4	150
8.7.6	SPPAS-4.5	150
8.7.7	SPPAS-4.6	151
8.7.8	SPPAS-4.7	152

<b>9</b>	<b>Appendix</b>	<b>155</b>
9.1	List of error messages . . . . .	155
9.1.1	Global errors . . . . .	155
9.1.2	From “annotations” package . . . . .	155
9.1.3	From “utils” package . . . . .	156
9.1.4	From “audiodata” package . . . . .	156
9.1.5	From “calculus” package . . . . .	157
9.1.6	From “plugins” package . . . . .	157
9.1.7	From “resources” package . . . . .	157
9.1.8	From “structs” package . . . . .	158
9.1.9	From “models” package . . . . .	158
9.2	GNU Free Documentation License . . . . .	158

# Introduction

## 1.1 About this documentation

This documentation is governed by [GNU Free Documentation License, version 1.3](#). It will assume that you are using a relatively recent version of SPPAS. There's no reason not to download the latest version whenever released: it's easy and fast!

**Any and all constructive comments are welcome.**

## 1.2 What is SPPAS?

### 1.2.1 Overview

*SPPAS - the automatic annotation and analyses of speech* is a scientific computer software package. SPPAS is daily developed with the aim to provide a robust and reliable software. Available for free, with open source code, there is simply no other package for linguists to simply use in both the automatic annotations of speech, the analyses of any kind of annotated data and the conversion of annotated file formats. You can imagine the annotations or analyses you need, SPPAS does the rest!

**SPPAS ensures flexibility rather than a one-size-fits all approach in the implementation of the proposed features. There are no fixed or unique solutions, but a bunch of customizable features to make choices according to the user requirements.**

Annotating recordings is very labor-intensive and cost-ineffective since it has to be performed manually by experienced researchers with many hours of work. As the primary functionality, SPPAS proposes a set of **automatic or semi-automatic annotations of recordings**. In the present context, annotations are “defined as the practice of adding interpretative, linguistic information to an electronic corpus of spoken and/or written language data. ‘Annotation’ can also refer to the end-product of this process” (Leech, 1997). SPPAS automatizes the annotation processes and allows users to save time. In order to be used efficiently, SPPAS expects a rigorous methodology to collect data and to prepare them.

Linguistics annotation, especially when dealing with multiple domains, makes use of different tools within a given project. This implies a rigorous annotation framework to ensure compatibilities between annotations and time-saving. SPPAS annotation files are in a specific XML format with extension `xra`. **Annotations**

**can be imported from and exported to a variety of other formats** including Praat (TextGrid, PitchTier, IntensityTier), Elan (eaf), Transcriber (trs), Annotation Pro (antx), Phonedit (mrk), Sclite (ctm, stm), HTK (lab, mlf), subtitles formats (srt, sub), CSV files...

“[...] when multiple annotations are integrated into a single data set, inter-relationships between the annotations can be explored both qualitatively (by using database queries that combine levels) and quantitatively (by running statistical analyses or machine learning algorithms)” (Chiarcos 2008). As a consequence, the annotations must be time-synchronized: annotations need to be time-aligned in order to be useful for purposes such as analyses. Some special features are offered in SPPAS for **managing annotated files and analyzing data**. Among others, it includes a tool to filter multi-levels annotations. Other included tools are to estimate descriptive statistics, to manage annotated files, to manage audio files, etc. These data analysis tools of SPPAS are mainly proposed in the Graphical User Interface. However, advanced users can also access directly the Application Programming Interface, for example to estimate statistics or to manipulate annotated data.

### 1.2.2 User engagement

By using SPPAS, **you agree to cite a reference in your publications**. The full list of references is available in Chapter 7.

### 1.2.3 Need help

1. Many problems can be solved by updating the version of SPPAS.
2. When looking for more detail about some subject, one can search this documentation. This documentation is available in-line - see the SPPAS website, it is also included in the package in PDF format.
3. There are: a F.A.Q., tutorials and slides in the SPPAS web site.

### 1.2.4 About the author

Since January 2011, **Brigitte Bigi** is the main author of SPPAS. She has a tenured position of researcher at the French CNRS - [Centre National de la Recherche Scientifique](#). She's working since 2009 at Laboratoire Parole et Langage in Aix-en-Provence, France.

More about the author:

- <http://www.sppas.org/big>
- <https://orcid.org/0000-0003-1834-6918>

Contact the author by e-mail:

- to improve the quality of the linguistic resources;
- to add new linguistic resources;
- to help in development;
- to add new annotations or analysis methods;
- to propose to collaborate for a research purpose;
- to declare an issue;
- to send a constructive comment.

Do not contact the author if:

- you failed to install Python or wxPython;
- you didn't follow the tutorials;
- you didn't read both the documentation and the tutorials.

Possible e-mails are:

- *contact@sppas.org* for general purposes;
- *develop@sppas.org* for developer questions or for a bug alert.

## Contributors

Here is the list of other contributors in programming:

- April 2012-June 2012: Alexandre Ranson (1st website)
- April 2012-July 2012: Cazembé Henry (1st GUI)
- April 2012-June 2013: Bastien Herbaut (help system)
- March 2013-March 2014: Tatsuya Watanabe (annotated data support)
- April 2015-June 2015: Nicolas Chazeau (audio support)
- April 2015-June 2015: Jibril Saffi (annotated data support)
- April 2019-June 2019: Barthélémy Drabczuk (workspaces and num2letter)
- April 2020-June 2020: Laurent Vouriot (a SPEAKER annotation)
- April 2020-June 2020: Florian Hocquet (installer and video support)

### 1.2.5 Licenses

SPPAS software, except documentation and resources, are distributed under the terms of the [GNU GENERAL PUBLIC LICENSE v3](#).

Linguistic resources of SPPAS are either distributed:

- under the terms of the “GNU GENERAL PUBLIC LICENSE, v3”, or
- on the terms of the “Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Public License”.

See the documentation of the resources for details about individual license.

To summarize, SPPAS users are:

- free to study the source code and the resources of the software they use,
- free to share the software and resources with other people,
- free to modify the software and resources.

### 1.2.6 Supports

SPPAS is regularly supported by the Laboratoire Parole et Langage in many ways.

### 2011-2012:

Partly supported by ANR OTIM project (Ref. Nr. ANR-08-BLAN-0239), Tools for Multimodal Information Processing. Read more at: <http://www.lpl-aix.fr/~otim/>

### 2013-2015:

Partly supported by ORTOLANG (Ref. Nr. ANR-11-EQPX-0032) funded by the « Investissements d’Avenir » French Government program managed by the French National Research Agency (ANR). Read more at: <http://www.ortolang.fr/>

### 2014-2016:

SPPAS was also partly carried out thanks to the support of the following projects or groups:

- CoFee - Conversational Feedback <http://cofee.hypotheses.org>
- Variamu - Variations in Action: a MULTilingual approach <http://variamu.hypotheses.org>
- Campus France, Procore PHC.

### 2017-2020:

The introduction of Naija language is supported by the ANR [NaijaSynCor](#).

### 2019-2020:

The introduction of workspaces to manage files and the SPEAKER annotation type were both supported by the Vapvisio ANR project (ANR-18-CE28-0011-01).

### 2020:

Both the Proof of Concept of the Cued Speech automatic annotation and the package for the dependencies’ installation was supported by the LPL.

## 1.3 Getting and installing

### 1.3.1 Websites

The main website of SPPAS is located at the following URL:

<http://www.sppas.org>

Click on “Get it -> Download” to get access to the last release and to some recent ones.

The source code is hosted by SourceForge at:

<http://sppas.sf.net/>

### 1.3.2 Download and install SPPAS

The main website contains the `Download` page to download recent versions of the SPPAS software package.

#### Download and decompress the SPPAS package:

There is a unique version of SPPAS which does not depend on the operating system. SPPAS is ready to run, so it does not need elaborate installation. All you need to do is to copy the SPPAS package from the website to somewhere on your computer. Choose *a location with preferably only US-ASCII characters in the name - it obviously includes the path*. The package of SPPAS is compressed and zipped, so you will need to *decompress and unpack* it once you've got it.

#### Install external programs

An installation guide is available and must be followed carefully. Notice that administrator rights can be required.

#### To summarize it:

- STEP 1: install the recommended version of **Python 3.x ONCE**
- STEP 2:
  - GUI: launch the `setup.bat` (Windows) or `setup.command` (Linux, MacOS) to continue the installation in GUI mode and follows instructions.
  - CLI: you have first to create a Python venv with name `“sppaspyenv~”`, then use the program `sppas\bin\preinstall.py`. It allows installing external programs to enable some features, and the linguistic resources.

In case of difficulty arising from this installation, you're invited to consult the web first. It probably will provide the solution. If, however, the problems were to persist, ask a technician for help.

### 1.3.3 The package

SPPAS is a Research Software, distributed in the context of the **“Open Science”**. Then, unlike many other software tool, SPPAS is not distributed as an executable program. Instead, **everything is done so that users can check / change operation**.

It is particularly suitable for automatic annotations: it allows anyone to customize automatic annotations to its own needs and to implement its custom annotation solution.

The package of SPPAS is a directory with content as files and folders. It contains the followings:

- the `README.txt` file, which aims to be read!
- the files `setup.bat` and `setup.command` to install some external programs
- the files `sppas.bat` and `sppas.command` to launch the Graphical User Interface
- the `samples` directory contains data of various languages: they are proposed in order to test various features of SPPAS

- the `demo` folder contains an audio file, its corresponding orthographic transcription and the video file. It also contains a large amount of annotations automatically created by SPPAS.
- the `plugins` directory with one plugin. Others are available on the website.
- the `workspaces` directory with an example of a workspace
- the `sppas` directory contains the program itself, it's the SPPAS software itself
- the `resources` directory contains data that are used by automatic annotations (lexicons, dictionaries, ...)
- the `documentation` directory contains:
  - the terms of the licenses
  - the printable documentation
  - the printable resources documentation
  - the printable XRA file format documentation
  - the printable version of the main reference published in “the Phonetician” journal
  - the orthographic transcription convention
  - the folder `scripting_solutions` is a set of python scripts corresponding to the exercises proposed in the chapter “Scripting with Python and SPPAS”

### 1.3.4 Update

SPPAS is constantly being improved and new packages are published frequently (about 10 versions a year). It is important to update regularly in order to get the latest features and corrections.

Updating SPPAS consists of downloading, decompressing and launching the setup.

## 1.4 Features

### 1.4.1 How to use SPPAS?

There are three main ways to use SPPAS:

1. The Graphical User Interface (GUI) is as user-friendly as possible:
  - double-click on the `sppas.bat` file, under Windows;
  - double-click on the `sppas.command` file, under MacOS or Linux.
2. The Command-line User Interface (CLI), with a set of programs, each one essentially independent of the others, that can be run on its own at the level of the shell.
3. Scripting with Python and SPPAS provides the more powerful way.

Features of SPPAS can then be used either with a Command-line User Interface (CLI) or a Graphical User Interface (GUI). This latter requires wxPython to be installed but not the former. So, there's no specific difficulty by using this software.

Advanced users can also access directly the Application Programming Interface - API.

### 1.4.2 What SPPAS can do?

Features of SPPAS can be divided into 3 main categories:

1. Annotate
2. Analyze
3. Convert

The SPPAS website contains the list of the features and how most of the automatic annotations can be organized in an annotation workflow. See <http://www.sppas.org/features.html>

## 1.5 Main and important recommendations

### 1.5.1 About files

There is a list of important things to keep in mind while annotating with SPPAS. They are summarized as follows and detailed in the chapters of this documentation:

1. Speech audio files for automatic annotations:
  - only `wav` and `au` files are supported
  - only mono (= one channel) files are supported
  - frame rate is *preferably* 16000hz
  - bit rate is *preferably* 16 bits
  - good recording quality is expected. It is obviously required to never convert from a compressed file, as mp3 for example.
2. Annotated data files:
  - *UTF-8* encoding only
3. It is recommended to use only US-ASCII characters in file names, obviously it includes its path.

### 1.5.2 About automatic annotations

The quality of the results for most of the automatic annotations is highly influenced by **the quality of the data the annotation takes in input**. This is a politically correct way to say: *Garbage in, garbage out!*

Annotations are based on the use of linguistic resources. Resources for several languages are gently shared and freely available or downloadable. The quality of the automatic annotations is largely influenced by **the quality of the linguistic resources**.

### 1.5.3 About linguistic resources

**Users are of crucial importance for resource development.**

Do not hesitate to help in checking resource files and sharing your corrections with the community.

The users of SPPAS are invited to contribute to improve them. They can release the improvements to the public, so that the whole community benefits.

Any help is welcome to improve existing resources or to create new ones.

# User interfaces

## 2.1 Introduction

The current chapter describes how to manage workspaces, i.e a set of files and thier description. Then it describes how to use the Graphical User Interface (GUI) and the Command-line User Interface (CLI).

## 2.2 Workspaces

### 2.2.1 Overview

A workspace is a manager of a set of files and of manually defined references. Several workspaces can be created but only one at a time is enabled when using SPPAS. All the features of SPPAS are using such currently enabled workspace.

### 2.2.2 Create and save a workspace

A workspace is described in a configuration file into the folder “workspaces” of the SPPAS package. This file is made of the name of the workspace followed by the extension “.json”. It includes all properties of a workspace that are described here.

This configuration file can be copied into any other directory and re-imported into SPPAS.

### 2.2.3 Files

A workspace is managing a set of files. Each time a file is added into the currently selected workspace, its “root” and its “path” are extracted and stored into a tree-like architecture.

The data are structured as follow:

- a workspace contains a list of paths,
- each path contains the list of roots sharing this path,

- each root contains the list of filenames sharing this root.

For example, when adding the file “/sppas/samples/samples-eng/oriana1.wav” into the workspace, the following next actions will be performed:

- the path “/sppas/samples/samples-eng/” is added,
- the root “/sppas/samples/samples-eng/oriana1” is added into the path,
- the filename “/sppas/samples/samples-eng/oriana1.wav” is added into the root. Several properties of the file are also stored, including its size or the date of modification.

When SPPAS is creating a new file - like when annotating automatically, the filename is automatically added into the workspace. For example, when the automatic annotation “Fill in IPUs” has finished, the newly created file “/sppas/samples/samples-eng/oriana1.xra” is added into the workspace:

- the path “/sppas/samples/samples-eng/” is already existing so it’s not added,
- the root “/sppas/samples/samples-eng/oriana1” is not added into the path because it’s also already existing,
- the filename “/sppas/samples/samples-eng/oriana1.xra” is added into the root.

All files sharing the same root are then stored in the same place in the tree because there’re sharing properties: they are all related to the same recording. The only difference between files sharing the same root are:

1. the extension (.wav, .mp4, .mkv, .xra, .TextGrid, ...)
2. the pattern (-ipus, -token, -phon, -palign, -lexm, -orepet, ...)

Notice that a pattern:

- must start by the character “-”;
- must contain at least 1 character.

Thanks to this new management of files, the patterns used by the automatic annotations are not fixed anymore. Both the input and output patterns of the annotations can be modified at the time of configuring the annotations.

The second advantage of such management of files into a workspace is to add all files of a corpus only once into a workspace and to save it. It is then ready-to-use each time SPPAS is started.

To summarize, the file “/sppas/samples/samples-eng/oriana1-token.TextGrid” will be represented as follow:

1. “/sppas/samples/samples-eng/” is the *path*
2. “/sppas/samples/samples-eng/oriana1” is the *root*
3. “/sppas/samples/samples-eng/oriana1-token.TextGrid” is the *filename*, in which: - “-token” is the *pattern*, - “.TextGrid” is the *extension*.

The workspace is also managing the state of each of these items: checked or not for filenames, checked/at least one checked/not checked for roots or paths.

## 2.2.4 References

The “references” are information that can be associated to files.

There are 3 different types of references: STANDALONE, SPEAKER and INTERACTION. Each of them corresponds to an annotation type. The annotations of both SPEAKER and INTERACTION type require to associate roots with these references.

Each reference can contain a list of attributes with an identifier key to define them, a value and its type, and a description. Here is an example of a reference “Oriana” of type SPEAKER with its list of attributes in the form “key: values (type)”:

- `firstname: Oriana (str)`
- `gender: female (str)`
- `recorded_in: 2012 (int)`
- `L1: eng (str)`
- `L2: fra (str)`

It is recommended to use only us-ascii characters and no whitespace for the key.

## 2.2.5 Associate files and references

Roots and references should be associated. In the previous example, it could allow to declare the list of files related to a given speaker. In that example, the reference “Oriana” should be associated to the roots: “/sppas/samples/samples-eng/oriana1” and “/sppas/samples/samples-eng/oriana2”.

In the specific case of interactions, a reference of the interaction has to be created and associated to the roots of the filenames.

The definition of such references, and their link to the corresponding roots allow to perform an elaborated way to check files of a workspace. It’s very easy for example to check all files with file extension .wav of “female” gender more than 30 years old.

## 2.3 The Command-Line user Interface - CLI

### 2.3.1 Overview

A command-line user interface (CLI) is a means of interacting with a computer program where the user issues commands to the program in the form of successive lines of text (command lines). Command-line interfaces provide a more concise and powerful means to control the program than the GUI.

Operating system command line interfaces are called a command-line interpreter, command processor or shell. It displays a prompt, accept a “command line” typed by the user terminated by the Enter key, then execute the specified command and provide textual display of results or error messages. When a shell is active a program is typically invoked by typing its name followed by command-line arguments (if any).

Such **CLI programs are located in the `bin` folder** of the `sppas` directory of the package. All these programs are written with the programming language Python and are compatible with version 3.4+.

**Installing Python is then the only required other program in order to work with SPPAS CLI** except the `alignment.py` program which requires either Julius CSR or HVite (see installation instructions).

### 2.3.2 Usage

It is usual for a program to be able to display a brief summary of its parameters. Each program included in SPPAS provides its usage by using the option `--help`, as for example:

```
prompt> python .\sppas\bin\trsconvert.py --help
usage: trsconvert.py [files] [options]

... a program to export annotated files.

optional arguments:
  -h, --help      show this help message and exit
  --quiet         Disable the verbosity
  --debug        Highest level of verbosity

Files:
  -i file         Input annotated file name.
  -o file         Output annotated file name.

Options:
  -n value        Number of a tier (use as many -n options as wanted). Positive
                  or negative value: 1=first tier, -1=last tier.
  -t tiername     Name of a tier (use as many -t options as wanted).

This program is part of SPPAS version 2.0. Copyright (C) 2011-2019 Brigitte
Bigi. Contact the author at: contact@sppas.org
```

### 2.3.3 Arguments for input/output

In most of the programs, there is an option `-i` for the input file. There's no specific constraint on this file name. For example, the following program will execute the Momel automatic annotation:

```
python .\sppas\bin\momel.py -i .\samples\samples-eng\ENG_M15_ENG_T02.PitchTier
```

With this option `-i`, a name of an output file can be given with the option `-o`; if not, the main part of the result is printed on the standard output.

In several programs, an option `-I` can also be available to execute the program, and several files can be processed with this option. Moreover, there is some flexibility in file names with this option. SPPAS will search for the appropriate file from the given file name. For example, the next commands will process the same:

```
python .\sppas\bin\intsint.py -I .\samples\samples-eng\ENG_M15_ENG_T02.wav
python .\sppas\bin\intsint.py -I .\samples\samples-eng\ENG_M15_ENG_T02.PitchTier
python .\sppas\bin\intsint.py -I .\samples\samples-eng\ENG_M15_ENG_T02-momel.xra
```

With the option `-I`, the name of the output file is fixed and can't be changed. For example, the previous example will create a file with name `.\samples\samples-eng\ENG_M15_ENG_T02-intsint.xra`. An option `-e` allows to choose the extension of the file, `.xra` is the default one.

The options to manage input/output files can be summarized as follow:

```
Files (manual mode):
  -i file           An input file.
  -o file           Output file name (optionnal).

Files (auto mode):
  -I file           Input file (append).
  -e .ext           Output file extension. One of: .xra .TextGrid .eaf
                   .csv .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx
                   .arff .xrff
```

### 2.3.4 preinstall.py

This program allows to:

1. install external programs to enable some of the features of SPPAS – including the GUI;
2. install linguistic resources for the supported languages;
3. install resources required to enable some of the annotations.

## 2.4 The graphical user interface

### 2.4.1 Important

Once the installation is completed - including wxpython, the GUI feature is enabled. Do not launch the GUI of SPPAS if the installation is not finished, i.e. launch the setup first: see the tutorial series 2 and to follow instructions.

### 2.4.2 Launch the GUI

#### Under Windows

Once the SPPAS package is opened in the File Explorer, double-click on the `sppas.bat` file.

In recent versions of Windows (e.g. 10), the first time you try to run SPPAS you may get a window with title “Windows protected your PC” and the following message: “Windows SmartScreen prevented an unrecognized app from starting. Running this app might put your PC at risk. More info”. Click `More info` message and then `Run anyway` button. The file will now run SPPAS, and you will now no longer get a Windows protected your PC prompt when you run this specific file next time. This warning message comes from the fact that SPPAS is a free software and we did not paid to Microsoft commercial fees which would remove the “Unknown Publisher” warnings.

#### Under MacOS

Once the SPPAS package is opened in the Finder, double-click on the `sppas.command` file.

The first time you try to run SPPAS you may get a message: “sppas.command can’t be opened because it is from an unidentified developer.”. This warning message comes from the fact that SPPAS is a free software and we did not paid to Apple commercial fees. The solution is to run SPPAS with a right click (alt-click) on `sppas.command` file. This time you will get a message: “sppas.command is from an unidentified developer. Are you sure you want to open it?” Then click on `Open`. It will also now work each time you run it.

## Under Linux

Once the SPPAS package is opened in the File Explorer, double-click on the `sppas.command` file. Either, open a terminal and run it. Last solution is to directly execute `sppas/bin/sppasgui.py`.

### 2.4.3 The GUI of SPPAS

Both the main windows and a “Log window” will open. The main frame is made of a menubar at top, a main content in the middle and buttons for actions at bottom.

Most of the existing/incoming tutorials are explaining how to access features of SPPAS with this GUI. This chapter only summarizes some of them.

### 2.4.4 The main frame

The top menu allows to open various pages like “Files”, “Annotate” or “Analyze”. There are 2 ways to browse through the items of the menu: with the mouse or with the keyboard. To open a page with the mouse, just click on the item button in the menubar. To open a page with the keyboard, under Windows, press CTRL and under MacOS, press COMMAND. Click left-right arrows to navigate to the previous/next page. Use up-arrow to go to the 1st page and Down-arrow to the last page. Use “CTRL+F”/“COMMAND+F” to go to the “Files” page.

To change colors and fonts, click the `Settings` icon at bottom. These settings are saved in a file which is used each time SPPAS is executed.

The `About` action button allows displaying information about SPPAS: author, license, a link to the web site, etc.

## Home page

The home page shows a welcome message and gives a quick access to a set of external links. Clicking on the buttons will open the default web-browser of the OS to this link.

The links are pointing to:

- the SPPAS home web site: <http://www.sppas.org>
- the in-line version of this documentation
- the in-line tutorials
- the frequently asked questions
- the home page of the author, including the whole list of publications related to SPPAS: <http://www2.lpl-aix.fr/~bigi/>

## Files page

The page “Files” is dealing with the workspaces (see the previous section about workspaces).

**Workspaces:** A column at left with green buttons contains a list of buttons to perform actions related to the workspaces then the list of existing ones.

The actions to perform are:

1. “Import from”: select a workspace file (.wjson) somewhere on the computer and add it into the SPPAS workspaces directory.
2. “Export to”: save the currently enabled workspace somewhere in the computer
3. “Pin & Save”: The currently selected workspace is saved. If the currently selected workspace is “Blank”, clicking this action will ask for a workspace name and create it (“Pin”), then select it.
4. “Rename”: change the name of the currently selected workspace. Can’t be applied on the “Blank” workspace.

The list of workspaces allows to select a workspace among the existing ones or to work with the “Blank” one (default).

**Files:** The files area with a pinky color displays the tree of files of the currently selected workspace, and a toolbar with action buttons.

The actions to perform are:

1. “Add”: add new files into the workspace
2. “Remove checked”: remove the checked files of the workspace
3. Remove missing files of the workspace. The tree of paths/roots/files is explored and if a filename does not correspond to an existing file on the disk, it is removed of the workspace tree
4. All files of the disk sharing the checked roots will be added to the workspace tree

**References:** The “References” area with a blue color displays the tree of references of the currently selected workspace, and a toolbar with action buttons.

The actions to perform are:

1. “Create”: Open a dialog to ask for a name and the reference type among: STANDALONE, INTER-ACTION, SPEAKER. It then creates the corresponding reference and add it to the workspace
2. “Edit”: Open a dialog to delete, add or modify the attributes of the checked references.
3. “Delete”: Definitively delete the checked references. There’s no way to recover them if needed.

**Column of action buttons:** Between the files and references, there’s a column with action buttons. The actions to perform are related to files:

1. Open the checked files with a text editor (under development)
2. “Delete checked”: remove the checked files of the workspace *and* delete them of the disk. Actually, there are moved and renamed into the trash of SPPAS. To recover a mistaken deleted file, use your OS explorer.
3. Check/uncheck all files
4. Check only files matching some patterns.

or to link files and references:

5. Associate the checked roots to the checked references
6. Dissociate the checked roots of the checked references

### **Annotate page**

This page allows annotating checked files of the currently enabled workspace.

The automatic annotation is a 5-steps process:

1. Choose the output file format. All the automatic annotations are creating files. Choosing the extension means to choose the format of the resulting file. Some formats are more “powerful” than others, some are more “sharable”...
2. Fix the language. Some annotations are language-independent, so they don’t need a specific language resource, like “Search for IPUs” or “Face detection”. But some annotations are requiring linguistic resources like a lexicon (Text Normalization), a dictionary (Phonetization), etc. This step allows to choose the language among the list of installed ones.
3. Select and configure annotations. Click any of the STANDALONE, SPEAKER or INTERACTION button. It will change the page content in order to select the annotations SPPAS will perform and to configure them. Click the top arrow to come back to the main page.
4. Perform the annotations. Click the “Let’s go” button to launch the automatic annotation process, and wait!
5. A procedure outcome report is displayed and saved in the logs folder of the SPPAS package.

The previously saved reports are listed in the column at right of the page. Click the button “Show report” to display the content of the checked one.

### **Analyze page**

This page is a merge of the “DataFilter”, the “DataStats” and the “DataRoamer” of the previous GUI version... with new features added.

It allows to display a summary of each file: either audio or annotated files can be opened and analyzed. The following analyses can be performed:

- manage tiers: duplicate, remove, copy, move, ...
- estimate distribution statistics on tiers;
- create new tiers with the filter systems, either the “simple” or the ”relation one”;
- check if an audio file is compatible with the automatic annotations.

### **Editor page**

This page is a merge the “IPUScriber” and the “Visualizer” of the previous GUI version... with new features added.

It allows to annotate manually and to view files in a timeline. The features “video” is required to play a video file and the feature “audioplay” is required to play an audio file.

This editor is very useful to play several media at a time synchronously.

## Convert page

The convert page allows to export the checked annotated files of the workspace in any of the supported file format.

The displayed table indicates all the format capabilities, including:

- Metadata: the file format supports metadata, i.e. it can save extra-information, not only annotations. If not, the metadata won't be exported;
- Multi tiers: the file format accepts to save more than one tier;
- No tier: the file format can save a file without tiers;
- Point: the file format can save tiers of type 'Point';
- Interval: the file format can save tiers of type 'Interval'.
- Gaps: the file format supports gaps between annotations of a tier. If not, when exporting the file, SPPAS will fill in the gaps by annotations with an empty label;
- Overlaps: the file format supports overlaps between annotations of a tier. If not, when exporting the file, SPPAS will split annotations to create a sequence of non-overlapping ones.
- Hierarchy: the format can deal with a hierarchy of tiers. If not, the hierarchy won't be saved.
- Ctrl vocab: the format can save a controlled vocabulary of tiers. If not, the controlled vocabularies won't be saved.
- Media: the format can save a reference to a media of tiers. If not, the media information won't be saved.
- Vagueness: the format supports a vagueness value of the localization. If not, the vagueness value won't be saved.
- Alt. loc: the format supports alternative annotation localizations. If not, the alternative localizations won't be saved.
- Alt. tag: the format supports alternative annotation tags. If not, when exporting the annotations, SPPAS will use a "{ | }" system to parse the alternative tags and = for scores.

## Plugins page

Installing plugins is a very useful solution to extend the features of SPPAS. Several plugins are available for download in the main site of SPPAS. The plugins of SPPAS are installed in a folder with name "plugins" in the main directory of the SPPAS package.

To install a new plugin, simply follow this workflow:

1. Create or download the plugin package - e.g. a zip file.
2. Execute SPPAS.
3. Click on 'Plugin' then click on the 'Install' button of the toolbar.
4. Browse to indicate the plugin package.
5. See the new plugin at bottom of the plugins list.

To delete a plugin, click on the "Delete" button of the toolbar. Choose the plugin in the given list then click on the "OK" button. Notice that the plugin is definitively deleted of the disk.

To execute a plug-in, check the file(s) of your current workspace in the File page, then, click on the icon of the plug-in and follow instructions of the plugged program.



# Automatic Annotations

## 3.1 Introduction

### 3.1.1 About this chapter

This chapter describes the followings:

- what is the goal of the annotation;
- what are the requirements (files, tiers);
- what kind of resources are needed;
- what is the expected result;
- how to perform it within SPPAS.

However, this chapter is not a description on how each automatic annotation is working and what to expect about it: there are references for that in chapter 8.

Among others, SPPAS is able to produce automatically annotations from a recorded speech sound and its orthographic transcription. Let us first introduce what is **the recommended way to annotate a corpus with SPPAS**.

### 3.1.2 THE recommended corpus construction workflow

The kind of process to implement in the perspective of obtaining rich and broad-coverage multimodal/multi-levels annotations of a corpus is illustrated in next Figure. It describes each step of the annotation workflow. Obviously, there are other ways to construct a corpus but 1/ do not blame SPPAS if you don't get the results you expected, and 2/ do not contact the author if you have chosen to not follow these recommendations.

This Figure must be read from top to bottom and from left to right, starting by the recordings and ending to the analysis of annotated files. Yellow boxes represent manual annotations, blue boxes represent automatic ones.

Annotation methodology

After the recording an audio file (see recordings recommendations), the first annotation to perform is to search for the IPU's. Indeed, at a first stage, the audio signal must be automatically segmented into Inter-Pausal Units (IPUs) which are sounding segments surrounded by silent pauses of more than X ms, and time-aligned on the speech signal.

An orthographic transcription has to be performed **manually inside the IPU's** (do not expect to use an automatic speech transcription system). Then text normalization automatic annotation will normalize the orthographic transcription. The phonetization process will convert the normalized text in a grammar of pronunciations using X-SAMPA standard. Alignment will perform segmentation at phonemes and tokens levels, etc.

At the end of each automatic annotation process, SPPAS produces a *Procedure Outcome Report*. It contains important information about the annotations. It includes all parameters and eventually warnings and errors that occurred during the annotation process. This window opens in the scope to be read by users (!) and should be saved with the annotated corpus.

### 3.1.3 Recordings

SPPAS performs automatic annotations: It does not make sense to hope for miracles but you can expect good enough results that will allow you to save your precious time! *And it begins by taking care of the recordings...*

#### Audio files

Only `wav` and `au` audio file formats are supported by Python, so does SPPAS;

Only mono audio files are supported by automatic annotations of SPPAS.

SPPAS verifies if the audio file is 16 bits sample rate and 16000 Hz frame rate. Otherwise it automatically creates a new converted audio file. For very long files, this process may take time. If Python can't read the audio file, an error message is displayed: you'll have to convert it with audacity, praat... A relatively good recording quality is expected (see next Figure).

For example, both "Search for IPU's" and "Fill in IPU's" require a better quality compared to what is expected by "Alignment", and for that latter, it depends on the language. The quality of the result of automatic annotations highly depends on the quality of the audio file.

Example of expected recorded speech

Providing a guideline or recommendation of good practices is impossible, because it depends on too many factors. However, the followings are obvious:

- Never, never, never records a lossy audio file. It means that extracting the audio file from a video is feasible only if the embedded audio is either lossless or not compressed: see this page [https://en.wikipedia.org/wiki/Comparison\\_of\\_video\\_container\\_formats](https://en.wikipedia.org/wiki/Comparison_of_video_container_formats).
- The better microphone, the better audio file! Using an headworn microphone is much more better than a clip-on one. At LPL, we get very good results with the AKG C520.
- The recorded volume must be high enough. Ideally, it should be in range [-0.5 ; 0.5]. If all the amplitude values are in range [-0.1 ; 0.1], the difference between speech and silence is very slight and it makes the search for silences very difficult.

- The audio file should not be 32 floating bits. For speech, 32 bits are totally unusefull and - worse, sometimes Python can't read it.
- as you probably don't plan to burn your audio file on a CD-ROM, 44100Hz framerate does not make sense. 48000hz is a more reasonable choice, particularly because it doesn't need of elabotated interpolation methods when it's converted to 16000hz for automatic annotations.

## Video files

SPPAS is proposing a few automatic annotations of a video if the Python library "opencv" is installed. All of them are annotating the face of recorded people.

### 3.1.4 File formats and tier names

When annotating with the GUI, the file names of each annotation is fixed and can't be customized. A filename is made of a root, followed by a pattern then an extension. For example "oriana1-align.TextGrid" is made of the root "oriana1", the pattern "-align" and the extension ".TextGrid". Each annotation allows to fix manually the pattern and to choose the extension among the list of the supported ones. Notice that the pattern must start with the "-" (minus) character. It means that the character "-" must only be used to separate the root to the pattern:

The character '.' can't be used in the root of a filename.

The name of the tiers the annotations are expecting for their input are fixed and can't be changed; so does the produced tier names.

## 3.2 Resources required to annotate

All the automatic annotations proposed by SPPAS are designed with language-independent algorithms, but some annotations are requiring language-knowledges. This linguistic knowledge is represented in external files so they can be added, edited or removed easily.

Adding a new language for a given annotation only consists in adding the linguistic resources the annotation needs, like lexicons, dictionaries, models, set of rules, etc. For exemple, see:

Mélanie Lancien, Marie-Hélène Côté, Brigitte Bigi (2020). Developing Resources for Automated Speech Processing of Quebec French. In Proceedings of The 12th Language Resources and Evaluation Conference, pp. 5323–5328, Marseille, France.

Brigitte Bigi, Bernard Caron, Abiola S. Oyelere (2017). Developing Resources for Automated Speech Processing of the African Language Naija (Nigerian Pidgin). In 8th Language and Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, pp. 441-445, Poznań, Poland.

### 3.2.1 Download and install linguistic resources

Since June 2020, the linguistic resources and models for some annotations are no longer distributed into the package of SPPAS. Instead, they are hosted by Ortolang repository with a public access.

They can be installed automatically into SPPAS by the `preinstall.py` program (CLI) or in the GUI by clicking “Add languages” or “Add annotations” in the toolbar of the “Annotate” page.

They can also be installed manually by downloading them at: <https://hdl.handle.net/11403/sppasresources> and unpacking the zip file into the `resources` folder of SPPAS package.

A full description of such resources and how to install them is available in the repository: download and read the file `Documentation.pdf`. It contains details about the list of phonemes, authors, licenses, etc.

### 3.2.2 New language support

Some of the annotations are requiring external linguistic resources in order to work efficiently on a given language: text normalization requires a lexicon, phonetization requires a pronunciation dictionary, etc. It is either possible to install and use the existing resources or to create and use custom ones.

When executing SPPAS, the list of available languages of each annotation is dynamically created by exploring the `resources` directory content. This means that:

- the resources you added or modified are automatically taken into account (ie. there’s no need to modify the program itself);
- SPPAS needs to be re-started if new resources are added when it was already being running.

## 3.3 Annotate with the GUI

Performing automatic annotations with SPPAS Graphical User Interface is a step-by-step process.

It starts by checking the list of paths and/or roots and/or files in the currently active workspace of the “Files” page. Then, in the “Annotate” page:

1. Select the output file format, i.e. the file format of the files SPPAS will create;
2. Select a language in the list;
3. Enable each annotation to perform by clicking on the button in red, among `STANDALONE`, `SPEAKER` and `INTERACTION` annotation types. Each button will be turned green if some annotations are selected.
  - 3.1 Configure each annotation by clicking on the “Configure...” link text in blue;
  - 3.2 The language of any annotation can be changed.
4. Click on the *Perform annotations* button, and wait. A progress bar should indicates the annotation steps and files. Some annotations are very very fast but some others are not. For example, Face Detection is 2.5 x real times, i.e. annotating a video of 1 minute will take 2 minutes 30 secs.
5. It is important to read the Procedure Outcome report. It allows to check that everything happened normally during the automatic annotations. This report is saved in the “logs” folder of the SPPAS package.

### 3.4 Annotate with the CLI

To perform automatic annotations with the Command-line User Interface, there is a main program `annotation.py`. This program allows to annotate in an easy-and-fast way but none of the annotations can be configured: their default parameters are used. This program performs automatic annotations on a given file or on all files of a directory. It strictly corresponds to the button `Perform annotations` of the GUI except that annotations are pre-configured: no specific option can be specified.

```
usage: python .\sppas\bin\annotation.py -I file|folder [options]
```

optional arguments:

```
-h, --help          show this help message and exit
--log file          File name for a Procedure Outcome Report (default: None)
--momel            Activate Momel
--intsint          Activate INTSINT
--fillipus         Activate Fill in IPUs
--searchipus       Activate Search for IPUs
--textnorm         Activate Text Normalization
--phonetize        Activate Phonetization
--alignment        Activate Alignment
--syllabify        Activate Syllabification
--tga              Activate Time Group Analysis
--activity         Activate Activity
--rms              Activate RMS
--selfrepet        Activate Self-Repetitions
--stopwords        Activate Stop Tags
--lexmetric        Activate LexMetric
--otherrepet       Activate Other-Repetitions
--reoccurrences    Activate Re-Occurrences
--merge            Create a merged file with all the annotations
```

Files:

```
-I file|folder      Input transcription file name (append).
-l lang             Language code (iso8859-3). One of: por eng ita kor deu nan
                   vie und hun spa cat pol yue fra pcm yue_chars cmn jpn.
-e .ext             Output file extension. One of: .xra .TextGrid .eaf .csv
                   .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx .arff .xrff
```

Examples of use:

```
./sppas/bin/annotation.py -I .\samples\samples-eng
                           -l eng
                           -e .TextGrid
                           --fillipus --textnorm --phonetize --alignment
```

A progress bar is displayed for each annotation if the Terminal is supporting it (`bash` for example). Instead, the progress is indicated line-by-line (`Windows PowerShell` for example).

CLI: `annotation.py` output example

Each annotation has also its own program and all options can be fixed. They are all located in the `sppas/bin` folder.

### 3.5 The procedure outcome report

It is very important to read conscientiously this report: it describes exactly what happened during the automatic annotation process. It is recommended to store a copy of the report within the corpus because it contains information that are interesting to know for anyone using the annotations.

By default, all reports are saved in the “logs” folder of the SPPAS package.

The text first indicates the version of SPPAS that was used. This information is very important. Annotations in SPPAS and their related resources are regularly improved and then, the result of the automatic process can change from one version to the other one.

Example:

```
SPPAS version 3.5
Copyright (C) 2011-2021 Brigitte Bigi
Web site: http://www.sppas.org/
Contact: Brigitte Bigi (contact@sppas.org)
```

Secondly, the text shows information related to the given input:

1. the selected language of each annotation - only if the annotation is language-dependent. For some language-dependent annotations, SPPAS can still perform the annotation even if the resources for a given language are not available: in that case, select “und”, which is the iso639-3 code for “undetermined”.
2. the selected files and folder to be annotated.
3. the list of annotations, and if each annotation was enabled. In that case, enabled means that the checkbox of the annotation was checked by the user and that the resources are available for the given language. On the contrary, disabled means that either the checkbox was not checked or the required resources are not available.
4. the file format of the resulting files.

Example:

```
Date: 2020-04-21T11:14:01+02:00
Input languages:
- Momel: ---
- INTSINT: ---
- Fill in IPUs: ---
- Search for IPUs: ---
- Text Normalization: eng
- Phonetization: eng
- Alignment: eng
- Syllabification:
- Time Group Analysis: ---
- Activity: ---
- RMS: ---
- Self-Repetitions:
- Stop Tags:
- LexMetric: ---
- Other-Repetitions:
```

```
- Re-Occurrences: ---

Selected files and folders:
- oriana1.wav

Selected annotations:
- Momel: enabled
- INTSINT: enabled
- Fill in IPUs: enabled
- Search for IPUs: disabled
- Text Normalization: enabled
- Phonetization: enabled
- Alignment: enabled
- Syllabification: disabled
- Time Group Analysis: disabled
- Activity: disabled
- RMS: disabled
- Self-Repetitions: disabled
- Stop Tags: disabled
- LexMetric: disabled
- Other-Repetitions: disabled
- Re-Occurrences: disabled

File extension: .xra
```

Thirdly, each automatic annotation is described in details, for each annotated file. At a first stage, the list of options and their value is summarized. Example:

#### Text Normalization

```
The vocabulary contains 121250 tokens.
The replacement dictionary contains 8 items.
Options:
... inputpattern:
... outputpattern: -token
... faked: True
... std: False
... custom: False
... occ_dur: True
```

Then, a diagnosis of the given file is printed. It can be: 1. “Valid”: the file is relevant 2. “Admit”: the file is not like expected but SPPAS will convert it and work on the converted file. 3. “Invalid”: SPPAS can’t work with that file. The annotation is then disabled. In case 2 and 3, a message indicates the origin of the problem.

Then, if any, the annotation procedure prints messages. Four levels of information must draw your attention:

1. “[ OK ]” means that everything happened normally. The annotation was performed successfully.
2. “[ IGNORE ]” means that SPPAS ignored the file and didn’t do anything.
3. “[ WARNING ]” means that something happened abnormally, but SPPAS found a solution, and the annotation was performed anyway.
4. “[ ERROR ]” means that something happened abnormally and SPPAS failed to found a solution. The annotation was either not performed, or performed with a wrong result.

Example of “Warning” message:

```
... .. Export AP_track_0711.TextGrid
... .. into AP_track_0711.xra
... .. [ IGNORE ] because a previous segmentation is existing.
```

Example of “Warning” message:

```
... .. [ WARNING ] chort- is missing of the dictionary and was
... .. automatically phonetized as S-O/-R-t
```

At the end of the report, the “Result statistics” section mentions the number of files that were annotated for each annotation, or -1 if the annotation was disabled.

### 3.6 Orthographic Transcription

An orthographic transcription is often the minimum requirement for a speech corpus so it is at the top of the annotation procedure, and it is the entry point for most of the automatic annotations. A *transcription convention* is designed to provide rules for writing speech corpora. This convention establishes what are the phenomena to transcribe and also how to mention them in the orthography.

From the beginning of its development it was considered to be essential for SPPAS to deal with an **Enriched Orthographic Transcription** (EOT). The transcription convention is summarized below and all details are given in the file `TOE-SPPAS.pdf`, available in the `documentation` folder. It indicates the rules and includes examples of what is expected or recommended.

Convention overview:

- truncated words, noted as a ‘-’ at the end of the token string (an ex-ample);
- noises, noted by a ‘\*’ (not available for some languages);
- laughter, noted by a ‘@’ (not available for some languages);
- short pauses, noted by a ‘+’;
- elisions, mentioned in parenthesis;
- specific pronunciations, noted with brackets [example,eczap];
- comments are preferably noted inside braces {this is a comment!};
- comments can be noted inside brackets without using comma [this and this];
- liaisons, noted between ‘=’ (this =n= example);
- morphological variants with <ice scream,I scream>;
- proper name annotation, like \$ John S. Doe \$.

The symbols \* + @ must be surrounded by whitespace.

SPPAS allows to include the regular punctuations. For some languages, it also allows to include numbers: they will be automatically converted to their written form during Text Normalization process.

From this EOT, several derived transcriptions can be generated automatically, including the two followings:

1. the standard transcription is the list of orthographic tokens (optional);
2. a specific transcription from which the phonetic tokens are obtained to be used by the grapheme-phoneme converter that is named faked transcription (the default).

For example, with the transcribed sentence: *This [is,iz] + hum... an enrich(ed) transcription {loud} number 1!*, the derived transcriptions are:

1. standard: *this is + hum an enriched transcription number one*
2. tokens: *this iz + hum an enrich transcription number one*

Notice that the convention allows to include a large scale of phenomena, for which most of them are optional. As a minimum, the **transcription must include**:

- filled pauses;
- short pauses;
- repeats;
- noises and laugh items (not available for Japanese and Cantonese).

Finally, it has to be noticed that this convention is not software-dependent. The orthographic transcription can be manually performed within SPPAS GUI in the “Edit” page, with Praat, with Annotation Pro, Audacity, ...

## 3.7 Search for Inter-Pausal Units (IPUs)

### 3.7.1 Overview

The Search for IPUs is a semi-automatic annotation process. This segmentation provides an annotated file with one tier named “IPUs”. The silence intervals are labelled with the “#” symbol, and IPUs intervals are labelled with “ipu\_” followed by the IPU number. This annotation is semi-automatic: **it should be verified manually**.

Notice that the better recording quality, the better IPUs segmentation.

### 3.7.2 The parameters

The following parameters must be properly fixed:

- Minimum volume value (in seconds): If this value is set to zero, the minimum volume is automatically adjusted for each sound file. Try with it first, then if the automatic value is not correct, set it manually. The Procedure Outcome Report indicates the value the system choose. The AudioRoamer component can also be of great help: it indicates min, max and mean volume values of the sound.
- Minimum silence duration (in seconds): By default, this is fixed to 0.2 sec. This duration mostly depends on the language. It is commonly fixed to at least 0.2 sec for French and at least 0.25 seconds for English language.

- Minimum speech duration (in seconds): By default, this value is fixed to 0.3 sec. A relevant value depends on the speech style: for isolated sentences, probably 0.5 sec should be better, but it should be about 0.1 sec for spontaneous speech.
- IPU boundary shift (in seconds) for start or end: a duration which is systematically added to IPU boundaries, to enlarge the IPU interval, and as a consequence, the neighboring silences are reduced.

The procedure outcome report indicates the values (volume, minimum durations) that were used by the system for each sound file.

### 3.7.3 Perform “Search for IPUs” with the GUI

It is an annotation of STANDALONE type.

Click on the “Search IPUs” activation button and on the “Configure...” blue text to fix options.

Example of result

Notice that the speech segments can be transcribed using SPPAS, in the “Analyze” page.

Orthographic transcription based on IPUs

### 3.7.4 Perform “Search for IPUs” with the CLI

`searchipus.py` is the program to perform this semi-automatic annotation, i.e. silence/IPUs segmentation, either on a single file (-i and optionally -o) or on a set of files (by using -I and optionally -e).

#### Usage

```
searchipus.py [files] [options]
```

Search for IPUs: Search for Inter-Pausal Units in an audio file.

optional arguments:

```
-h, --help          show this help message and exit
--quiet             Disable the verbosity
--log file          File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file             Input wav file name.
-o file             Annotated file with silences/units segmentation
                    (default: None)
```

Files (auto mode):

```
-I file             Input wav file name (append).
-e .ext             Output file extension. One of: .xra .TextGrid .eaf
                    .csv .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx
                    .arff .xrff
```

Options:

```
--outputpattern OUTPUTPATTERN
                    Output file pattern (default: )
--win_length WIN_LENGTH
```

---

```

                                Window size to estimate rms (in seconds) (default:
                                0.020)
--threshold THRESHOLD          Threshold of the volume value (rms) for the detection
                                of silences, 0=automatic (default: 0)
--min_ipu MIN_IPU              Minimum duration of an IPU (in seconds) (default:
                                0.300)
--min_sil MIN_SIL              Minimum duration of a silence (in seconds) (default:
                                0.200)
--shift_start SHIFT_START      Systematically move at left the boundary of the
                                beginning of an IPU (in seconds) (default: 0.01)
--shift_end SHIFT_END          Systematically move at right the boundary of the end
                                of an IPU (in seconds) (default: 0.02)

```

This program is part of SPPAS version 2.4. Copyright (C) 2011-2019 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

## Examples of use

A single input file and output on stdout:

```

python .\sppas\bin\searchipus.py -i .\samples\samples-eng\oriana1.wav
2018-12-19 10:49:32,782 [INFO] Logging set up level=15
2018-12-19 10:49:32,790 [INFO] ... Information:
2018-12-19 10:49:32,792 [INFO] ... .. Number of IPUs found:          3
2018-12-19 10:49:32,792 [INFO] ... .. Threshold volume value:        0
2018-12-19 10:49:32,792 [INFO] ... .. Threshold silence duration: 0.200
2018-12-19 10:49:32,792 [INFO] ... .. Threshold speech duration: 0.300
0.000000 1.675000 #
1.675000 4.580000 ipu_1
4.580000 6.390000 #
6.390000 9.880000 ipu_2
9.880000 11.430000 #
11.430000 14.740000 ipu_3
14.740000 17.792000 #

```

Idem without logs:

```

python .\sppas\bin\searchipus.py -i .\samples\samples-eng\oriana1.wav --quiet
0.000000 1.675000 #
1.675000 4.580000 ipu_1
4.580000 6.390000 #
6.390000 9.880000 ipu_2
9.880000 11.430000 #
11.430000 14.740000 ipu_3
14.740000 17.792000 #

```

Several input files, output in Praat-TextGrid file format:

```
python .\sppas\bin\searchipus.py -I .\samples\samples-eng\oriana1.wav \
-I .\samples\samples-eng\oriana3.wave -e .TextGrid
2018-12-19 10:48:16,520 [INFO] Logging set up level=15
2018-12-19 10:48:16,522 [INFO] File oriana1.wav: Valid.
2018-12-19 10:48:16,532 [INFO] ... Information:
2018-12-19 10:48:16,532 [INFO] ... .. Number of IPUs found:          3
2018-12-19 10:48:16,532 [INFO] ... .. Threshold volume value:         0
2018-12-19 10:48:16,532 [INFO] ... .. Threshold silence duration: 0.200
2018-12-19 10:48:16,533 [INFO] ... .. Threshold speech duration: 0.300
2018-12-19 10:48:16,538 [INFO] ... E:\bigi\Projets\sppas\samples\samples-eng\oriana1.Text
2018-12-19 10:48:16,538 [INFO] File oriana3.wave: Invalid.
2018-12-19 10:48:16,539 [ERROR] ... .. An audio file with only one channel is expected.
2018-12-19 10:48:16,540 [INFO] ... No file was created.
```

## 3.8 Fill in Inter-Pausal Units (IPUs)

### 3.8.1 Overview

This automatic annotation consists in aligning macro-units of a document with the corresponding sound.

IPUs are blocks of speech bounded by silent pauses of more than X ms. This annotation searches for a silences/IPUs segmentation of a recorded file (see previous section) and fill in the IPUs with the transcription given in a `txt` file.

### 3.8.2 How does it work

SPPAS identifies silent pauses in the signal and attempts to align them with the units proposed in the transcription file, under the assumption that each such unit is separated by a silent pause. It is based on the search of silences described in the previous section, but in this case, the number of units to find is known. The system adjusts automatically the volume threshold and the minimum durations of silences/IPUs to get the right number of units. The content of the units has no regard, because SPPAS does not interpret them: it can be the orthographic transcription, a translation, numbers, ... This algorithm is language-independent: it can work on any language.

In the transcription file, **silent pauses must be indicated** using both solutions, which can be combined:

- with the symbol '#';
- with newlines.

A recorded speech file must strictly correspond to a `txt` file of the transcription. The annotation provides an annotated file with one tier named "Transcription". The silence intervals are labelled with the "#" symbol, as IPUs are labelled with "ipu\_" followed by the IPU number then the corresponding transcription.

The same parameters than those indicated in the previous section must be fixed.

Remark: This annotation was tested on read speech no longer than a few sentences (about 1 minute speech) and on recordings of very good quality.

Fill in IPUs

### 3.8.3 Perform “Fill in IPU” with the GUI

It is an annotation of STANDALONE type.

Click on the “Fill in IPU” activation button and on the “Configure...” blue text to fix options.

### 3.8.4 Perform “Fill in IPU” with the CLI

`fillipus.py` is the program to perform this IPU segmentation, i.e. silence/ipus segmentation, either on a single file (-i and optionnally -o) or on a set of files (by using -I and optionnally -e).

#### Usage

```
fillipus.py [files] [options]
```

Fill in IPU: Search for Inter-Pausal Units and fill in with a transcription. Requires an audio file and a .txt file with the transcription.

optional arguments:

```
-h, --help          show this help message and exit
--quiet            Disable the verbosity
--log file        File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file            Input wav file name.
-t file            Input transcription file name.
-o file            Annotated file with filled IPU
```

Files (auto mode):

```
-I file            Input wav file name (append).
-e .ext           Output file extension. One of: .xra .TextGrid .eaf .csv
                  .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx .arff .xrff
```

Options:

```
--outputpattern OUTPUTPATTERN
                  Output file pattern (default: )
--min_ipu MIN_IPU Initial minimum duration of an IPU (in seconds) (default:
                  0.300)
--min_sil MIN_SIL Initial minimum duration of a silence (in seconds)
                  (default: 0.200)
```

This program is part of SPPAS version 3.0. Copyright (C) 2011-2020 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

#### Examples of use

A single input file with an input in manual mode:

```
python .\sppas\bin\fillipus.py -i .\samples\samples-eng\oriana1.wav -t .\samples\samples-en-
2018-12-19 11:03:15,614 [INFO] Logging set up level=15
2018-12-19 11:03:15,628 [INFO] ... Information:
```

```

2018-12-19 11:03:15,628 [INFO] ... .. Threshold volume value:      122
2018-12-19 11:03:15,630 [INFO] ... .. Threshold silence duration: 0.200
2018-12-19 11:03:15,630 [INFO] ... .. Threshold speech duration: 0.300
0.000000 1.675000 #
1.675000 4.570000 the flight was 12 hours long and we really got bored
4.570000 6.390000 #
6.390000 9.870000 they only played two movies + which we had both already seen
9.870000 11.430000 #
11.430000 14.730000 I never get to sleep on the airplane because it's so uncomfortable
14.730000 17.792000 #

```

A single input file in automatic mode:

```

python .\sppas\bin\fillipus.py -I .\samples\samples-eng\oriana1
python .\sppas\bin\fillipus.py -I .\samples\samples-eng\oriana1.wav
python .\sppas\bin\fillipus.py -I .\samples\samples-eng\oriana1.txt

```

## 3.9 Text normalization

### 3.9.1 Overview

In principle, any system that deals with unrestricted text need the text to be normalized. Texts contain a variety of “non-standard” token types such as digit sequences, words, acronyms and letter sequences in all capitals, mixed case words, abbreviations, roman numerals, URL’s and e-mail addresses... Normalizing or rewriting such texts using ordinary words is then an important issue. The main steps of the text normalization implemented in SPPAS (Bigi 2011) are:

- Replace symbols by their written form, thanks to a “replacement” dictionary, located into the folder “repl” in the “resources” directory.
- Word segmentation based on the content of a lexicon.
- Convert numbers to their written form.
- Remove punctuation.
- Lower the text.

### 3.9.2 Adapt Text normalization

Word segmentation of SPPAS is mainly based on the use of a lexicon. If a segmentation is not as expected, it is up to the user to modify the lexicon: Lexicons of all supported languages are all located in the folder “vocab” of the “resources” directory. They are in the form of “one word at a line” with [UTF-8 encoding](#) and “LF” for newline.

### 3.9.3 Support of a new language

Adding a new language in Text Normalization consists in the following steps:

1. Create a lexicon. Fix properly its encoding (utf-8), its newlines (LF), and fix the name and extension of the file as follow:

- language name with iso639-3 standard
  - extension “.vocab”
2. Put this lexicon in the `resources/vocab` folder
  3. Create a replacement dictionary for that language (take a look on the ones of the other language!)
  4. Optionally, the language can be added into the `num2letter.py` program

That’s it for most of the languages! If the language requires more steps, simply write to the author to collaborate, find some funding, etc. like it was already done for Cantonese (Bigi & Fung 2015) for example.

### 3.9.4 Perform Text Normalization with the GUI

It is an annotation of STANDALONE type.

The SPPAS Text normalization system takes as input a file (or a list of files) for which the name strictly match the name of the audio file except the extension. For example, if a file with name “oriana1.wav” is given, SPPAS will search for a file with name “oriana1.xra” at a first stage if “.xra” is set as the default extension, then it will search for other supported extensions until a file is found.

This file must include a tier with an orthographic transcription. At a first stage, SPPAS tries to find a tier with `transcription` as name. If such a tier does not exist, the first tier that is matching one of the following strings is used (case-insensitive search):

1. `trans`
2. `trs`
3. `ipu`
4. `ortho`
5. `toe`

Text normalization produces a file with “-token” appended to its name, i.e. “oriana1-token.xra” for the previous example. By default, this file is including only one tier with the resulting normalization and with name “Tokens”. To get other versions of the normalized transcription, click on the “Configure” text then check the expected tiers.

Read the “Introduction” of this chapter for a better understanding of the difference between “standard” and “faked” results.

To perform the text normalization process, click on the Text Normalization activation button, select the language and click on the “Configure...” blue text to fix options.

### 3.9.5 Perform Text Normalization with the CLI

`normalize.py` is the program to perform Text Normalization, i.e. the text normalization of a given file or a raw text.

#### Usage

```
normalize.py [files] [options]
```

Text Normalization: Text normalization segments the orthographic transcription

into tokens and remove punctuation, convert numbers, etc. Requires an orthographic transcription into IPU.

optional arguments:

```
-h, --help          show this help message and exit
--quiet            Disable the verbosity
--log file         File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file            Input transcription file name.
-o file            Annotated file with normalized tokens.
```

Files (auto mode):

```
-I file           Input transcription file name (append).
-l lang           Language code (iso8859-3). One of: cat cmn deu eng fra hun
                  ita jpn kor nan pcm pol por spa vie yue yue_chars.
-e .ext           Output file extension. One of: .xra .TextGrid .eaf .csv
                  .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx .arff .xrff
```

Resources:

```
-r vocab          Vocabulary file name
```

Options:

```
--inputpattern INPUTPATTERN
                  Input file pattern (orthographic transcription)
                  (default: )
--outputpattern OUTPUTPATTERN
                  Output file pattern (default: -token)
--faked FAKED     Create a tier with the faked tokens (required for
                  phonetization) (default: True)
--std STD         Create a tier with the standard tokens (useful if EOT)
                  (default: False)
--custom CUSTOM   Create a customized tier (default: False)
--occ_dur OCC_DUR Create tiers with number of tokens and duration of
                  each IPU (default: True)
```

This program is part of SPPAS version 2.4. Copyright (C) 2011–2019 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

## Examples of use

A single input file with a raw transcription input in manual mode:

```
python .\sppas\bin\normalize.py -r .\resources\vocab\eng.vocab -i .\samples\samples-eng\orianan
2018-12-19 11:48:34,151 [INFO] Logging set up level=15
2018-12-19 11:48:34,473 [INFO] ... ... Intervalle numéro 1.
2018-12-19 11:48:34,477 [INFO] ... ... Intervalle numéro 2.
2018-12-19 11:48:34,480 [INFO] ... ... Intervalle numéro 3.
Tokens
1, the flight was twelve hours long and we really got bored
2, they only played two movies + which we had both already seen
3, i never get to sleep on the airplane because it's so uncomfortable
```

A single input file with a transcription time-aligned into the IPUS, in manual mode and no logs:

```
python .\sppas\bin\normalize.py -r .\resources\vocab\eng.vocab
-i .\samples\samples-eng\oriana1.xra --quiet
Tokens
0.000000, 1.675000 #
1.675000, 4.570000 the flight was twelve hours long and we really got bored
4.570000, 6.390000 #
6.390000, 9.870000 they only played two movies + which we had both already seen
9.870000, 11.430000 #
11.430000, 14.730000 i never get to sleep on the airplane because it's so uncomfortable
14.730000, 17.792000 #
```

The same file in automatic mode can be annotated with one of the following commands:

```
python .\sppas\bin\normalize.py -I .\samples\samples-eng\oriana1.xra -l eng
python .\sppas\bin\normalize.py -I .\samples\samples-eng\oriana1.txt -l eng
python .\sppas\bin\normalize.py -I .\samples\samples-eng\oriana1.wav -l eng
python .\sppas\bin\normalize.py -I .\samples\samples-eng\oriana1 -l eng
```

This program can also normalize data from the standard input. Example of use, using stdin/stdout under Windows:

```
Write-Output "The flight was 12 HOURS {toto} long." |
python .\sppas\bin\normalize.py -r .\resources\vocab\eng.vocab --quiet
the
flight
was
twelve
hours
long
```

In that case, the comment mentioned with the braces is removed and the number is converted to its written form. The character “\_” is used for compound words (it replaces the whitespace).

## 3.10 Phonetization

### 3.10.1 Overview

Phonetization, also called grapheme-phoneme conversion, is the process of representing sounds with phonetic signs. However, converting from written text into actual sounds, for any language, cause several problems that have their origins in the relative lack of correspondence between the spelling of the lexical items and their sound contents. As a consequence, SPPAS implements a dictionary based-solution which consists in storing a maximum of phonological knowledge in a lexicon. This approach is then language-independent. SPPAS phonetization process is the equivalent of a sequence of dictionary look-ups.

Most of the other systems assume that all words of the speech transcription are mentioned in the pronunciation dictionary. On the contrary, SPPAS includes a language-independent algorithm which is able to phonetize unknown words of any language as long as a (minimum) dictionary is available (Bigi 2013). The Procedure Outcome Report reports on such cases with a WARNING message.

### 3.10.2 Adapt Phonetization

Since Phonetization is only based on the use of a pronunciation dictionary, the quality of the result only depends on this resource. If a pronunciation is not as expected, it is up to the user to change it in the dictionary: Dictionaries are located in the folder “dict” of the “resources” directory. They are all with **UTF-8 encoding** and **“LF” for newline**. The format of the dictionaries is HTK-like. As example, below is a piece of the `eng.dict` file:

```

THE           [THE]           D @
THE (2)      [THE]           D V
THE (3)      [THE]           D i :
THEA         [THEA]          T i : @
THEALL       [THEALL]        T i : l
THEANO       [THEANO]        T i : n @U
THEATER      [THEATER]       T i : @ 4 3:r
THEATER'S    [THEATER'S]      T i : @ 4 3:r z

```

The first column indicates the word, followed by the variant number (except for the first one). The second column indicates the word between brackets. The last columns are the succession of phones, separated by a whitespace. SPPAS is relatively compliant with the format and accept empty brackets or missing brackets.

The phoneset of the languages are mainly based on **X-SAMPA** international standard. See the chapter “Resources” of this documentation to know the list of accepted phones for a given language. This list can’t be extended nor modified by users. However, new phones can be added: Send an e-mail to the author to collaborate in that way.

Actually, some words can correspond to several entries in the dictionary with various pronunciations. These pronunciation variants are stored in the phonetization result. By convention, whitespace separate words, minus characters separate phones and pipe character separate phonetic variants of a word. For example, the transcription utterance:

- **Transcription:** The flight was 12 hours long.
- **Text Normalization:** the flight was twelve hours long
- **Phonetization:** D-@|D-V|D-i: f-l-aI-t w-A-z|w-V-z|w-@-z|w-O:-z t-w-E-l-v aU-3:r-z|aU-r-z l-O:-N

### 3.10.3 Support of a new language

The support of a new language in Phonetization only consists in: 1. creating the pronunciation dictionary. The following constraints on the file must be respected: - its format (HTK-like), - its encoding (UTF-8), - its newlines (LF), - its phone set (X-SAMPA), - its file name (iso639-3 of the language and “.dict” extension). 2. adding the dictionary in the “dict” folder of the “resources” directory.

### 3.10.4 Perform Phonetization with the GUI

It is an annotation of STANDALONE type.

The Phonetization process takes as input a file that strictly match the audio file name except for the extension and that “-token” is appended. For example, if the audio file name is “oriana1.wav”, the expected input file name is “oriana1-token.xra” if .xra is the default extension for annotations. This file must include a **normalized** orthographic transcription. The name of such tier must contains one of the following strings:

1. “tok”
2. “trans”

The first tier that matches one of these requirements is used (this match is case-insensitive).

Phonetization produces a file with “-phon” appended to its name, i.e. “oriana1-phon.xra” for the previous example. This file contains only one tier with the resulting phonetization and with name “Phones”.

To perform the annotation, click on the Phonetization activation button, select the language and click on the “Configure...” blue text to fix options.

### 3.10.5 Perform Phonetization with the CLI

phonetize.py is the program to perform Phonetization on a given file, i.e. the grapheme-conversion of a file or a raw text.

#### Usage

```
phonetize.py [files] [options]
```

Phonetization: Grapheme to phoneme conversion represents sounds with phonetic signs. Requires a Text Normalization.

optional arguments:

```
-h, --help          show this help message and exit
--quiet             Disable the verbosity
--log file          File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file             Input tokenization file name.
-o file             Annotated file with phonetization.
```

Files (auto mode):

```
-I file             Input transcription file name (append).
-l lang             Language code (iso8859-3). One of: cat cmn deu eng fra
                   ita jpn kor nan pcm pol por spa yue yue_chars.
-e .ext             Output file extension. One of: .xra .TextGrid .eaf
                   .csv .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx
                   .arff .xrff
```

Resources:

```
-r dict             Pronunciation dictionary (HTK-ASCII format).
-m map_file         Pronunciation mapping table. It is used to generate
                   new pronunciations by mapping phonemes of the
                   dictionary.
```

Options:

```
--inputpattern INPUTPATTERN
                   Input file pattern (tokenization) (default: -token)
--outputpattern OUTPUTPATTERN
                   Output file pattern (default: -phon)
--unk UNK           Try to phonetize unknown words (default: True)
--usestdtokens USESTDTOKENS
                   Phonetize from standard spelling (default: False)
```

This program is part of SPPAS version 2.4. Copyright (C) 2011-2019 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

## Examples of use

A single input file with a normalized text in manual mode:

```
python .\sppas\bin\phonetize.py -r .\resources\dict\eng.dict
-i .\samples\samples-eng\oriana1-token.xra --quiet
Phones
0.000000, 1.675000, sil
1.675000, 4.570000, {D-@|D-i|D-V} f-l-aI-t {w-@-z|w-V-z|w-O:-z|w-A-z} t-w-E-l-v
{aU-3:r-z|aU-r\~z} l-O:-N {{-n-d|@-n-d} w-i: {r\~I-l-i:|r\~i:-l-i:} g-A-t b-O:-r\~d
4.570000, 6.390000, sil
6.390000, 9.870000, D-eI @U-n-l-i: p-l-eI-d t-u m-u-v-i:-z sil {h-w-I-tS|w-I-tS}
w-i: h-{-d b-@U-T {O:-l-r\~E-4-i:|O:-r\~E-4-i:} s-i:-n
9.870000, 11.430000, sil
11.430000, 14.730000, aI n-E-v-3:r {g-I-t|g-E-t} {t-@|t-i|t-u} s-l-i:-p
{O:-n|A-n} {D-@|D-i|D-V} E-r\~p-l-eI-n {b-i-k-O:-z|b-i-k-V-z} {i-t-s|I-t-s}
s-@U @-n-k-V-m-f-3:r-4-@-b-@-l
14.730000, 17.792000, sil
```

The same file in automatic mode can be annotated with one of the following commands:

```
python .\sppas\bin\phonetize.py -l eng -I .\samples\samples-eng\oriana1-token.xra
python .\sppas\bin\phonetize.py -l eng -I .\samples\samples-eng\oriana1.xra
python .\sppas\bin\phonetize.py -l eng -I .\samples\samples-eng\oriana1.txt
python .\sppas\bin\phonetize.py -l eng -I .\samples\samples-eng\oriana1.wav
python .\sppas\bin\phonetize.py -l eng -I .\samples\samples-eng\oriana1
```

This program can also phonetize data from the standard input. Example of use, using stdin/stdout under Windows:

```
Write-Output "The flight was 12 HOURS {toto} long." |
python .\sppas\bin\normalize.py -r .\resources\vocab\eng.vocab --quiet |
python .\sppas\bin\phonetize.py -r .\resources\dict\eng.dict --quiet
D-@|D-V|D-i:
f-l-aI-t
w-A-z|w-V-z|w-@-z|w-O:-z
t-w-E-l-v
aU-3:r-z|aU-r\~z
l-O:-N
```

## 3.11 Alignment

### 3.11.1 Overview

Alignment, also called phonetic segmentation, is the process of aligning speech with its corresponding transcription at the phone level. The alignment problem consists in a time-matching between a given speech unit along with a phonetic representation of the unit.

**SPPAS Alignment does not perform the segmentation itself. It is a wrapper either for the Julius Speech Recognition Engine (SRE) or for the `fvite` command of HTK-Toolkit.** In addition, SPPAS can perform a “basic” alignment, assigning the same duration to each sound.

Speech Alignment requires an Acoustic Model in order to align speech. An acoustic model is a file that contains statistical representations of each of the distinct sounds of one language. Each sound is represented by one of these statistical representations. The quality of the alignment result only depends on both this resource and on the aligner. From our past experiences, we got better results with Julius. See the chapter 4 “Resources for Automatic Annotations” to get the list of sounds of each language.

Notice that SPPAS allows to time-align automatically laugh, noises, or filled pauses (depending on the language): No other system is able to achieves this task!

SPPAS alignment output example

### 3.11.2 Adapt Alignment

The better Acoustic Model, the better alignment results. Any user can append or replace the acoustic models included in the “models” folder of the “resources” directory. Be aware that SPPAS only supports HTK-ASCII acoustic models, trained from 16 bits, 16000 Hz wave files.

The existing models can be improved if they are re-trained with more data. To get a better alignment result, any new data is then welcome: send an e-mail to the author to share your recordings and transcripts.

### 3.11.3 Support of a new language

The support of a new language in Alignment only consists in adding a new acoustic model of the appropriate format, in the appropriate directory, with the appropriate phone set.

The articulatory representations of phonemes are so similar across languages that phonemes can be considered as units which are independent from the underlying language (Schultz et al. 2001). In SPPAS package, 9 acoustic models of the same type - i.e. same HMMs definition and acoustic parameters, are already available so that the phoneme prototypes can be extracted and reused to create an initial model for a new language.

Any new model can also be trained by the author, as soon as enough data is available. It is difficult to estimate exactly the amount of data a given language requires. That is said, we can approximate the minimum as follow:

- 3 minutes altogether of various speakers, manually time-aligned at the phoneme level.
- 10 minutes altogether of various speakers, time-aligned at the ipus level with the enriched orthographic transcription.
- more data is good data.

### 3.11.4 Perform Alignment with the GUI

It is an annotation of STANDALONE type.

The Alignment process takes as input one or two files that strictly match the audio file name except for the extension and that “-phon” is appended for the first one and “-token” for the optional second one. For example, if the audio file name is “oriana1.wav”, the expected input file name is “oriana1-phon.xra” with phonetization and optionally “oriana1-token.xra” with text normalization, if .xra is the default extension for annotations.

The speech segmentation process provides one file with name “-palign” appended to its name, i.e. “oriana1-palign.xra” for the previous example. This file includes one or two tiers:

- “PhonAlign” is the segmentation at the phone level;
- “TokensAlign” is the segmentation at the word level (if a file with tokenization was found).

The following options are available to configure Alignment:

- choose the speech segmentation system. It can be either: julius, hvite or basic
- perform basic alignment if the aligner failed, instead such intervals are empty.
- remove working directory will keep only alignment result: it will remove working files. Working directory includes one wav file per unit and a set of text files per unit.
- create the PhnTokAlign will append another tier with intervals of the phonetization of each word.

To perform the annotation, click on the Alignment activation button, select the language and click on the “Configure...” blue text to fix options.

### 3.11.5 Perform Alignment with the CLI

`alignment.py` is the program to perform automatic speech segmentation of a given phonetized file.

#### Usage

```
alignment.py [files] [options]
```

Alignment: Time-alignment of speech audio with its corresponding transcription at the phone and token levels. Requires a Phonetization.

optional arguments:

```
-h, --help          show this help message and exit
--quiet            Disable the verbosity
--log file         File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file            Input wav file name.
-p file            Input file name with the phonetization.
-t file            Input file name with the tokenization.
-o file            Output file name with estimated alignments.
```

Files (auto mode):

```
-I file            Input transcription file name (append).
-l lang           Language code (iso8859-3). One of: cat cmn deu eng
```

```

-e .ext                eng-cd fra ita jpn kor nan pcm pol por spa yue.
                       Output file extension. One of: .xra .TextGrid .eaf
                       .csv .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx
                       .arff .xrff

Resources:
-r model              Directory of the acoustic model of the language of the
                       text
-R model              Directory of the acoustic model of the mother language
                       of the speaker (under development)

Options:
--inputpattern INPUTPATTERN
                       Input file pattern (phonetization) (default: -phon)
--inputoptpattern INPUTOPTPATTERN
                       Optional input file pattern (tokenization) (default:
                       -token)
--outputpattern OUTPUTPATTERN
                       Output file pattern (default: -palign)
--aligner ALIGNER    Speech automatic aligner system (julius, hvite,
                       basic): (default: julius)
--basic BASIC        Perform basic alignment if the aligner fails (default:
                       False)
--clean CLEAN        Remove working directory (default: True)
--activity ACTIVITY  Create the Activity tier (default: True)
--activityduration ACTIVITYDURATION
                       Create the ActivityDuration tier (default: False)

```

This program is part of SPPAS version 2.4. Copyright (C) 2011-2019 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

### Example of use

```

python .\sppas\bin\alignment.py -I .\samples\samples-eng\oriana1.wav -l eng
2018-12-19 18:33:38,842 [INFO] Logging set up level=15
2018-12-19 18:33:38,844 [INFO] Options
2018-12-19 18:33:38,844 [INFO] ... activityduration: False
2018-12-19 18:33:38,845 [INFO] ... activity: True
2018-12-19 18:33:38,845 [INFO] ... aligner: julius
2018-12-19 18:33:38,845 [INFO] ... clean: True
2018-12-19 18:33:38,845 [INFO] ... basic: False
2018-12-19 18:33:38,845 [INFO] File oriana1.wav: Valid.
2018-12-19 18:33:38,845 [INFO] File oriana1-phon.xra: Valid.
2018-12-19 18:33:38,846 [INFO] File oriana1-token.xra: Valid.
2018-12-19 18:33:38,846 [WARNING] ... .. A file with name E:\bigi\Projets\sppas\sampl
2018-12-19 18:33:38,855 [INFO] ... Découpage en intervalles.
2018-12-19 18:33:38,901 [INFO] ... Intervalle numéro 1.
2018-12-19 18:33:38,904 [INFO] ... Intervalle numéro 2.
2018-12-19 18:33:38,908 [INFO] ... Intervalle numéro 3.
2018-12-19 18:33:38,913 [INFO] ... Intervalle numéro 4.
2018-12-19 18:33:38,917 [INFO] ... Intervalle numéro 5.
2018-12-19 18:33:38,921 [INFO] ... Intervalle numéro 6.
2018-12-19 18:33:38,926 [INFO] ... Intervalle numéro 7.
2018-12-19 18:33:38,928 [INFO] ... Fusion des alignements des intervalles.

```

```
2018-12-19 18:33:38,969 [INFO] ... Création de la tier des activités.  
2018-12-19 18:33:38,993 [INFO] ... E:\bigi\Projets\sppas\samples\samples-eng\oriana1-pal1
```

## 3.12 Activity

### 3.12.1 Overview

Activity tier represents speech activities, i.e. speech, silences, laughter, noises... It is based on the analysis of the time-aligned tokens.

### 3.12.2 Perform Activity with the GUI

It is an annotation of STANDALONE type.

The Activity process takes as input a file that strictly match the audio file name except for the extension and that “-palign” is appended. For example, if the audio file name is “oriana1.wav”, the expected input file name is “oriana1-palign.xra” if .xra is the default extension for annotations. This file must include time-aligned phonemes in a tier with name “PhonAlign”.

The annotation provides an annotated file with “-activity” appended to its name, i.e. “oriana1-activity.xra” for the previous example. This file is including 1 or 2 tiers: Activity, ActivityDuration.

To perform the annotation, click on the Activity activation button and click on the “Configure...” blue text to fix options.

### 3.12.3 Perform Alignment with the CLI

No CLI is available for this annotation.

## 3.13 RMS

### 3.13.1 Overview

The Root-Mean Square - RMS is a measure of the power in an audio signal. It is estimated from the amplitude values by:  $\sqrt{\sum(S_i^2)/n}$ .

RMS automatic annotation estimates the rms value on given intervals of an audio file. Empty intervals - i.e. intervals without labels, are ignored. By default, the RMS is estimated on a tier with name “PhonAlign” of an annotated file with pattern “-palign”. Both can be modified by configuring the annotations. The annotation provides an annotated file with “-rms” appended to its name. This file is including 3 tiers:

- RMS: indicates the RMS value estimated on each non-empty interval;
- RMS-values: indicates RMS values estimated every 10 ms in each interval;
- RMS-mean: indicates the mean of the previous values.

### 3.13.2 Perform RMS with the GUI

It is an annotation of STANDALONE type.

To perform the annotation, click on the RMS activation button and click on the “Configure...” blue text to fix options.

### 3.13.3 Perform RMS with the CLI

`rms.py` is the program to perform this annotation, either on a single given file (-i and -t) or on a set of files (-I).

## 3.14 Interval Values Analysis - IVA

### 3.14.1 Overview

The Interval Values Analysis - IVA is producing statistical information about a set of values in given intervals. IVA can for example estimate the mean/stddev values on given intervals (IPUs, ...) of a pitch file. Empty intervals - i.e. unlabelled intervals, are ignored and a list of tags to be ignored can be fixed.

By default, the IVA is estimated with the values of a PitchTier inside the intervals defined in a tier with name “TokensAlign” of a file with pattern “-palign”. If a list of separators is given, the intervals are created: an IVA segment is a set of consecutive annotations without separators. Default separators are “# + @ \* dummy” in order to ignore silences, laughter items, noises and untranscribed speech. However, if no separator is given, IVA segments are matching the intervals of the given input tier. In the latter case, be aware that some file formats - including TextGrid, are not supporting holes: they create unlabelled intervals between the labelled ones.

Both tiernames and patterns can be modified by configuring the annotation. The annotation provides an annotated file with the “-iva” pattern. This file includes the tiers:

- IVA-Segments: defined intervals;
- IVA-Values: the values extracted for each segment;
- IVA-Occurrences: indicates the number of values of each segment;
- IVA-Total: indicates the sum of values of each segment;
- IVA-Mean: indicates the mean of values of each segment;
- IVA-Median: indicates the median of values of each segment;
- IVA-StdDev: indicates the standard deviation of values of each segment;
- IVA-Intercept: indicates the intercept value of the linear regression of values of each segment;
- IVA-Slope: indicates the slope value of the linear regression of values of each segment.

### 3.14.2 Perform IVA with the GUI

It is an annotation of STANDALONE type.

To perform the annotation, click on the IVA activation button and click on the “Configure...” blue text to fix options.

### 3.14.3 Perform IVA with the CLI

`iva.py` is the program to perform this annotation, either on a single given file (`-i` and `-s`) or on a set of files (`-I`).

## 3.15 Lexical Metric

### 3.15.1 Overview

The Lexical Metric is producing information about the number of occurrences and the rank of each e occurrences of annotation labels.

By default, the lexical metrics are estimated on a tier with name “TokensAlign” of a file with pattern “-palign”. If a list of separators is given, segments are created to estimate a number of occurrences. Default separators are “# + @ \* dummy” in order to ignore silences, laughter items, noises and untranscribed speech.

Both the tiername and the pattern can be modified by configuring the annotation. The annotation provides an annotated file with the “-lexm” pattern. This file includes the tiers:

- LM-OccAnnInSegments: defined intervals with number of occurrences of annotations;
- LM-OccLabInSegments: defined intervals with number of occurrences of labels;
- LM-Occ: the number of occurrences of each label the annotation is representing;
- LM-Rank: the rank of each label the annotation is representing.

### 3.15.2 Perform Lexical Metric with the GUI

It is an annotation of STANDALONE type.

To perform the annotation, click on the Lexical Metric activation button and click on the “Configure...” blue text to fix options.

## 3.16 Syllabification

### 3.16.1 Overview

The syllabification of phonemes is performed with a rule-based system from time-aligned phonemes. This phoneme-to-syllable segmentation system is based on 2 main principles:

- a syllable contains a vowel, and only one;
- a pause is a syllable boundary.

These two principles focus the problem of the task of finding a syllabic boundary between two vowels. Phonemes were grouped into classes and rules are established to deal with these classes.

Syllabification example

For each language, the automatic syllabification requires a configuration file to fix phonemes, classes and rules.

### 3.16.2 Adapt Syllabification

Any user can change the set of rules by editing and modifying the configuration file of a given language. Such files are located in the folder “syll” of the “resources” directory. Files are all with [UTF-8 encoding](#) and [“LF” for newline](#).

At first, the list of phonemes and the class symbol associated with each of the phonemes are described as, for example:

- PHONCLASS e V
- PHONCLASS p P

Each association phoneme/class definition is made of 3 columns: the first one is the key-word PHONCLASS, the second is the phoneme symbol (like defined in the tier with the phonemes, commonly X-SAMPA), the last column is the class symbol. The constraints on this definition are that a class-symbol is only one upper-case character, and that the character X is forbidden, and the characters V and W are reserved for vowels.

The second part of the configuration file contains the rules. The first column is a keyword, the second one describes the classes between two vowels and the third column is the boundary location. The first column can be:

- GENRULE
- EXCRULE
- OTHRULE.

In the third column, a “0” means the boundary is just after the first vowel, “1” means the boundary is one phoneme after the first vowel, etc. Here are some examples of the file for French language:

- GENRULE VXV 0
- GENRULE VXXV 1
- EXCRULE VFLV 0
- EXCRULE VOLGV 0

Finally, to adapt the rules to specific situations that the rules failed to model, we introduced some phoneme sequences and the boundary definition. Specific rules contain only phonemes or the symbol “ANY” which means any phoneme. It consists of 7 columns: the first one is the key-word OTHRULE, the 5 following columns are a phoneme sequence where the boundary should be applied to the third one by the rules, the last column is the shift to apply to this boundary. In the following example:

```
OTHRULE ANY ANY p s k -2
```

More information are available in (Bigi et al. 2010).

### 3.16.3 Support of a new language

The support of a new language in this automatic syllabification only consists in adding a configuration file (see previous section). Fix properly the encoding (utf-8) and newlines (LF) of this file; then fix the name and extension of the file as follow:

- “syllConfig-” followed by language name with iso639-3 standard,
- with extension “.txt”.

### 3.16.4 Perform Syllabification with the GUI

It is an annotation of STANDALONE type.

The Syllabification process takes as input a file that strictly match the audio file name except for the extension and that “-palign” is appended. For example, if the audio file name is “oriana1.wav”, the expected input file name is “oriana1-palign.xra” if .xra is the default extension for annotations. This file must include time-aligned phonemes in a tier with name “PhonAlign”.

The annotation provides an annotated file with “-salign” appended to its name, i.e. “oriana1-salign.xra” for the previous example. This file is including 2 tiers: SyllAlign, SyllClassAlign. Optionally, the program can add a tier with the syllable structures (V, CV, CCV...).

To perform the annotation, click on the Syllabification activation button, select the language and click on the “Configure...” blue text to fix options.

### 3.16.5 Perform Syllabification with the CLI

syllabify.py is the program to perform automatic syllabification of a given file with time-aligned phones.

#### Usage

```
syllabify.py [files] [options]
```

Syllabification: Syllabification is based on a set of rules to convert phonemes into classes and to group them. Requires time-aligned phones.

optional arguments:

```
-h, --help          show this help message and exit
--quiet            Disable the verbosity
--log file         File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file            Input time-aligned phonemes file name.
-o file            Output file name with syllables.
```

Files (auto mode):

```
-I file            Input transcription file name (append).
-l lang           Language code (iso8859-3). One of: fra ita pol.
-e .ext           Output file extension. One of: .xra .TextGrid .eaf
                  .csv .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx
                  .arff .xrff
```

Resources:

```
-r rules           Configuration file with syllabification rules
```

Options:

```
--inputpattern INPUTPATTERN
                  Input file pattern (time-aligned phonemes) (default:
                  -palign)
--outputpattern OUTPUTPATTERN
                  Output file pattern (default: -syll)
--usesphons USESPHONS
                  Syllabify inside the IPU intervals (default: True)
```

```

--usesintervals USESINTERVALS
                        Syllabify inside an interval tier (default: False)
--tiername TIERNAME    Tier name for such interval tier: (default:
                        TokensAlign)
--createclasses CREATECLASSES
                        Create a tier with syllable classes (default: True)

```

This program is part of SPPAS version 2.4. Copyright (C) 2011-2019 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

### Examples of use

```

python .\sppas\bin\syllabify.py -i .\samples\samples-fra\F_F_B003-P8-palign.xra
-r .\resources\syll\syllConfig-fra.txt --quiet
SyllAlign
2.497101 2.717101 j-E-R
2.717101 2.997101 s-w-A/-R
...
19.412000 19.692000 P-L-V-P
19.692000 20.010000 P-V-L-P

```

All the following commands will produce the same result:

```

python .\sppas\bin\syllabify.py -I .\samples\samples-fra\F_F_B003-P8-palign.xra -l fra
python .\sppas\bin\syllabify.py -I .\samples\samples-fra\F_F_B003-P8.TextGrid -l fra
python .\sppas\bin\syllabify.py -I .\samples\samples-fra\F_F_B003-P8.wav -l fra
python .\sppas\bin\syllabify.py -I .\samples\samples-fra\F_F_B003-P8 -l fra

```

## 3.17 TGA - Time Groups Analyzer

### 3.17.1 Overview

TGA is originally available at <http://wwwhomes.uni-bielefeld.de/gibbon/TGA/>. It's a tool developed by Dafydd Gibbon, emeritus professor of English and General Linguistics at Bielefeld University.

*Dafydd Gibbon (2013). TGA: a web tool for Time Group Analysis, Tools and Resources for the Analysis of Speech Prosody, Aix-en-Provence, France, pp. 66-69.*

The original TGA is an online batch processing tool which provides a parametrised mapping from time-stamps in speech annotation files in various formats to a detailed analysis report with statistics and visualisations. TGA software calculates, inter alia, mean, median, rPVI, nPVI, slope and intercept functions within inter-pausal groups, provides visualizations of timing patterns, as well as correlations between these, and parses inter-pausal groups into hierarchies based on duration relations. Linear regression is selected mainly for the slope function, as a first approximation to examining acceleration and deceleration over large data sets.

The TGA online tool was designed to support phoneticians in basic statistical analysis of annotated speech data. In practice, the tool provides not only rapid analyses but also the ability to handle larger data sets than can be handled manually.

In addition to the original one, a second version of TGA was implemented in the AnnotationPro software:

*Katarzyna Klessa, Dafydd Gibbon (2014). **Annotation Pro + TGA: automation of speech timing analysis**, 9th International conference on Language Resources and Evaluation (LREC), Reykjavik (Iceland). pp. 1499-1505, ISBN: 978-2-9517408-8-4.*

The integrated Annotation Pro + TGA tool incorporates some TGA features and is intended to support the development of more robust and versatile timing models for a greater variety of data. The integration of TGA statistical and visualisation functions into Annotation Pro+TGA results in a powerful computational enhancement of the existing AnnotationPro phonetic workbench, for supporting experimental analysis and modeling of speech timing.

So, what's the novelty into the third version implemented into SPPAS...

First of all, it has to be noticed that TGA is only partly implemented into SPPAS. The statistics analyses tool of SPPAS allows to estimate TGA within the SPPAS framework; and it results in the following advantages:

- it can read either TextGrid, csv, Elan, or any other file format supported by SPPAS,
- it can save TGA results in any of the annotation file supported by SPPAS,
- it estimates the two versions of the linear regression estimators: the original one and the one implemented into AnnotationPro:
  1. in the original TGA, the x-axis is based on positions of syllables,
  2. in the AnnotationPro+TGA, the x-axis is based on time-stamps.

### 3.17.2 Result of TGA into SPPAS

The annotation provides an annotated file with “-tga” appended to its name, i.e. “oriana1-tga.xra” for the example. This file is including 10 tiers:

1. TGA-TimeGroups: intervals with the time groups
2. TGA-TimeSegments: same intervals, indicate the syllables separated by whitespace
3. TGA-Occurrences: same intervals, indicate the number of syllables
4. TGA-Total: same intervals, indicate interval duration
5. TGA-Mean: same intervals, indicate mean duration of syllables
6. TGA-Median: same intervals, indicate median duration of syllables
7. TGA-Stdev: same intervals, indicate stdev of duration of syllables
8. TGA-nPVI: same intervals, indicate nPVI of syllables
9. TGA-Intercept: same intervals, indicate the intercept
10. TGA-Slope: same intervals, indicate the slope

Both tiers 9 and 10 can be estimated in 2 ways (so 2 more tiers can be generated).

### 3.17.3 Perform TAG with the GUI

It is an annotation of STANDALONE type.

The TGA process takes as input a file that strictly match the audio file name except for the extension and that “-salign” is appended. For example, if the audio file name is “oriana1.wav”, the expected input file name is “oriana1-salign.xra” if .xra is the default extension for annotations. This file must include time-aligned syllables in a tier with name “SyllAlign”.

To perform the annotation, click on the TGA activation button and click on the “Configure...” blue text to fix options.

### 3.17.4 Perform TGA with the CLI

tga.py is the program to perform TGA of a given file with time-aligned syllables.

#### Usage

```
tga.py [files] [options]
```

TimeGroupAnalysis: Proposed by D. Gibbon, Time Group Analyzer calculates mean, median, nPVI, slope and intercept functions within inter-pausal groups. Requires time aligned syllables.

optional arguments:

```
-h, --help          show this help message and exit
--quiet            Disable the verbosity
--log file         File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file            An input time-aligned syllables file.
-o file            Output file name with TGA.
```

Files (auto mode):

```
-I file            Input time-aligned syllables file (append).
-e .ext            Output file extension. One of: .xra .TextGrid .eaf
                  .csv .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx
                  .arff .xrff
```

Options:

```
--original ORIGINAL Use the original estimation of intercept and slope
                    (default: False)
--annotationpro ANNOTATIONPRO
                    Use the estimation of intercept and slope proposed in
                    AnnotationPro (default: True)
--tg_prefix_label TG_PREFIX_LABEL
                    Prefix of each time group label: (default: tg_)
--with_radius WITH_RADIUS
                    Duration estimation: Use 0 to estimate syllable
                    durations with midpoint values, use -1 for Radius-, or
                    1 for Radius+. (default: 0)
```

This program is part of SPPAS version 2.0. Copyright (C) 2011-2019 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

## Example of use

```
python .\sppas\bin\tga.py -i .\samples\samples-fra\F_F_B003-P8-syll.xra
2018-12-20 08:35:21,219 [INFO] Logging set up level=15
TGA-TimeGroups
2.497101 5.683888 tg_1
5.743603 8.460596 tg_2
9.145000 11.948531 tg_3
12.494000 13.704000 tg_4
13.784000 15.036000 tg_5
16.602000 20.010000 tg_6
TGA-TimeSegments
...
13.784000 15.036000 -0.03063
16.602000 20.010000 0.00468
```

## Other commands:

```
python .\sppas\bin\tga.py -I .\samples\samples-fra\F_F_B003-P8-syll.xra
python .\sppas\bin\tga.py -I .\samples\samples-fra\F_F_B003-P8.TextGrid
python .\sppas\bin\tga.py -I .\samples\samples-fra\F_F_B003-P8.wav
```

## 3.18 Stop words

Create a tier with True/False indicating if a token is a stop-word or not.

## 3.19 Self-Repetitions

### 3.19.1 Overview

This automatic detection focus on word self-repetitions which can be exact repetitions (named strict echos) or repetitions with variations (named non-strict echos). The system is based only on lexical criteria. The algorithm is focusing on the detection of the source.

This system can use a list of stop-words of a given language. This is a list of very frequent words like adjectives, pronouns, etc. Obviously, the result of the automatic detection is significantly better if such list of stop-words is available.

Optionnally, SPPAS can add new stop-words in the list: they are deduced from the given data. These new entries in the stop-list are then different for each file (Bigi et al. 2014).

The annotation provides one annotated file with 2 to 4 tiers:

1. TokenStrain: if a replacement file was available, it's the entry used by the system
2. StopWord: if a stop-list was used, it indicates if the token is a stop-word (True or False)
3. SR-Sources: tags of the annotations are prefixed by "S" followed an index
4. SR-Repetitions: tags of the annotations are prefixed by "R" followed an index

### 3.19.2 Adapt to a new language

The list of stop-words of a given language must be located in the “vocab” folder of the “resources” directory with “.stp” extension. This file is with [UTF-8 encoding](#) and “LF” for newline.

### 3.19.3 Perform Self-Repetitions with the GUI

It is an annotation of STANDALONE type.

The automatic annotation takes as input a file with (at least) one tier containing the time-aligned tokens. The annotation provides one annotated file with 2 tiers: Sources and Repetitions.

Click on the Self-Repetitions activation button, select the language and click on the “Configure...” blue text to fix options.

### 3.19.4 Perform SelfRepetitions with the CLI

`selfrepetition.py` is the program to perform automatic detection of self-repetitions.

#### Usage

```
selfrepetition.py [files] [options]
```

Self-repetitions: Self-repetitions searches for sources and echos of a speaker. Requires time-aligned tokens.

optional arguments:

```
-h, --help          show this help message and exit
--quiet            Disable the verbosity
--log file         File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file            Input time-aligned tokens file name.
-o file            Output file name with syllables.
```

Files (auto mode):

```
-I file            Input transcription file name (append).
-e .ext            Output file extension. One of: .xra .TextGrid .eaf
                  .csv .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx
                  .arff .xrff
```

Resources:

```
-r file            List of stop-words
```

Options:

```
--inputpattern INPUTPATTERN
                  Input file pattern (time-aligned words or lemmas)
                  (default: -palign)
--outputpattern OUTPUTPATTERN
                  Output file pattern (default: -srepet)
--span SPAN        Span window length in number of IPUs (default: 3)
--stopwords STOPWORDS
                  Add stop-words estimated from the given data (default:
```

```
True)
--alpha ALPHA      Coefficient to add data-specific stop-words (default:
                   0.5)
```

This program is part of SPPAS version 3.0. Copyright (C) 2011-2020 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

## Examples of use

```
python .\sppas\bin\selfrepetition.py -i .\samples\samples-fra\F_F_B003-P8-palign.xra
-r .\resources\vocab\fra.stp
python .\sppas\bin\selfrepetition.py -I .\samples\samples-fra\F_F_B003-P8.wav -l fra
```

## 3.20 Other-Repetitions

### 3.20.1 Overview

This automatic detection focus on other-repetitions, which can be either exact repetitions (named strict echos) or repetitions with variations (named non-strict echos). The system is based only on lexical criteria (Bigi et al. 2014). Notice that the algorithm is focusing on the detection of the source.

This system can use a list of stop-words of a given language. This is a list of very frequent words like adjectives, pronouns, etc. Obviously, the result of the automatic detection is significantly better if such list of stop-words is available.

Optionnally, SPPAS can add new stop-words in the list: they are deduced from the given data. These new entries in the stop-list are then different for each file (see Bigi et al. 2014).

The detection of the ORs is performed in a span window of N IPU; by default, N is fixed to 5. It means that if a repetition is after these N IPU, it won't be detected. Technically, it also means that SPPAS needs to identify the boundaries of the IPU from the time-aligned tokens: the tier must indicate the silences with the '#' symbol.

A file with the following tiers will be created:

- OR-Source: intervals with the number of the sources
- OR-SrcStrain: intervals with the tokens of the sources
- OR-SrcLen: intervals with the number of tokens in the source
- OR-SrcType: intervals with the type of the echos of the sources
- OR-Echo: intervals with the number of the echos

### 3.20.2 Adapt to a language and support of a new one

This system can use a list of stop-words of a given language. This is a list of very frequent words like adjectives, pronouns, etc. Obviously, the result of the automatic detection is significantly better if such list of stop-words is available. It must be located in the "vocab" folder of the "resources" directory with ".stp" extension. This file is with [UTF-8 encoding](#) and "LF" for newline.

### 3.20.3 Perform Other-Repetitions with the GUI

It is an annotation of INTERACTION type.

The automatic annotation takes as input a file with (at least) one tier containing the time-aligned tokens of the main speaker, and another file/tier with tokens of the interlocutor. The annotation provides one annotated file with 2 tiers: Sources and Repetitions.

Click on the Other-Repetitions activation button, select the language and click on the “Configure...” blue text to fix options.

### 3.20.4 Perform Other-Repetitions with the CLI

```
usage: otherrepetition.py -r stopwords [files] [options]
```

Files:

```
-i file           Input file name with time-aligned tokens of the main
                  speaker.
-s file           Input file name with time-aligned tokens of the
                  echoing speaker
-o file           Output file name with ORs.
```

Options:

```
--inputpattern INPUTPATTERN
                        Input file pattern (time-aligned words or lemmas)
                        (default: -palign)
--outputpattern OUTPUTPATTERN
                        Output file pattern (default: -orepet)
--span SPAN            Span window length in number of IPUs (default: 3)
--stopwords STOPWORDS Add stop-words estimated from the given data
                        (default: True)
--alpha ALPHA          Coefficient to add data-specific stop-words
                        (default: 0.5)
```

## 3.21 Re-Occurrences

This annotation is searching for re-occurrences of an annotation of a speaker in the next N annotations of the interlocutor. It is originally used for gestures in (M. Karpinski et al. 2018).

Maciej Karpinski, Katarzyna Klessa Methods, Tools and Techniques for Multimodal Analysis of Accommodation in Intercultural Communication CMST 24(1) 29–41 (2018), DOI:10.12921/cmst.2018.0000006

### 3.21.1 Perform Re-Occurrences with the GUI

The automatic annotation takes as input any annotated file with (at least) one tier, and another file+tier of the interlocutor. The annotation provides one annotated file with 2 tiers: Sources and Repetitions.

Click on the Re-Occurrences activation button, and click on the “Configure...” blue text to fix options.

### 3.21.2 Perform Re-Occurrences with the CLI

```
usage: reoccurrences.py [files] [options]
```

Files:

```
-i file           Input file name with time-aligned annotations of
                  the main speaker.
-s file           Input file name with time-aligned annotations of
                  the interlocutor
-o file           Output file name with re-occurrences.
```

Options:

```
--inputpattern INPUTPATTERN
                        Input file pattern (default: )
--outputpattern OUTPUTPATTERN
                        Output file pattern (default: -reocc)
--tiername TIERNAME    Tier to search for re-occurrences (default: )
--span SPAN            Span window length in number of annotations (default:
                        10)
```

This program is part of SPPAS version 2.4. Copyright (C) 2011-2019 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

## 3.22 Momel (modelling melody)

Momel is an algorithm for the automatic modeling of fundamental frequency (F0) curves using a technique called asymmetric modal quadratic regression.

This technique makes it possible by an appropriate choice of parameters to factor an F0 curve into two components:

- a macro-prosodic component represented by a quadratic spline function defined by a sequence of target points < ms, hz >.
- a micro-prosodic component represented by the ratio of each point on the F0 curve to the corresponding point on the quadratic spline function.

For details, see the following reference:

**Daniel Hirst and Robert Espesser** (1993). *Automatic modelling of fundamental frequency using a quadratic spline function*. Travaux de l'Institut de Phonétique d'Aix. vol. 15, pages 71-85.

The SPPAS implementation of Momel requires a file with the F0 values **sampled at 10 ms**. Two file formats are supported:

- “.PitchTier”, from Praat.
- “.hz”, from any tool. It is a file with one F0 value per line.

The following options can be fixed:

- Window length used in the “cible” method
- F0 threshold: Maximum F0 value
- F0 ceiling: Minimum F0 value
- Maximum error: Acceptable ratio between two F0 values
- Window length used in the “reduc” method
- Minimal distance
- Minimal frequency ratio
- Eliminate glitch option: Filter f0 values before ‘cible’

### 3.22.1 Perform Momel with the GUI

It is an annotation of STANDALONE type.

Click on the Momel activation button then click on the “Configure...” blue text to fix options.

### 3.22.2 Perform Momel with the CLI

`momel.py` is the program to perform Momel annotation of a given file with F0 values sampled at 10ms.

#### Usage

```
momel.py [files] [options]
```

```
Momel: Proposed by D. Hirst and R. Espesser, Momel - Modelling of fundamental
frequency (F0) curves is using a technique called assymetric modal quaratic
regression. Requires pitch values.
```

```
optional arguments:
```

```
-h, --help          show this help message and exit
--quiet            Disable the verbosity
--log file         File name for a Procedure Outcome Report (default: None)
```

```
Files (manual mode):
```

```
-i file            Input file name (extension: .hz or .PitchTier)
-o file            Output file name (default: stdout)
```

```
Files (auto mode):
```

```

-I file          Input file name with pitch (append).
-e .ext         Output file extension. One of: .xra .TextGrid .eaf .csv
               .mrk .txt .stm .ctm .lab .mlf .sub .srt .antx .arff .xrff

```

## Options:

```

--outputpattern OUTPUTPATTERN
                Output file pattern (default: -momel)
--win1 WIN1     Target window length (default: 30)
--lo LO         F0 threshold (default: 50)
--hi HI         F0 ceiling (default: 600)
--maxerr MAXERR Maximum error (default: 1.04)
--win2 WIN2     Reduce window length (default: 20)
--mind MIND     Minimal distance (default: 5)
--minr MINR    Minimal frequency ratio (default: 0.05)

```

This program is part of SPPAS version 2.4. Copyright (C) 2011–2019 Brigitte Bigi. Contact the author at: [contact@sppas.rg](mailto:contact@sppas.rg)

**Examples of use**

```

python .\sppas\bin\momel.py -i .\samples\samples-eng\ENG_M15_ENG_T02.PitchTier
2018-12-19 15:44:00,437 [INFO] Logging set up level=15
2018-12-19 15:44:00,674 [INFO] ... .. 41 anchors found.
1.301629 109.285503
1.534887 126.157058
1.639614 143.657446
1.969234 102.911464
2.155284 98.550759
2.354162 108.250869
2.595364 87.005994
2.749773 83.577924
2.933222 90.218382
3.356651 119.709142
3.502254 104.104568
3.707747 132.055286
4.000578 96.262109
4.141915 93.741407
4.383332 123.996736
4.702203 89.152708
4.987086 101.561180
5.283864 87.499710
5.538984 92.399690
5.707147 95.411586
5.906895 87.081095
6.705373 121.396919
7.052992 130.821479
7.218415 120.917642
7.670083 101.867028
7.841935 109.094053
8.124574 90.763267
8.455182 114.261067
8.746016 93.704705
9.575359 101.108444
9.996245 122.488120

```

```

10.265663 105.244429
10.576394 94.875460
11.730570 99.698799
12.083323 124.002313
12.411790 108.563104
12.707442 101.928297
12.963805 113.980850
13.443483 90.782781
13.921939 90.824376
14.377324 60.126506

```

Apply Momel on all files of a given folder:

```
python .\sppas\bin\momel.py -I .\samples\samples-eng
```

### 3.23 INTSINT: Encoding of F0 anchor points

INTSINT assumes that pitch patterns can be adequately described using a limited set of tonal symbols, T,M,B,H,S,L,U,D (standing for : Top, Mid, Bottom, Higher, Same, Lower, Up-stepped, Down-stepped respectively) each one of which characterises a point on the fundamental frequency curve.

The rationale behind the INTSINT system is that the F0 values of pitch targets are programmed in one of two ways : either as absolute tones T, M, B which are assumed to refer to the speaker's overall pitch range (within the current Intonation Unit), or as relative tones H, S, L, U, D assumed to refer only to the value of the preceding target point.

INTSINT example

The rationale behind the INTSINT system is that the F0 values of pitch targets are programmed in one of two ways : either as absolute tones T, M, B which are assumed to refer to the speaker's overall pitch range (within the current Intonation Unit), or as relative tones H, S, L, U, D assumed to refer only to the value of the preceding target point.

A distinction is made between non-iterative H, S, L and iterative U, D relative tones since in a number of descriptions it appears that iterative raising or lowering uses a smaller F0 interval than non-iterative raising or lowering. It is further assumed that the tone S has no iterative equivalent since there would be no means of deciding where intermediate tones are located.

**D.-J. Hirst** (2011). *The analysis by synthesis of speech melody: from data to models*, Journal of Speech Sciences, vol. 1(1), pages 55-83.

#### 3.23.1 Perform INTSINT with the GUI

It is an annotation of STANDALONE type.

Click on the INTSINT activation button and click on the "Configure..." blue text to fix options.

#### 3.23.2 Perform INTSINT with the CLI

`intsint.py` is the program to perform INTSINT annotation of a given file with momel anchors.

**Usage**

```
intsint.py [files] [options]
```

INTSINT: International Transcription System for INTonation codes the intonation of an utterance by means of an alphabet of 8 discrete symbols. Requires Momel targets.

optional arguments:

```
-h, --help  show this help message and exit
--quiet     Disable the verbosity
--log file  File name for a Procedure Outcome Report (default: None)
```

Files (manual mode):

```
-i file      Input file name with anchors.
-o file      Output file name (default: stdout)
```

Files (auto mode):

```
-I file      Input file name with anchors (append).
-e .ext      Output file extension. One of: .xra .TextGrid .eaf .csv .mrk
             .txt .stm .ctm .lab .mlf .sub .srt .antx .arff .xrff
```

Options:

```
--inputpattern INPUTPATTERN
             Input file pattern (momel anchors) (default: -momel)
--outputpattern OUTPUTPATTERN
             Output file pattern (default: -intsint)
```

This program is part of SPPAS version 2.4. Copyright (C) 2011-2019 Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

**Examples of use**

Apply INTSINT on a single file and print the result on the standard output:

```
python .\sppas\bin\intsint.py -i .\samples\samples-eng\ENG_M15_ENG_T02-momel.xra --quiet
1.301629 M
1.534887 U
1.639614 H
1.969234 L
2.155284 S
2.354162 U
2.595364 L
2.749773 S
2.933222 S
3.356651 H
3.502254 D
3.707747 H
4.000578 L
4.141915 S
4.383332 H
4.702203 L
4.987086 U
```

```
5.283864 L
5.538984 U
5.707147 D
5.906895 S
6.705373 M
7.052992 U
7.218415 S
7.670083 D
7.841935 S
8.124574 D
8.455182 U
8.746016 D
9.575359 M
9.996245 U
10.265663 D
10.576394 D
11.730570 M
12.083323 U
12.411790 D
12.707442 S
12.963805 U
13.443483 L
13.921939 S
14.377324 B
```

Apply INTSINT in auto mode:

```
python .\sppas\bin\intsint.py -I .\samples\samples-eng\ENG_M15_ENG_T02.wav
python .\sppas\bin\intsint.py -I .\samples\samples-eng\ENG_M15_ENG_T02.PitchTier
python .\sppas\bin\intsint.py -I .\samples\samples-eng\ENG_M15_ENG_T02-momel.xra
```

## 3.24 Hand&Pose

SPPAS is a wrapper for MediaPipe Hand detection and Mediapipe Pose detection. It also proposes a custom solution in order to detect right-left hands of a person.

### 3.24.1 Overview

MediaPipe Hands is a high-fidelity hand and finger tracking solution. It employs machine learning (ML) to infer 21 3D landmarks of a hand from just a single frame. MediaPipe Hands utilizes an ML pipeline consisting of multiple models working together: A palm detection model that operates on the full image and returns an oriented hand bounding box. A hand landmark model that operates on the cropped image region defined by the palm detector and returns high-fidelity 3D hand key points. For details about hand detection, see: <https://google.github.io/mediapipe/solutions/hands.html>

MediaPipe Pose is a ML solution for high-fidelity body pose tracking, inferring 33 3D landmarks. For details about pose detection, see: <https://google.github.io/mediapipe/solutions/pose.html>

Important:

- SPPAS is only exporting (x,y)
- SPPAS can export the results in either XRA or CSV formats.

Three detection “mode” are available:

1. hands only: when enabling hands and disabling pose, SPPAS is a simple wrapper for MediaPipe Hand detection. The result of the system is a list of 21 sights for each detected hand.
2. pose only: when disabling hand detection and enabling pose detection, SPPAS is a simple wrapper for MediaPipe Pose detection. Notice that if several persons are on the image, it will detect only one human body. The result of the system is a list of the 33 sights of the human body.
3. hand&pose: when enabling both hand and pose, SPPAS is combining both pose detections and hand detections. It results in two lists of sights: the first list for the right hand and the second list of the left hand. Each list is made either of 21 sights if the hand was properly detected by the hand prediction system, or made of the 4 sights of the pose detection system.

Here is the match between the indexes of the 8 sights of the pose detection and the indexes of the sights of the hand detection:

- right hand: pose 16 is hand 0, pose 18 is hand 8, pose 20 is hand 20 and pose 22 is hand 4.
- left hand: pose 15 is hand 0, pose 17 is hand 8, pose 19 is hand 20 and pose 21 is hand 4.

### 3.24.2 Perform annotation with the GUI

It is a STANDALONE annotation.

The annotation process takes as input an image or a video file. To perform the annotation, click on its activation button and click on the “Configure...” blue text to fix options.

### 3.24.3 Perform with the CLI

The CLI does not work on an image but video recordings only. To perform this annotation on an image, use the script “annotation.py” instead.

```
usage: handpose.py [files] [options]

optional arguments:
  -h, --help            show this help message and exit
  --quiet               Disable the verbosity
  --log file            File name for a Procedure Outcome Report
                       (default: None)

Files:
  -i file              Input video.
  -o file              Output base name.
  -I file              Input file name (append).
  -e .ext              Output file extension. One of: .mp4 .avi .mkv .mov

Options:
  --inputpattern1 INPUTPATTERN1
                       Input pattern of the video file. (default: )
```

---

```

--outputpattern OUTPUTPATTERN          Output file pattern with the sights. (default: -hands)
--hand HAND                             Enable hands detection. (default: True)
--pose POSE                             Enable pose detection -- for only one human body. (default: True)
--csv CSV                               Save points in a CSV file instead of XRA (default: False)
--tag TAG                                Draw points on the video (default: True)
--folder FOLDER                         Save result in a folder with image files -- if video input only (de

```

## 3.25 Face Detection

This is a state-of-the-art implementation of face detection performed with freely available models and tools. We introduced an original method in order to use any or all of them and to combine their results. This combined result was not evaluated.

### 3.25.1 Overview

FaceDetection annotation allows to search for coordinates of faces in an image or in all images of a video. It requires both to enable the “video” feature in the setup to install the external libraries ‘numpy’, ‘opencv-contrib’ and ‘mediapipe’. It also requires to check “facedetect” in the list of annotations to be installed at the time of the setup.

On the basis of the detection methods implemented in the opencv library, SPPAS is able to use several systems and to combine their results. These systems are based on 2 different methods:

1. an Artificial Neural Network (DNN);
2. an Haar Cascade Classifier (HCC).

The linguistic resources of this annotation include two DNN models and two models for HCC (a frontal-face model and a profile-face one). SPPAS can also launch the MediaPipe face detection system which is much more faster than the previous ones but results sounds worse.

By default, SPPAS launches two of these detectors: 1 DNN and 1 HCC, and it combines their results. This annotation is about 2.5x real time. Even if it can increase the quality of the final result, other models are not used by default because the detection is very slow: 15x real time to use all of the 5 models. The options allow to choose the models to be used.

### 3.25.2 Result of Face Detection

There are several output files that can be created:

- a copy of the image/video with all the detected faces surrounded by a square indicating a confidence score;
- as many cropped image files as the number of detected faces;
- an XRA or a CSV file with coordinates and confidence score of each detected face.

There’s also the possibility to consider the selfie (portrait size) instead of the face.

### 3.25.3 Perform Face Detection with the GUI

It is a STANDALONE annotation.

The Face Detection process takes as input an image file and/or a video. To perform the annotation, click on the FaceDetection activation button and click on the “Configure...” blue text to fix options.

### 3.25.4 Perform Face Detection with the CLI

facetedetection.py is the program to perform Face Detection annotation of a given media file.

#### Usage

```
usage: facetedetection.py [files] [options]
```

optional arguments:

```
-h, --help          show this help message and exit
--quiet            Disable the verbosity
--log file         File name for a Procedure Outcome Report (default: None)
```

Files:

```
-i file            Input image.
-o file            Output base name.
-I file            Input file name (append).
-r model           Model base name (.caffemodel or .xml models as wishes)
-e .ext            Output file extension (image or video)
```

Options:

```
--inputpattern INPUTPATTERN
                    Input file pattern (default: )
--nbest NBEST       Number of faces to select among those
                    detected (0=auto) (default: 0)
--score SCORE       Minimum confidence score to select detected
                    faces (default: 0.2)
--portrait PORTRAIT Consider the portrait instead of the face in
                    outputs (default: False)
--csv CSV           Save coordinates of detected faces in a CSV
                    file instead of XRA (default: False)
--folder FOLDER     Save result in a folder with image files --
                    if video input only (default: False)
--tag TAG           Surround the detected faces in the output
                    image (default: True)
--crop CROP         Save detected faces in cropped images
                    (default: False)
--width WIDTH       Resize all the cropped images to a fixed
                    width (0=no) (default: 0)
--height HEIGHT     Resize all the cropped images to a fixed
                    height (0=no) (default: 0)
--model:opencv_face_detector_uint8.pb MODEL
                    Enable the opencv's ANN TensorFlow model.
                    (default: True)
--model:haarcascade_frontalface_alt.xml MODEL
                    Enable the opencv's HaarCascade Frontal face model.
                    (default: True)
```

```

--model:res10_300x300_ssd_iter_140000_fp16.caffemodel MODEL
    Enable the opencv's ANN Caffe model.
    (default: False)
--model:haarcascade_profileface.xml MODEL
    Enable the opencv's HaarCascade Profile face model.
    (default: False)
--model:mediapipe MODEL:MEDIAPIPE
    Enable the MediaPipe Face Detection system.
    (default: False)

```

This program is part of SPPAS version 4.2. Copyright (C) 2011-2021  
Brigitte Bigi. Contact the author at: [contact@sppas.org](mailto:contact@sppas.org)

### Examples of use

```

python3 ./sppas/bin/facedetection.py -I ./samples/faces/BrigitteBigi_Aix2020.png --tag=True
[INFO] Logging redirected to StreamHandler (level=0).
[INFO] SPPAS version 3.5
[INFO] Copyright (C) 2011-2021 Brigitte Bigi
[INFO] Web site: http://www.sppas.org/
[INFO] Contact: Brigitte Bigi (contact@sppas.org)
[INFO] * * * Annotation step 0 * * *
[INFO] Number of files to process: 1
[INFO] Options:
[INFO] ... inputpattern:
[INFO] ... outputpattern: -face
[INFO] ... nbest: 0
[INFO] ... score: 0.2
[INFO] ... portrait: True
[INFO] ... csv: True
[INFO] ... tag: True
[INFO] ... crop: True
[INFO] ... width: 0
[INFO] ... height: 0
[INFO] File BrigitteBigi_Aix2020.png: Valid.
[INFO] ... .. 3 faces found.
[INFO] ... ./samples/faces/BrigitteBigi_Aix2020-face.jpg

```

It creates the following 5 files in the samples/faces folder:

- BrigitteBigi\_Aix2020-face.jpg with the 3 detected faces surrounded by a square and indicating the detection score
- BrigitteBigi\_Aix2020\_1-face.jpg: image of the face with the highest score
- BrigitteBigi\_Aix2020\_2-face.jpg: image of the face with the medium score
- BrigitteBigi\_Aix2020\_3-face.jpg: image of the face with the worse score
- BrigitteBigi\_Aix2020-face.csv contains 3 lines with coordinates in columns 3-7 with (x,y) and (w,h) then the confidence score in column 8 ranging [0., 1.]

Notice that the image contains 3 faces and their positions are properly found.

## 3.26 Face Identity

This is a new and original automatic annotation but it's still in progress. It has to be evaluated.

### 3.26.1 Overview

Face Identity automatic annotation assigns a person identity to detected faces of a video. It takes as input a video and a CSV file with coordinates of the detected faces. It produces a CSV file with coordinates of the identified faces. Assigned persons names are "id-00x". Obviously, the CSV file can be edited and such names can be changed *a posteriori*.

This annotation requires to enable the "video" feature in the setup, so it will install the external python libraries 'numpy' and 'opencv-contrib'.

No external resources are needed.

### 3.26.2 Perform annotation with the GUI

It is a STANDALONE annotation.

The Face Identity process takes as input a video file. To perform the annotation, click on the Face Identity activation button and click on the "Configure..." blue text to fix options.

### 3.26.3 Perform with the CLI

`faceidentity.py` is the program to perform Face Identity annotation of a given video file, if the corresponding CSV file with detected faces is existing.

#### Usage

```
usage: faceidentity.py [files] [options]

optional arguments:
  -h, --help            show this help message and exit
  --quiet               Disable the verbosity
  --log file            File name for a Procedure Outcome Report (default: None)

Files:
  -i file              Input video.
  -c file              Input CSV file with face coordinates and sights.
  -o file              Output base name.
  -I file              Input file name (append).
  -e .ext              Output file extension. One of: .mp4 .avi .mkv

Options:
  --inputpattern INPUTPATTERN
  --inputoptpattern INPUTOPTPATTERN (default: -face)
  --outputpattern OUTPUTPATTERN (default: -ident)
```

## 3.27 Face Landmarks

This is a state-of-the-art implementation of face landmark performed with freely available models and tools. We introduced a solution to combine the results when several methods are used. The combined result was not evaluated.

### 3.27.1 Overview

SPPAS is using both the MediaPipe Face Mesh and the OpenCV's facial landmark API called Facemark. It includes three different implementations of landmark detection based on three different papers:

- **FacemarkKazemi:** This implementation is based on a paper titled “One Millisecond Face Alignment with an Ensemble of Regression Trees” by V.Kazemi and J. Sullivan published in CVPR 2014.
- **FacemarkAAM:** This implementation uses an Active Appearance Model and is based on an the paper titled “Optimization problems for fast AAM fitting in-the-wild” by G. Tzimiropoulos and M. Pantic, published in ICCV 2013.
- **FacemarkLBF:** This implementation is based a paper titled “Face alignment at 3000 fps via regressing local binary features” by S. Ren published in CVPR 2014.

The fundamental concept is that any person will have 68 particular points on the face (called sights). SPPAS is able to launch several of them and to combine their results in a single and hopefully better one. Actually, SPPAS is launching MediaPipe Face Mesh and extracting the 68 sights among the 468 that are detected ; then this result is combined (weight=6) with the 68 sights of the LBF detection method (weight=1).

This annotation requires both to enable the “video” feature in the setup in order to install the external libraries ‘numpy’ and ‘opencv-contrib’, and to check “facemark” in the list of annotations to be installed. Two different models will be downloaded and used: a Kazemi one and a LBF one.

### 3.27.2 Perform annotation with the GUI

It is a STANDALONE annotation.

The Face Sights process takes as input an image file and/or a video. To perform the annotation, click on the Face Sights activation button and click on the “Configure...” blue text to fix options.

### 3.27.3 Perform with the CLI

```
usage: facesights.py [files] [options]

optional arguments:
  -h, --help            show this help message and exit
  --quiet              Disable the verbosity
  --log file           File name for a Procedure Outcome Report

Files:
  -i file              Input image.
  -o file              Output base name.
  -I file              Input file name (append).
  -r model             Landmark model name (Kazemi, LBF or AAM)
  -R model             FaceDetection model name
```

-e .ext                      Output file extension.

Options:

```
--inputpattern INPUTPATTERN
--inputoptpattern INPUTOPTPATTERN (default: -face)
--outputpattern OUTPUTPATTERN (default: -sights)
```

## 3.28 Cued speech - LfPC

This automatic annotation is currently under development. It is in progress and it's currently only a Proof of Concept. **It mustn't be used neither for any final application nor evaluation.** It can only be used in order to test it and to contribute in improving it.

### 3.28.1 Overview

Speech reading or lip-reading requires watching the lips of a speaker and is used for the understanding of the spoken sounds. However, various sounds have the same lips movement which implies a lot of ambiguity. In 1966, R. Orin Cornett invented the Cued Speech, a visual system of communication. It adds information about the pronounced sounds that are not visible on the lips.

Thanks to this code, speech reading is encouraged since the Cued Speech (CS) keys match all the spoken phonemes but phonemes with the same movement have different keys. Actually, from both the hand position on the face (representing vowels) and hand shapes, known as cues (representing consonants), CV syllables can be represented. So, a single CV syllable will be generated or decoded through both the lips position and the key of the hand.

LfPC is the French acronym for “Langue française Parlée Complétée”.

The conversion of phonemes into keys of CS is performed using a rule-based system. This RBS phoneme-to-key segmentation system is based on the only principle that a key is always of the form CV.

This annotation requires both to enable the “video” feature in the setup to install the external libraries ‘numpy’ and ‘opencv-contrib’ and to check “cuedspeech” in the list of annotations.

### 3.28.2 Perform annotation with the GUI

It is a STANDALONE annotation.

The annotation process takes as input a “-palign” file and optionally a video. To perform the annotation, click on its activation button and click on the “Configure...” blue text to fix options.

### 3.28.3 Perform with the CLI

```
usage: cuedspeech.py [files] [options]
```

optional arguments:

```
-h, --help                      show this help message and exit
--quiet                          Disable the verbosity
--log file                        File name for a Procedure Outcome Report
                                  (default: None)
```

Files:

```
-i file           Input time-aligned phonemes file name.
-v file           Input video file name.
-o file           Output file name with Cued Speech key codes.
-r rules          File with Cued Speech keys description

-I file           Input file name (append).
-l lang           Language code (iso8859-3). One of: fra.
-e .ext           Output file extension. One of: .xra .TextGrid
                  .eaf .ant .antx .mrk .lab .srt .sub .vtt .ctm
                  .stm .csv .txt
```

Options:

```
--inputpattern1 INPUTPATTERN1
                    Pattern of the file with time-aligned phonemes
                    (default: -palign)
--inputpattern2 INPUTPATTERN2
                    Pattern of the video file (default: )
--inputpattern3 INPUTPATTERN3
                    Pattern of the file with sights of the face of
                    the person (default: -sights)
--outputpattern OUTPUTPATTERN
                    Pattern of the output files (default: -cuedsp)
--createvideo CREATEVIDEO
                    Tag the video with the code of the key
                    (needs video+csv) (default: False)
```



## Convert files

### 4.1 Interoperability and compatibility: an introduction

The conversion of a file to another file is the process of changing the form of the presentation of the data, and not the data itself. Every time, when data file is to be used, they must be converted to a readable format for the next application. A data conversion is normally an automated process to some extent. SPPAS provide the possibility to automatically import and export the work done on some various file formats from a wide range of other software tools. For the users, the visible change will be only a different file extension but for software it is the difference between understanding of the contents of the file and the inability to read it.

The conversion of file formats is then a difficult task and it can imply that some data are left. Representing annotated data in SPPAS is of crucial importance for its automatic annotations, its analysis of annotations and for the software to be able to automatically annotate and analyze any kind of data files. SPPAS then includes an original and generic enough annotation representation framework. This framework for annotations is very rich and contain several information like alternative labels or alternative localizations of annotations, a tier hierarchy, controlled vocabularies, etc. A native format named XRA was then developed to fit in such data representation. The physical level of representation of XRA obviously makes use of XML, XML-related standards and stand-off annotation. Due to an intuitive naming convention, XRA documents are human readable as far as possible within the limits of XML.

### 4.2 SPPAS conversion method

In the scope of the compatibility between SPPAS data and annotated data from other software tools or programs, SPPAS is able to open/save and import/export several file formats.

SPPAS makes use of its internal data representation to convert files. A conversion then consists of two steps:

1. the incoming file is loaded and mapped to the SPPAS data framework;
2. such data is saved to the expected format.

These two steps are applied whatever the organization structure of annotations in the original or in the destination file format. This process allows SPPAS to import from and export to a variety of file formats. This is

illustrated by the next Figure which includes the list of file formats and the corresponding software that are supported. Arrows are representing the following actions:

- import from, represented by a continued line with an arrow from the file format to the SPPAS framework;
- partially import from, represented by a dash line with an arrow from the file format to the SPPAS framework;
- export to, represented by an arrow from the SPPAS framework to the file format.

SPPAS conversion method and formats

### 4.3 Supported file formats

SPPAS supports the following software with their file extensions:

- Praat: TextGrid, PitchTier, IntensityTier
- Elan: eaf
- Annotation Pro: ant, antx
- Phonedit: mrk
- Sc-lite: ctm, stm
- HTK: lab, mlf
- Subtitles: srt, sub
- Signaux: hz
- Spreadsheets, R, ...: csv
- Audacity, notepads: txt

The followings can be imported:

- ANVIL: anvil
- Transcriber: trs
- Xtrans: tdf

The followings can be exported:

- Weka: arff, xrff
- Subtitles: vtt

The support of external formats is regularly extended to new formats by the author on-demand from the users and distributed to the community in the SPPAS regular updates.

### 4.4 Convert with the UI of SPPAS

The page ‘Convert’ of the GUI allows to convert checked files of the workspace into the selected format. A table is indicating which of the features are supported by each format. It allows to know what to expect when converting from one format to the other one. For example, a Praat TextGrid file does not contain metadata nor media information neither a controlled vocabulary nor a hierarchy. If the file to be converted contains such information, they’ll be lost in the TextGrid converted file.

The program `bin/trsconvert.py` allows to convert files with the CLI.

# Data Analyses

## 5.1 Introduction

The analyses of annotated files includes the descriptive statistics, the filtering of the annotated data to get only the annotations you are interested in, see/edit the information about files, etc.

Like the other features of SPPAS, analyzing data can be performed in three different ways:

- the Application Programming Interface, in Python language;
- the Command-Line User Interface;
- the Graphical User Interface.

Among the features implemented in the API, a big majority are included in the GUI but just a few can be performed with the CLI. This chapter then describes only the page “Analyze” of the GUI.

## 5.2 The page Analyze of the GUI

The page “Analyze” of the GUI is divided into 2 main areas: a toolbar and a content to represent files.

### 5.2.1 Displayed files content

Each opened file is displayed in a panel in the content area. The next Figure shows how panels of 4 different files look like. Panels of the annotated files have a yellow background, panels of the audio files are blue and panels of unknown files are pinky-red.

Analyze: open files

Some actions can be performed individually on the panel of a file. A mouse click on the filename or on the arrow at left will show or hide the content of the file.

For transcription files, the icons at top-right of the panel allow to:

- edit the metadata of the file - notice that only the XRA file format allows to save the metadata;

- select the file (for example, it's used to paste tiers into);
- save the file - only if it has changed;
- close the file and unlock it in the workspace.

For audio files, the icons at top-right of the panel allow to:

- view audio content and manage channels - available only for files less than a few minutes;
- close the file.

### 5.2.2 The toolbar

The toolbar of the page is made of 3 different parts:

1. "Files:" click on these buttons to perform an action on the checked filenames of the page "Files";
2. "Tiers:" click on these buttons to perform an action on the checked tiers of the opened files;
3. "Annotations:" click on these buttons to analyze the annotations of the checked tiers of the opened files.

#### **Files: Open files**

Open, load and display a panel for each checked file of the current workspace. Some files could need a long time to be loaded (like TextGrid files with a lot of annotations), so a scrollbar should indicate the progression. As soon as a file is opened, it is locked and no other page can perform an action on it.

To open new files, check new files in the current workspace and click again on the "Open files" button. The panels of the newly opened files will be appended to the existing ones.

#### **Files: New file**

Click on this button to create a new annotated file. A dialog box will ask for a path/filename and for a file extension. The extension defines the file format; any of the supported file format can be used. The file will be created on disk when it will be saved for the first time.

#### **Files: Save all**

Save all files for which some changes were done, without confirmation.

#### **Files: Close all**

Close all the opened files. If some files were changed and not saved, a dialog will ask for confirmation.

#### **Tiers: Metadata**

Open a dialog to edit the metadata of the checked tiers. Notice that most of the file format don't allow to save metadata or it allows only some specific ones; only the XRA format allows to save any tier metadata.

**Tiers: Check**

A click on the “Check” button will open a dialog to enter a tier name and it will check all tiers matching it. The entry is a regular expression.

**Tiers: Uncheck**

A click on the “Uncheck” button will un-check all tiers.

**Tiers: Rename**

A mouse click on the “Rename” button will open a dialog to fix a name of a tier and then to rename all checked tiers. If a file has already a tier with the given name, an index number will be appended to the new name.

**Tiers: Delete**

A mouse click on the “Delete” button will delete all checked tiers. This process is irreversible. To recover a deleted tier, the only way is to close the file and to re-open it but all the un-saved changes are lost.

**Tiers: Cut/Copy/Paste**

The Cut/Copy/Paste buttons allow to use a clipboard to manage tiers. Tiers can then be copied from files to other ones. The files to paste in must be selected first with the “select” button of the individual panels.

**Tiers: Duplicate**

The duplicate button allows to copy/paste a tier into the same file. The name of the duplicated tier will be the same as the original one with an index number at the end.

**Tiers: Move Up/Move Down**

These buttons allow to move checked tiers into a file.

**Annotations: Radius**

The “Radius” button allows to fix a radius for all the localizations of the annotations of the checked tiers, i.e. the vagueness around the fixed point in time. Notice that only XRA file format allows to save it.

Read the following paper for details:

*Brigitte Bigi, Tatsuya Watanabe, Laurent Prévot* (2014). **Representing Multimodal Linguistics Annotated Data**, 9th International conference on Language Resources and Evaluation (LREC), Reykjavik (Iceland), pages 3386-3392. ISBN: 978-2-9517408-8-4.

**Annotations: View**

Click the button “View” to see all the annotations of the checked tiers in a table.

**Annotations: Statistics**

It allows to estimate the number of occurrences, the duration, etc. of the annotations of the checked tiers, and allows to save result in CSV (for Excel, OpenOffice, R, MatLab,...).

It offers a serie of sheets organized in a notebook. The first tab is displaying a summary of descriptive statistics of the set of given tiers. The other tabs are indicating one of the statistics over the given tiers. The followings are estimated:

- occurrences: the number of observations
- total durations: the sum of the durations
- mean durations: the arithmetic mean of the duration
- median durations: the median value of the distribution of durations
- std dev. durations: the standard deviation value of the distribution of durations

All of them can be estimated on a single annotation label or on a serie of them. The length of this context can be optionally changed while fixing the “N-gram” value (available from 1 to 5), just above the sheets.

Each displayed sheet can be saved as a CSV file, which is a useful file format to be read by R, Excel, OpenOffice, LibreOffice, and so... To do so, display the sheet you want to save and click on the button “Save sheet”, just below the sheets. If you plan to open this CSV file with Excel under Windows, it is recommended to change the encoding to UTF-16. For the other cases, UTF-8 is probably the most relevant.

The annotation durations are commonly estimated on the Midpoint value, without taking the radius into account; see (Bigi et al, 2012) for explanations about the Midpoint/Radius. Optionally, the duration can either be estimated by taking the vagueness into account, then check “Add the radius value” button, or by ignoring the vagueness and estimating only on the central part of the annotation, then check “Deduct the radius value”.

For those who are estimating statistics on XRA files, you can either estimate stats only on the best label (the label with the higher score) or on all labels, i.e. the best label and all its alternatives (if any).

Descriptive statistics

**Annotations: Single filter**

It allows to create a new tier with only selected annotations: define filters in order to create new tiers with only the annotations you are interested in!

Pattern selection is an important part to extract data of a corpus and is obviously and important part of any filtering system. Thus, if the label of an annotation is a string, the following filters are proposed in DataFilter:

- exact match: an annotation is selected if its label strictly corresponds to the given pattern;
- contains: an annotation is selected if its label contains the given pattern;
- starts with: an annotation is selected if its label starts with the given pattern;
- ends with: an annotation is selected if its label ends with the given pattern.

All these matches can be reversed to represent respectively: does not exactly match, does not contain, does not start with or does not end with. Moreover, this pattern matching can be case sensitive or not.

For complex search, a selection based on regular expressions is available for advanced users.

A multiple pattern selection can be expressed in both ways:

- enter multiple patterns at the same time (separated by commas) to mention the system to retrieve either one pattern or the other, etc.
- enter one pattern at a time and choose the appropriate button: “Apply All” or “Apply any”.

Frame to create a filter on annotation label tags: filter annotations that exactly match either a, @ or E

Another important feature for a filtering system is the possibility to retrieve annotated data of a certain duration, and in a certain range of time in the timeline.

Frame to create a filter on annotation durations: filter annotations that are during more than 80 ms

Search can also start and/or end at specific time values in a tier.

Frame to create a filter on annotation time values: filter annotations that are starting after the 5th minute

In SPPAS 3.7, a new filter is added: it can select annotation depending on their number of labels. For example, the automatic annotation “Normalization” creates a tier “Tokens” in which each annotation contains a list of labels - one per token; then it is possible to get the annotations with more than 3 tokens, with only one token, etc.

All the given filters are summarized in the “SingleFilter” dialog. To complete the filtering process, it must be clicked on one of the apply buttons and the new resulting tiers are added in the annotation file(s).

In the given example:

- click on “Apply All” to get either a, @ or E vowels during more than 80ms, after the 5th minute.
- click on “Apply Any” to get a, @ or E vowels, and all annotations during more than 80 ms, and all annotations after the 5th minute.

DataFilter: SingleFilter frame

Read the following publications for details:

*Brigitte Bigi* (2019). **Filtering multi-levels annotated data**. In 9th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, pp. 13-14, Poznań, Poland.

*Brigitte Bigi, Jorane Saubesty* (2015). **Searching and retrieving multi-levels annotated data**, Proceedings of Gesture and Speech in Interaction, Nantes (France).

### **Annotations: Relation filter**

Regarding the searching problem, linguists are typically interested in locating patterns on specific tiers, with the possibility to relate different annotations a tier from another. The proposed system offers a powerful way to request/extract data, with the help of Allen’s interval algebra.

In 1983 James F. Allen published a paper in which he proposed 13 basic relations between time intervals that are distinct, exhaustive, and qualitative:

- distinct because no pair of definite intervals can be related by more than one of the relationships;
- exhaustive because any pair of definite intervals are described by one of the relations;
- qualitative (rather than quantitative) because no numeric time spans are considered.

These relations and the operations on them form Allen's interval algebra. These relations were extended to Interval-Tiers as Point-Tiers to be used to find/select/filter annotations of any kind of time-aligned tiers.

For the sake of simplicity, only the 13 relations of the Allen's algebra are available in the GUI. But actually, we implemented the 25 relations proposed Pujari and al. (1999) in the INDU model. This model is fixing constraints on INtervals (with Allen's relations) and on DUration (duration are equals, one is less/greater than the other). Such relations are available while requesting with Python.

At a first stage, the user must select the tiers to be filtered and click on "RelationFilter". The second stage is to select the tier that will be used for time-relations.

Fix time-relation tier name

The next step consists in checking the Allen's relations that will be applied. The last stage is to fix the name of the resulting tier. The above screenshots illustrates how to select the first phoneme of each token, except for tokens that are containing only one phoneme (in this later case, the "equal" relation should be checked).

DataFilter: RelationFilter frame

To complete the filtering process, it must be clicked on the "Apply" button and the new resulting tiers are added in the annotation file(s).

# Scripting with Python and SPPAS

## 6.1 Introduction

Since version 1.8.7, SPPAS implements an Application Programming Interface (API), named *anndata*, to deal with annotated files. Previously, SPPAS was based on another API named *annotationdata* which is still distributed in the package but no longer updated or maintained.

*anndata* API is a free and open source Python library to access and search data from annotated data of any of the supported formats (xra, TextGrid, eaf...). It can either be used with the Programming Language Python 2.7 or Python 3.4+. This API is PEP8 and PEP257 compliant and the internationalization of the messages is implemented (English and French are available in the “po” directory).

In this chapter, it is assumed that a version of Python is installed and configured. It is also assumed that the Python IDLE is ready-to-use. For more details about Python, see:

The Python Website: <http://www.python.org>

This chapter firstly introduces basic programming concepts, then it gradually introduces how to write scripts with Python. Those who are familiar with programming in Python can directly go to the last section related to the description of the *anndata* API and how to use it in Python scripts.

This API can convert file formats like Elan’s EAF, Praat’s TextGrid and others into a `sppasTranscription` object and convert this object into any of these formats. This object allows unified access to linguistic data from a wide range sources.

This chapter includes exercises. The solution scripts are included in the package directory `*documentation*`, folder `*scripting_solutions*`.

## 6.2 A gentle introduction to programming

## 6.2.1 Introduction

This section includes examples in Python programming language. You may want to try out some of the examples that come with the description. In order to do this, execute the Python IDLE - available in the Application-Menu of your operating system, and write the examples after the prompt “>>>”.

The Python IDLE logo

To get information about the IDLE, get access to [the IDLE documentation](#)

Writing any program consists of writing statements so using a programming language. A *statement* is often known as a line of code that can be one of:

- comment,
- documentation,
- assignment,
- conditions,
- iterations, etc.

Lines of code are grouped in *blocks*. Depending on the programming language, blocks delimited by brackets, braces or by the indentation.

Each language has its own syntax to write these lines and the user has to follow strictly this syntax for the program to be able to interpret the program. However, the amount of freedom the user has to use capital letters, whitespace and so on is very high. Recommendations for Python language are available in the [PEP8 - Style Guide for Python Code](#).

## 6.2.2 Variables: Assignment and Typing

A *variable* is a name to give to a piece of memory with some information inside. Assignment is then the action of setting a variable to a value. The equal sign (=) is used to assign values to variables.

```
>>>a = 1
>>>b = 1.0
>>>c = "c"
>>>hello = "Hello world!"
>>>vrai = True
```

In the previous example, a, b, c, hello and vrai are variables, a = 1 is a declaration.

Variable declarations and print in the Python IDLE

Assignments to variables with Python language can be performed with the following operators:

```
>>> a = 10 # simple assignment operator
>>> a += 2 # add and assignment operator, so a is 12
>>> a -= 7 # minus and assignment, so a is 5
>>> a *= 20 # multiply and assignment, so a is 100
>>> a /= 10 # divide and assignment, so a is 10
>>> a # verify the value of a...
10
```

### 6.2.3 Basic Operators

Basic operators are used to manipulate variables. The following is the list of operators that can be used with Python, i.e. equal (assignment), plus, minus, multiply, divide:

```
>>> a = 10
>>> b = 20 # assignment
>>> a + b # addition
>>> a - b # subtraction
>>> a * b # multiplication
>>> a / b # division
```

### 6.2.4 Data types

The variables are of a data-type. For example, the declarations `a=1` and `a=1.0` are respectively assigning an integer and a real number. In Python, the command `type` allows to get the type of a variable, like in the following:

```
>>> type(a)
<type 'int'>
>>> type(b)
<type 'float'>
>>> type(c)
<type 'str'>
>>> type(cc)
<type 'unicode'>
>>> type(vrai)
<type 'bool'>
```

Here is a list of some fundamental data types, and their characteristics:

- *str* String of characters
- *unicode* Unicode string of characters
- *int* Integer in range from -2147483648 to 2147483647
- *bool* Boolean value, can take value `True` (=1) or `False` (=0)
- *float* Floating point number (max 15 digits)

Python is assigning data types dynamically. As a consequence, the result of the sum between an `int` and a `float` is a `float`. The next examples illustrate that the type of the variables have to be carefully managed.

```
>>> a = 10
>>> a += 0.
>>> a
10.0
>>> a += True
>>> a
11.0
```

```

>>> a += "a"
Traceback (most recent call last):
  File "<input>", line 1, in <module>
TypeError: unsupported operand type(s) for +=: 'float' and 'str'
>>> a = "a"
>>> a *= 5
>>> a
'aaaaa'

```

The type of a variable can be explicitly changed. This is called a “cast”:

```

>>> a = 10
>>> b = 2
>>> a/b
5
>>> float(a) / float(b)
5.0
>>> a = 1
>>> b = 1
>>> str(a) + str(b)
'11'

```

Complex data types are often used to store variables sharing the same properties like a list of numbers, and so on. Common types in languages are lists/arrays and dictionaries. The following is the assignment of a list with name `fruits`, then the assignment of a sub-part of the list to the `to_buy` list:

```

>>> fruits = ['apples', 'tomatoes', 'peers', 'bananas', 'lemons']
>>> to_buy = fruits[1:3]
>>> to_buy
['tomatoes', 'peers']

```

### 6.2.5 Conditions

Conditions aim to test whether a statement is True or False. The statement of the condition can include a variable, or be a variable and is written with operators. The following shows examples of conditions/comparisons in Python. Notice that the comparison of variables of a different data-type is possible (but not recommended!).

```

>>> var = 100
>>> if var == 100:
...     print("Value of expression is 100.")
...
Value of expression is 100.

>>> if var == "100":
...     print("This message won't be printed out.")
...

```

Conditions can be expressed in a more complex way like:

```
>>> if a == b:
...     print('a and b are equals')
... elif a > b:
...     print('a is greater than b')
... else:
...     print('b is greater than a')
```

The simple operators for comparisons are summarized in the next examples:

```
>>> a == b    # check if equals
>>> a != b    # check if different
>>> a > b     # check if a is greater than b
>>> a >= b    # check if a is greater or equal to b
>>> a < b     # check if a is lesser than b
>>> a <= b    # check if a is lesser or equal to b
```

It is also possible to use the following operators:

- **and**: called “Logical AND operator”. If both the operands are true then the condition becomes true.
- **or**: called “Logical OR operator”. If any of the two operands are non zero then the condition becomes true.
- **not** called “Logical NOT operator”. Use to reverses the logical state of its operand. If a condition is true then Logical NOT operator will make false.
- **in**: evaluates to true if it finds a variable in the specified sequence and false otherwise.

```
>>> if a == "apples" and b == "peers":
...     print("You need to buy fruits.")

>>> if a == "apples" or b == "apples":
...     print("You already have bought apples.")

>>> if "tomatoas" not in to_buy:
...     print("You don't have to buy tomatoes.")
```

### 6.2.6 Loops

The `for` loop statement iterates over the items of any sequence. The next Python lines of code print items of a list on the screen:

```
>>> to_buy = ['fruits', 'viande', 'poisson', 'oeufs']
>>> for item in to_buy:
...     print(item)
...
fruits
viande
poisson
oeufs
```

A `while` loop statement repeatedly executes a target statement as long as a given condition returns `True`. The following example prints exactly the same result as the previous one:

```
>>> to_buy = ['fruits', 'viande', 'poisson', 'oeufs']
>>> i = 0
>>> while i < len(to_buy):
...     print(to_buy[i])
...     i += 1
...
fruits
viande
poisson
oeufs
```

### 6.2.7 Dictionaries

A dictionary is a very useful data type. It consists of pairs of keys and their corresponding values.

```
>>> fruits = dict()
>>> fruits['apples'] = 3
>>> fruits['peers'] = 5
>>> fruits['tomatoas'] = 1
```

`fruits['apples']` is a way to get the value - i.e. 3, of the apple key. However, an error is sent if the key is unknown, like `fruits[bananas]`. Alternatively, the `get` function can be used, like `fruits.get("bananas", 0)` that returns 0 instead of an error.

The next example is showing how use a simple dictionary:

```
>>> for key in fruits:
...     value = fruits.get(key, 0)
...     if value < 3:
...         print("You have to buy new {s}.".format(key))
...
You have to buy new tomatoes.
```

To learn more about data structures and how to manage them, get access to [the Python documentation](#)

## 6.3 Scripting with Python

This section describes how to create simple Python lines of code in separated files commonly called *scripts*, and run them. Some practical exercises, appropriate to the content of each action, are proposed and test exercises are suggested at the end of the section.

To practice, you have first to create a new folder in your computer - on your Desktop for example; with name “pythonscripts” for example, and to execute the python IDLE.

For an advanced use of Python, the installation of a dedicated IDE is very useful. SPPAS is developed with PyCharm: See [the PyCharm Help webpage](#)

### 6.3.1 Comments and documentation

Comments are not required by the program to work. But comments are necessary! Comments are expected to be appropriate, useful, relevant, adequate and always reasonable.

```
# This script is doing this and that.
# It is under the terms of a license.
# and I can continue to write what I want after the # symbol
# except that it's not the right way to tell the story of my life
```

The documentation of a program complements the comments. Both are not sharing the same goal: comments are used in all kind of programs but documentation is appended to comments for the biggest programs and/or projects. Documentation is automatically extracted and formatted thanks to dedicated tools. Documentation is required for sharing the program. See the [Docstring Conventions](#) for details. Documentation must follow a convention like for example the markup language [reST - reStructured Text](#). Both conventions are used into SPPAS API, programs and scripts.

### 6.3.2 Getting started with scripting in Python

In the IDLE, create a new empty file either by clicking on “File” menu, then “New File”, or with the shortcut “CTRL”+N.

Copy the following line of code in this newly created file:

```
1 print("Hello world!")
```

Hello world! in a Python script

Then, save the file in the “pythonscripts” folder. By convention, Python source files end with a *.py* extension, and so the name `01_helloworld.py` could be fine.

To execute the program, you can do one of:

- with the mouse: Click on the Menu “Run”, then “Run module”
- with the keyboard: Press F5

The expected output is as follow:

Output of the first script

A better practice while writing scripts is to describe by who, what and why this script was done. A nifty trick is to create a skeleton for any future script that will be written. Such ready-to-use script is available in the SPPAS package with the name `skeleton.py`.

### 6.3.3 Blocks

Blocks in Python are created from the indentation. Tab and spaces can be used but using spaces is recommended.

```
>>> if a == 3:
...     # this is a block using 4 spaces for indentation
...     print("a is 3")
```

### 6.3.4 Functions

#### Simple function

A function does something: it starts with its definition then is followed by its lines of code in a block.

Here is an example of function:

```
27 def print_vowels():
28     """ Print the list of French vowels on the screen. """
29
30     vowels = ['a', 'e', 'E', 'i', 'o', 'u', 'y', '@', '2', '9', 'a~', 'o~', 'U~']
31     print("List of French vowels:")
32     for v in vowels:
33         print(v)
```

What the `print_vowels()` function is doing? This function declares a list with name `vowels`. Each item of the list is a string representing a vowel in French encoded in X-SAMPA. Of course, this list can be overridden with any other set of strings. The next line prints a message. Then, a loop prints each item of the list.

At this stage, if a script with this function is executed, it will do... nothing! Actually, the function is created, but it must be invoked in the main function to be interpreted by Python. The `main` is as follow:

```
34 if __name__ == '__main__':
35     print_vowels()
```

*Practice:* create a copy of the file `skeleton.py`, then make a function to print “Hello World!”. (solution: `ex01_hello_world.py`).

*Practice:* Create a function to print plosives and call it in the main function (solution: `ex02_functions.py`).

Output of the second script

One can also create a function to print glides, another one to print affricates, and so on. Hum... this sounds a little bit fastidious!

#### Function with parameters

Rather than writing the same lines of code with only a minor difference over and over, we can declare *parameters* to the function to *make it more generic*. Notice that the number of parameters of a function is not limited!

In the example, we can replace the `print_vowels()` function and the `print_plosives()` function by a single function `print_list(mylist)` where `mylist` can be any list containing strings or characters. If the list contains other typed-variables like numerical values, they must be converted to string to be printed out. This can result in the following function:

```

27     def print_list(mylist, message=" -"):
28         """ Print a list on the screen.
29
30         :param mylist: (list) the list to print
31         :param message: (string) an optional message to print before each element
32
33         """
34         for item in mylist:
35             print("{:s} {:s}".format(message, item))

```

### Function return values

Functions are used to do a specific job and the result of the function can be captured by the program. In the following example, the function would return a boolean value, i.e. True if the given string has no character.

```

27     def is_empty(mystr):
28         """ Return True if mystr is empty. """
29
30         return len(mystr.strip()) == 0

```

*Practice:* Add this function in a new script and try to print various lists (solution: ex03\_functions.py)

Expected output of the 3rd script

### 6.3.5 Reading/Writing files

#### Reading data from a file

Now, we'll try to get data from a file. Create a new empty file with the following lines - and add as many lines as you want; then, save it with the name "phonemes.csv" by using UTF-8 encoding:

```

occlusives ; b ; b
occlusives ; d ; d
fricatives ; f ; f
liquids ; l ; l
nasals ; m ; m
nasals ; n ; n
occlusives ; p ; p
glides ; w ; w
vowels ; a ; a
vowels ; e ; e

```

The following statements are typical statements used to read the content of a file. The first parameter of the open function is the name of the file, including the path (relative or absolute); and the second argument is the opening mode ('r' is the default value, used for reading).

*Practice:* Add these lines of code in a new script and try it (solution: ex04\_reading\_simple.py)

```

21 fp = open("phonemes.csv", 'r')
22 for line in fp:
23     # do something with the line stored in variable l
24     print(line.strip())
25 f.close()

```

The following is a solution with the ability to deal with various file encodings, thanks to the `codecs` library:

```

21 def read_file(filename):
22     """ Get the content of file.
23
24     :param filename: (string) Name of the file to read, including path.
25     :returns: List of lines
26
27     """
28     with codecs.open(filename, 'r', encoding="utf8") as fp:
29         return fp.readlines()

```

In the previous code, the `codecs.open` functions got 3 parameters: the name of the file, the mode to open, and the encoding. The `readlines()` function gets each line of the file and store it into a list.

*Practice:* Write a script to print the content of a file (solution: `ex05_reading_file.py`)

Notice that Python `os` module provides useful methods to perform file-processing operations, such as renaming and deleting. See Python documentation for details: <https://docs.python.org/3.8/>

### Writing data to a file

Writing a file requires to open it in a writing mode:

- 'w' is the mode to write data; it will erase any existing file;
- 'a' is the mode to append data in an existing file.

A file can be opened in an encoding and saved in another one. This could be useful to write a script to convert the encoding of a set of files. The following could help to create such script:

```

10 # getting all files of a given folder:
11 path = 'C:\Users\Me\data'
12 dirs = os.listdir( path )
13
14
15 # Converting the encoding of a file:
16 file_stream = codecs.open(file_location, 'r', file_encoding)
17 file_output = codecs.open(file_location+'utf8', 'w', 'utf-8')
18
19 for line in file_stream:
20     file_output.write(line)

```

### 6.3.6 Python tutorials

Here is a list of web sites with tutorials, from the easiest to the most complete:

1. [Learn Python, by DataCamp](#)
2. [Tutorial Points](#)
3. [The Python documentation](#)

### 6.3.7 Exercises to practice

Exercise 1: How many vowels are in a list of phonemes? (solution: `ex06_list.py`)

Exercise 2: Write a X-SAMPA to IPA converter. (solution: `ex07_dict.py`)

Exercise 3: Compare 2 sets of data using NLP techniques (Zipf law, Tf.Idf) (solution: `ex08_counter.py`)

## 6.4 `anndata`, an API to manage annotated data

### 6.4.1 Overview

We are now going to write Python scripts using the `anndata` API included in SPPAS. This API is useful to read/write and manipulate files annotated from various annotation tools like SPPAS, Praat or Elan.

First of all, it is important to understand the data structure included into the API to be able to use it efficiently.

### 6.4.2 Why developing a new API?

In the Linguistics field, multimodal annotations contain information ranging from general linguistic to domain specific information. Some are annotated with automatic tools, and some are manually annotated. In annotation tools, annotated data are mainly represented in the form of “tiers” or “tracks” of annotations. Tiers are mostly series of intervals defined by:

- two time points to represent the beginning and the end of the interval;
- a label to represent the annotation itself.

Of course, depending on the annotation tool, the internal data representation and the file formats are different. In Praat, tiers can be represented either by a single point in time (such tiers are named `PointTiers`) or two (`IntervalTiers`) In Elan, points are not supported; but contrariwise to Praat, unlabelled intervals are not represented nor saved.

The `anndata` API was designed to be able to manipulate all data in the same way, regardless of the file type. It supports to merge data and annotations from a wide range of heterogeneous data sources.

### 6.4.3 The “anndata” API class diagram

After opening/loading a file, its content is stored in a `sppasTranscription` object. A `sppasTranscription` has a name, and a list of `sppasTier` objects. Tiers can't share the same name, the list of tiers can be empty, and a hierarchy between tiers can be defined. Actually, subdivision relations can be established between tiers. For example, a tier with phonemes is a subdivision reference for syllables, or for tokens; and tokens are a subdivision reference for the orthographic transcription in IPUs. Such subdivisions can be of two categories: alignment or association.

A `sppasTier` object has a name, and a list of `sppasAnnotation` objects. It can also be associated to a controlled vocabulary, or a media.

All these objects contain a set of meta-data.

An annotation is made of 2 objects:

- a `sppasLocation` object,
- a list of `sppasLabel` objects.

A `sppasLabel` object is representing the “content” of the annotation. It is a list of `sppasTag` each one associated to a score.

A `sppasLocation` is representing where this annotation occurs in the media. Then, a `sppasLocation` is made of a list of localization each one associated with a score. A localization is one of:

- a `sppasPoint` object; or
- a `sppasInterval` object, which is made of 2 `sppasPoint` objects; or
- an un-used `sppasDisjoint` object which is a list of `sppasInterval`.

API class diagram

#### Label representation

Each annotation holds a serie of 0..N labels. A label is also an object made of a list of `sppasTag`, each one with a score. A `sppasTag` is mainly represented in the form of a string, freely written by the annotator but it can also be a boolean (True/False), an integer, a floating number, a point with (x, y) coordinates with an optional radius or a rectangle with (x, y, w, h) coordinates with an optional radius value.

#### Location representation

In the *anndata* API, a `sppasPoint` is considered *as an imprecise value*. It is possible to characterize a point in a space immediately allowing its vagueness by using:

- a midpoint value (center) of the point;
- a radius value.

Representation of a `sppasPoint`

### Example

The screenshot below shows an example of multimodal annotated data, imported from 3 different annotation tools. Each `sppasPoint` is represented by a vertical dark-blue line with a gradient color to refer to the radius value.

In the screenshot the following radius values were assigned:

- 0ms for prosody,
- 5ms for phonetics, discourse and syntax
- 40ms for gestures.

Example of multimodal data

## 6.5 Creating scripts with `anndata`

### 6.5.1 Preparing the data

To practice, you have first to create a new folder in your computer - on your Desktop for example; with name “sppasscripts” for example, and to execute the python IDLE.

Open a File Explorer window and go to the SPPAS folder location. Then, copy the `sppas` directory into the newly created “sppasscripts” folder. Then, go to the solution directory and copy/paste the files `skeleton-sppas.py` and `F_F_B003-P9-merge.TextGrid` into your “sppasscripts” folder. Then, open the skeleton script with the python IDLE and execute it. It will do... nothing! But now, you are ready to do something with the API of SPPAS!

When using the API, if something forbidden is attempted, the object will raise an Exception which means the program will stop.

### 6.5.2 Read/Write annotated files

We are being to Open/Read an annotated file of any format (XRA, TextGrid, Elan, ...) and store it into a `sppasTranscription` object instance. Then, the object will be saved into another file.

```
1     # Create a parser object then parse the input file.
2     parser = sppasRW(input_filename)
3     trs = parser.read()
4
5     # Save the sppasTranscription object into a file.
6     parser.set_filename(output_filename)
7     parser.write(trs)
```

Only these few lines of code are required to convert a file from a format to another one! The appropriate parsing system is extracted from the extension of file name.

To get the list of accepted extensions that the API can read, just use `parser.extensions_in()`. The list of accepted extensions that the API can write is given by `parser.extensions_out()`.

*Practice:* Write a script to convert a TextGrid file into CSV (solution: `ex10_read_write.py`)

### 6.5.3 Manipulating a `sppasTranscription` object

The most useful functions to manage the tiers of a `sppasTranscription` object are:

- `create_tier()` to create an empty tier and to append it,
- `append(tier)` to add a tier into the `sppasTranscription`,
- `pop(index)` to remove a tier of the `sppasTranscription`,
- `find(name, case_sensitive=True)` to find a tier from its name.

Below is a piece of code to browse through the list of tiers:

```

15     for tier in trs:
16         # below, do something with the tier:
17         print(tier.get_name())
18     # Search for a specific tier,
19     # None is returned if not found.
20     phons_tier = trs.find("PhonAlign")

```

*Practice:* Write a script to select a set of tiers of a file and save them into a new file (solution: `ex11_transcription.py`).

### 6.5.4 Manipulating a `sppasTier` object

A tier is made of a name, a list of annotations, and optionally a controlled vocabulary and a media. To get the name of a tier, or to fix a new name, the easier way is to use `tier.get_name()`. The following block of code allows to get a tier and change its name.

```

43     # Get the first tier, with index=0
44     tier = trs[0]
45     print(tier.get_name())
46     tier.set_name("NewName")
47     print(tier.get_name())

```

The most useful functions to manage annotations of a `sppasTier` object are:

- `create_annotation(location, labels)` to create and add a new annotation
- `append(annotation)` to add a new annotation at the end of the list
- `add(annotation)` to add a new annotation
- `pop(index)` to delete the annotation of a given index
- `remove(begin, end)` to remove annotations of a given localization range
- `is_disjoint()`, `is_interval()`, `is_point()` to know the type of location
- `is_string()`, `is_int()`, `is_float()`, `is_bool()`, `is_fuzzypoint()`, `is_fuzzyrect()` to know the type of labels
- `find(begin, end)` to get annotations in a given localization range
- `get_first_point()`, `get_last_point()` to get respectively the point with the lowest or highest localization
- `set_radius(radius)` to fix the same vagueness value to each localization point

*Practice:* Write a script to open an annotated file and print information about tiers (solution: `ex12_tiers_info.py`)

### 6.5.5 Manipulating a `sppasAnnotation` object

An annotation is a container for a location and optionally a list of labels. It can be used to manage the labels and tags with the following methods:

- `is_labelled()` returns `True` if at least a `sppasTag` exists and is not `None`
- `append_label(label)` to add a label at the end of the list of labels
- `get_labels_best_tag()` returns a list with the best tag of each label
- `add_tag(tag, score, label_index)` to add a tag into a label
- `remove_tag(tag, label_index)` to remove a tag of a label

An annotation object can also be copied with the method `copy()`. The location, the labels and the metadata are all copied; and the 'id' of the returned annotation is then the same. It is expected that each annotation of a tier as its own 'id', but the API doesn't check this.

*Practice:* Write a script to print information about annotations of a tier (solution: `ex13_tiers_info.py`)

### 6.5.6 Search in annotations: Filters

#### Overview

This section focuses on the problem of *searching and retrieving* data from annotated corpora.

The filter implementation can only be used together with the `sppasTier()` class. The idea is that each `sppasTier()` can contain a set of filters, that each reduce the full list of annotations to a subset.

SPPAS filtering system proposes 2 main axis to filter such data:

- with a boolean function based either on the content, the duration or on the time of annotations,
- with a relation function between annotation locations of 2 tiers.

A set of filters can be created and combined to get the expected result. To be able to apply filters to a tier, some data must be loaded first. First, a new `sppasTranscription()` has to be created when loading a file. Then, the tier(s) to apply filters on must be fixed. Finally, if the input file was NOT an XRA, it is widely recommended to fix a radius value before using a relation filter.

```
f = sppasFilter(tier)
```

When a filter is applied, it returns an instance of `sppasAnnSet` which is the set of annotations matching with the request. It also contains a 'value' which is the list of functions that are truly matching for each annotation. Finally, `sppasAnnSet` objects can be combined with the operators '!' and '&', and expected to a `sppasTier` instance.

### Filter on the tag content

The following matching names are proposed to select annotations:

- ‘exact’: means that a tag is valid if it strictly corresponds to the expected pattern;
- ‘contains’ means that a tag is valid if it contains the expected pattern;
- ‘startswith’ means that a tag is valid if it starts with the expected pattern;
- ‘endswith’ means that a tag is valid if it ends with the expected pattern.
- ‘regexp’ to define regular expressions.

All these matches can be reversed, to represent does not exactly match, does not contain, does not start with or does not end with. Moreover, they can be case-insensitive by adding ‘i’ at the beginning like ‘iexact’, etc. The full list of tag matching functions is obtained by invoking `sppasTagCompare().get_function_names()`.

The next examples illustrate how to work with such pattern matching filter. In this example, `f1` is a filter used to get all phonemes with the exact label ‘a’. On the other side, `f2` is a filter that ignores all phonemes matching with ‘a’ (mentioned by the symbol ‘~’) with a case insensitive comparison (iexact means insensitive-exact).

```
tier = trs.find("PhonAlign")
f = sppasFilter(tier)
ann_set_a = f.tag(exact='a')
ann_set_aA = f.tag(iexact='a')
```

The next example illustrates how to write a complex request. Notice that `r1` is equal to `r2`, but getting `r1` is faster:

```
tier = trs.find("TokensAlign")
f = sppasFilter(tier)
r1 = f.tag(startswith="pa", not_endswith='a', logic_bool="and")
r2 = f.tag(startswith="pa") & f.tag(not_endswith='a')
```

With this notation in hands, it is easy to formulate queries like for example: *Extract words starting by “ch” or “sh”*:

```
result = f.tag(startswith="ch") | f.tag(startswith="sh")
```

*Practice::* Write a script to extract phonemes /a/ then phonemes /a/, /e/, /A/ and /E/. (solution: `ex15_annotation_label_filter.py`).

### Filter on the duration

The following matching names are proposed to select annotations:

- ‘lt’ means that the duration of the annotation is lower than the given one;
- ‘le’ means that the duration of the annotation is lower or equal than the given one;
- ‘gt’ means that the duration of the annotation is greater than the given one;

- ‘ge’ means that the duration of the annotation is greater or equal than the given one;
- ‘eq’ means that the duration of the annotation is equal to the given one;
- ‘ne’ means that the duration of the annotation is not equal to the given one.

The full list of duration matching functions is obtained by invoking `sppasDurationCompare().get_function_names`

Next example shows how to get phonemes during between 30 ms and 70 ms. Notice that `r1` and `r2` are equals!

```
tier = trs.find("PhonAlign")
f = sppasFilter(tier)
r1 = f.dur(ge=0.03) & f.dur(le=0.07)
r2 = f.dur(ge=0.03, le=0.07, logic_bool="and")
```

*Practice:* Extract phonemes ‘a’ or ‘e’ during more than 100ms (solution: `ex16_annotation_dur_filter.py`).

### Filter on position in time

The following matching names are proposed to select annotations:

- `rangefrom` allows to fix the begin time value,
- `rangeto` allows to fix the end time value.

Next example allows to extract phonemes ‘a’ of the 5 first seconds:

```
tier = trs.find("PhonAlign")
f = sppasFilter(tier)
result = f.tag(exact='a') & f.loc(rangefrom=0., rangeto=5., logic_bool="and")
```

### Creating a relation function

Relations between annotations is crucial if we want to extract multimodal data. The aim here is to select intervals of a tier depending on what is represented in another tier.

James Allen, in 1983, proposed an algebraic framework named Interval Algebra (IA), for qualitative reasoning with time intervals where the binary relationship between a pair of intervals is represented by a subset of 13 atomic relation, that are:

- distinct because no pair of definite intervals can be related by more than one of the relationships;
- exhaustive because any pair of definite intervals are described by one of the relations;
- qualitative (rather than quantitative) because no numeric time spans are considered.

These relations and the operations on them form “Allen’s Interval Algebra”.

Pujari, Kumari and Sattar proposed INDU in 1999: an Interval & Duration network. They extended the IA to model qualitative information about intervals and durations in a single binary constraint network. Duration relations are: greater, lower and equal. INDU comprises of 25 basic relations between a pair of two intervals.

`anndata` implements the 13 Allen interval relations: before, after, meets, met by, overlaps, overlapped by, starts, started by, finishes, finished by, contains, during and equals; and it also contains the relations proposed in the INDU model. The full list of matching functions is obtained by invoking `sppasIntervalCompare().get_function_names()`.

Moreover, in the implementation of `anndata`, some functions accept options:

- before and after accept a `max_delay` value,
- overlaps and overlappedby accept an `overlap_min` value and a boolean `percent` which defines whether the value is absolute or is a percentage.

The next example returns monosyllabic tokens and tokens that are overlapping a syllable (only if the overlap is during more than 40 ms):

```
tier = trs.find("TokensAlign")
other_tier = trs.find("Syllables")
f = sppasFilter(tier)
f.rel(other_tier, "equals", "overlaps", "overlappedby", min_overlap=0.04)
```

Below is another example of implementing a request. *Which syllables stretch across 2 words?*

```
1 # Get tiers from a sppasTranscription object
2 tier_syll = trs.find("Syllables")
3 tier_toks = trs.find("TokensAlign")
4 f = sppasFilter(tier_syll)
5
6 # Apply the filter with the relation function
7 ann_set = f.rel(tier_toks, "overlaps", "overlappedby")
8
9 # To convert filtered data into a tier:
10 tier = ann_set.to_tier("SyllStretch")
```

*Practice 1:* Create a script to get tokens followed by a silence. (solution: `ex17_annotations_relation_filter1.py`).

*Practice 2:* Create a script to get tokens preceded by OR followed by a silence. (solution: `ex17_annotations_relation_filter2.py`).

*Practice 3:* Create a script to get tokens preceded by AND followed by a silence. (solution: `ex17_annotations_relation_filter3.py`).

## 6.6 More with SPPAS...

In addition to `anndata`, SPPAS contains several other API. They are all free and open source Python libraries, with a documentation and a set of tests.

Among others:

- *audiodata* to manage digital audio data: load, get information, extract channels, re-sample, search for silences, mix channels, etc.
- *calculus* to perform some math on data, including descriptive statistics.
- *resources* to access and manage linguistic resources like lexicons, dictionaries, etc.



## References

### 7.1 References

#### 7.1.1 How to cite SPPAS?

For a general citation of SPPAS, use the reference below. A PDF version of this publication is available in the folder `documentation` of the package.

*Brigitte Bigi (2015). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech.* In “the Phonetician” - International Society of Phonetic Sciences, ISSN 0741-6164, Number 111-112 / 2015-I-II, pages 54-69.

For a specific feature of SPPAS, like the automatic syllabification, the forced-alignment, or for the request system, please refer to the related specific publication. The full list of references and the corresponding PDF files are available on [the author website](http://www.lpl-aix.fr/~bigi/publications.html), at the following URL: <http://www.lpl-aix.fr/~bigi/publications.html>

#### 7.1.2 SPPAS software description

*Brigitte Bigi (2015). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech.* In “the Phonetician” - International Society of Phonetic Sciences, ISSN 0741-6164, Number 111-112 / 2015-I-II, pages 54-69.

*Brigitte Bigi, Daniel Hirst (2013). What’s new in SPPAS 1.5?, Tools and Resources for the Analysis of Speech Prosody,* Aix-en-Provence, France, pp. 62-65.

*Brigitte Bigi (2012). SPPAS: a tool for the phonetic segmentations of Speech,* The eight international conference on Language Resources and Evaluation, Istanbul (Turkey), pages 1748-1755, ISBN 978-2-9517408-7-7.

*Brigitte Bigi, Daniel Hirst (2012) SPeech Phonetization Alignment and Syllabification (SPPAS): a tool for the automatic analysis of speech prosody,* Speech Prosody, Tongji University Press, ISBN 978-7-5608-4869-3, pages 19-22, Shanghai (China).

### 7.1.3 About linguistic resources

*Brigitte Bigi, Abiola S. Oyelere, Bernard Caron* (2020). **Resources for Automated Speech Segmentation of the African Language Naija (Nigerian Pidgin)**. Human Language Technologies Challenges for Computer Science and Linguistics, LNAI 12598, pp. 164-173, Springer, Heidelberg.

*Mélanie Lancien, Marie-Hélène Côté, Brigitte Bigi* (2020). **Developing Resources for Automated Speech Processing of Quebec French**. In Proceedings of The 12th Language Resources and Evaluation Conference, pp. 5323–5328, Marseille, France.

*Brigitte Bigi, Bernard Caron, Abiola S. Oyelere* (2017). **Developing Resources for Automated Speech Processing of the African Language Naija (Nigerian Pidgin)**. In 8th Language and Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, pp. 441-445, Poznań, Poland.

### 7.1.4 About Search for IPUs

*Brigitte Bigi and Béatrice Priego-Valverde* (accepted). **The automatic search for sounding segments of SPPAS: application to Cheese! corpus** Human Language Technologies Challenges for Computer Science and Linguistics, LNAI, Springer, Heidelberg.

*Brigitte Bigi, Béatrice Priego-Valverde* (2019). **Search for Inter-Pausal Units: application to Cheese! corpus**. In 9th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, pp. 289-293, Poznań, Poland.

### 7.1.5 About Text Normalization

Roxana Fung, Brigitte Bigi\* (2015). **Automatic Word Segmentation for Spoken Cantonese**, The Oriental Chapter of COCODA (International Committee for the Co-ordination and Standardization of Speech Databases and Assessment Techniques).

*Brigitte Bigi* (2014). **A Multilingual Text Normalization Approach**, Human Language Technologies Challenges for Computer Science and Linguistics LNAI 8387, Springer, Heidelberg. ISBN: 978-3-319-14120-6. Pages 515-526.

*Brigitte Bigi* (2011). **A Multilingual Text Normalization Approach**, 2nd Less-Resourced Languages workshop, 5th Language & Technology Conference, Poznan (Poland).

### 7.1.6 About Phonetization

*Brigitte Bigi* (2016). **A phonetization approach for the forced-alignment task in SPPAS**, Human Language Technologies Challenges for Computer Science and Linguistics, LNAI 9561, , pp. 515–526, Springer, Heidelberg. ISBN: 978-3-319-14120-6.

*Brigitte Bigi* (2013). **A phonetization approach for the forced-alignment task**, 3rd Less-Resourced Languages workshop, 6th Language & Technology Conference, Poznan (Poland).

*Brigitte Bigi, Pauline Péri, Roxane Bertrand* (2012). **Orthographic Transcription: Which Enrichment is required for Phonetization?**, Language Resources and Evaluation Conference, Istanbul (Turkey), pages 1756-1763, ISBN 978-2-9517408-7-7.

### 7.1.7 About Forced-Alignment

*Brigitte Bigi, Christine Meunier* (2018). **Automatic speech segmentation of spontaneous speech**. Revista de Estudos da Linguagem. International Thematic Issue: Speech Segmentation. Editors: Tommaso Raso, Heliana Mello, Plinio Barbosa, Volume 26, number 4, pages 1489-1530, e-ISSN 2237-2083.

*Brigitte Bigi, Christine Meunier* (2018). **euh, rire et bruits en parole spontanée : application à l'alignement forcé**, Actes des 32èmes Journées d'études sur la Parole, Aix-en-Provence, France.

*Brigitte Bigi* (2014). **The SPPAS participation to Evalita 2014**, Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 and the Fourth International Workshop EVALITA 2014, Pisa (Italy). Editors R. Basili, A. Lenci, B. Magnini. ISBN 978-886741-472-7. Volume 2. Pages 127-130.

*Brigitte Bigi* (2014). **Automatic Speech Segmentation of French: Corpus Adaptation**. 2nd Asian Pacific Corpus Linguistics Conference, p. 32, Hong Kong.

*Brigitte Bigi* (2012). **The SPPAS participation to Evalita 2011**, Working Notes of EVALITA 2011, Rome (Italy), ISSN: 2240-5186.

### 7.1.8 About Syllabification

*Brigitte Bigi, Katarzyna Klessa* (2015). **Automatic Syllabification of Polish**, 7th Language and Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, pp. 262–266, Poznan, Poland.

*Brigitte Bigi, Caterina Petrone, Leonardo Lancia* (2014). **Automatic Syllabification of Italian: adaptation from French**. Laboratory Approaches to Romance Phonology VII, Aix-en-Provence (France).

*Brigitte Bigi, Caterina Petrone* (2014). **A generic tool for the automatic syllabification of Italian**, Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 and the Fourth International Workshop EVALITA 2014, Pisa (Italy). Editors R. Basili, A. Lenci, B. Magnini. ISBN 978-886741-472-7. Volume 1. Pages 73-77.

*Brigitte Bigi, Christine Meunier, Irina Nesterenko, Roxane Bertrand* (2010). **Automatic detection of syllable boundaries in spontaneous speech**, Language Resource and Evaluation Conference, pages 3285-3292, La Valetta, Malte.

### 7.1.9 About Repetitions

*Brigitte Bigi, Roxane Bertrand, Mathilde Guardiola* (2014). **Automatic detection of other-repetition occurrences: application to French conversational speech**, 9th International conference on Language Resources and Evaluation (LREC), Reykjavik (Iceland), pp. 2648-2652. ISBN: 978-2-9517408-8-4.

### 7.1.10 About analyses tools

*Brigitte Bigi* (2019). **Filtering multi-levels annotated data**. In 9th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, pp. 13-14, Poznań, Poland.

*Brigitte Bigi, Jorane Saubesty* (2015). **Searching and retrieving multi-levels annotated data**, Proceedings of Gesture and Speech in Interaction, Nantes (France).

### 7.1.11 About the API

*Brigitte Bigi* (2018). **Annotation representation and file conversion tool**. Contributi del Centro Linceo Interdisciplinare 'Beniamino Segre' ISSN 0394-0705), 137, pp. 99-116.

*Brigitte Bigi, Tatsuya Watanabe, Laurent Prévot* (2014). **Representing Multimodal Linguistics Annotated Data**, 9th International conference on Language Resources and Evaluation (LREC), Reykjavik (Iceland), pages 3386-3392. ISBN: 978-2-9517408-8-4.

### 7.1.12 Related references

#### Some results of analyses

*Christine Meunier, Brigitte Bigi* (to be published). **Propriétés phonétiques et lexicales dans les variations temporelles des phonèmes en parole conversationnelle**, Studii de lingvistică, SL11: « Variation phonétique en parole continue », coord.: Yaru Wu, Martine Adda-Decker, Lori Lamel, Ioana Vasilescu

*Christine Meunier, Morgane Peirola, Brigitte Bigi* (2020). **La mobilisation du tractus vocal est-elle variable selon les langues en parole spontanée ?**, In Journées d'Etudes sur la Parole, 33e édition, pp. 433-441, Nancy, France.

*Béatrice Priego-Valverde, Brigitte Bigi, Salvatore Attardo, Lucy Pickering, Elisa Gironzetti* (2018). **Is smiling during humor so obvious? A cross-cultural comparison of smiling behavior in humorous sequences in American English and French interactions**. Intercultural Pragmatics. 10/2018; 15(4):563-591.

*Brigitte Bigi, Roxane Bertrand* (2016). **Laughter in French Spontaneous Conversational Dialogs**, Tenth International Conference on Language Resources and Evaluation (LREC 2016), pp. 2168-2174, Portoroz, Slovenia.

*Laurent Prevot, Jan Gorisch, Roxane Bertrand, Emilien Gorene and Brigitte Bigi* (2015). **A SIP of CoFee: A Sample of Interesting Productions of Conversational Feedback**, 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue, Prague (Czech Republic), pp. 149–153.

*Brigitte Bigi, Katarzyna Klessa, Laurianne Georgeton, Christine Meunier* (2015). **A syllable-based analysis of speech temporal organization: a comparison between speaking styles in dysarthric and healthy populations**. Interspeech2015, Dresden (Germany), pages 2977-2981.

*Marion Tellier, Gale Stam, Brigitte Bigi* (2013). **Gesturing While Pausing In Conversation: Self-oriented Or Partner-oriented?**, TIGER, Tillburg.

*Laurent Prévot, Brigitte Bigi, Roxane Bertrand* (2013). **A quantitative view of feedback lexical markers in conversational French**, 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue, Metz, France.

*Marion Tellier, Gale Stam, Brigitte Bigi* (2012). **Same speech, different gestures?**, 5th International Society for Gesture Studies (ISGS), Lund, Sweden.

### Some corpora

*Birgit Rauchbauer, Youssef Hmamouche, Brigitte Bigi, Laurent Prévot, Magalie Ochs, Thierry Chaminade* (2020). **Multimodal Corpus of Bidirectional Conversation of Human-human and Human-robot Interaction during fMRI Scanning**. In Proceedings of The 12th Language Resources and Evaluation Conference, pp. 668–675, Marseille, France.

*Béatrice Priego-Valverde, Brigitte Bigi, Mary Amoyal* (2020). **“Cheese!?”: a Corpus of Face-to-face French Interactions. A Case Study for Analyzing Smiling and Conversational Humor**. In Proceedings of The 12th Language Resources and Evaluation Conference, pp. 467–475, Marseille, France.

*Christine Meunier, Cécile Fougeron, Corinne Fredouille, Brigitte Bigi, Lise Crevier-Buchman, Elisabeth Delais-Roussarie, Laurianne Georgeton, Alain Ghio, Imed Laaridh, Thierry Legou, Claire Pillot-Loiseau, Gilles Pouchoulin* (2016). **The TYPALOC Corpus: A Collection of Various Dysarthric Speech Recordings in Read and Spontaneous Styles**, Tenth International Conference on Language Resources and Evaluation (LREC 2016), pp. 4658-4665, Portoroz, Slovenia.

*Jan Gorish, Corine Astésano, Ellen Gurman Bard, Brigitte Bigi, Laurent Prévot* (2014). **Aix Map Task corpus: The French multimodal corpus of task-oriented dialogue**, 9th International conference on Language Resources and Evaluation (LREC), Reykjavik (Iceland), pages 2648-2652. ISBN: 978-2-9517408-8-4.

*Daniel Hirst, Brigitte Bigi, Hyongsil Cho, Hongwei Ding, Sophie Herment, Ting Wang* (2013). **Building OMProDat: an open multilingual prosodic database**, Proceedings of Tools and Resources for the Analysis of Speech Prosody, Aix-en-Provence, France, Eds B. Bigi and D. Hirst, ISBN: 978-2-7466-6443-2, pp. 11-14.

*Ellen Gurman Bard, Corine Astésano, Alice Turk, Mariapaola D'imperio, Noel Nguyen, Laurent Prévot, Brigitte Bigi* (2013). **Aix MapTask: A (rather) new French resource for prosodic and discourse studies**, Proceedings of Tools and Resources for the Analysis of Speech Prosody, Aix-en-Provence, France, Eds B. Bigi and D. Hirst, ISBN: 978-2-7466-6443-2, pp. 15-19.

*Sophie Herment, Anastasia Loukina, Anne Tortel, Daniel Hirst, Brigitte Bigi* (2012). **AixOx, a multi-layered learners corpus: automatic annotation**, Proceedings of international conference on corpus linguistics, Jaèn (Spain).

#### **Other tools re-implemented into SPPAS**

*Maciej Karpinski, Katarzyna Klessa* (2018). **\*\* Methods, Tools and Techniques for Multimodal Analysis of Accommodation in Intercultural Communication\*\***, CMST 24(1) 29–41 (2018), DOI:10.12921/cmst.2018.0000006

*Dafydd Gibbon* (2013). **TGA: a web tool for Time Group Analysis**, Tools and Resources for the Analysis of Speech Prosody, Aix-en-Provence, France, pp. 66-69.

*D.-J. Hirst* (2011). **The analysis by synthesis of speech melody: from data to models**, Journal of Speech Sciences, vol. 1(1), pages 55-83.

*Daniel Hirst and Robert Espesser* (1993). **Automatic modelling of fundamental frequency using a quadratic spline function**. Travaux de l'Institut de Phonétique d'Aix. vol. 15, pages 71-85.

## SPPAS Release notes

### 8.1 The early versions

Versions 1.0 to 1.3 was made only of `tcsh` and `gawk` scripts. It was developed under Linux system and was efficiently tested under Windows with Cygwin.

#### 8.1.1 Version 1.0

(2011, 9th March)

The only feature was that it was able to perform speech segmentation of English read speech.

#### 8.1.2 Version 1.1

(2011, 7th June)

This was mainly the debug of the previous version and some code re-organization and cleaning.

#### 8.1.3 Version 1.2

(2011, 23th July)

The support of English, French and Italian was added: a lexicon, a pronunciation dictionary and an acoustic model of each language was created. Three annotations were implemented: Tokenization, Phonetization, Alignment.

The basis for a *multi-lingual methodology* was already there.

#### 8.1.4 Version 1.3

(2011, 12th December)

This is a transitional version, from the scripts language to Python programming language. The main fixes and improvements were:

**Development:**

- MacOS support
- bugs corrected in many scripts
- check.csh -> check.py
- sppas.csh -> sppas.py
- GUI: zenity -> pyGtk
- sox limited to the use of re-sampling audio signal
- a python library to manage wav files is added.
- IPU's segmentation automatic annotation

**Resources:**

- Italian dictionary improved
- Italian acoustic model changed (triphones trained from map-task dialogues)
- French acoustic model changed

But the program still haven't a name and isn't distributed.

## 8.2 The birth of SPPAS

### 8.2.1 SPPAS 1.4.0

(2012, 14th June)

It's the official birth of SPPAS: the software has a name, a license, a GUI, a web-page and is freely distributed to the community.

The source code is only based on Python 2.7 language and the GUI is based on WxPython 2.8.x or 2.9.x (on MacOS, wxpython must be 32 bits). SPPAS requires also `sox` and `julius` software to be installed.

**Automatic annotations:**

- add Momel: Momel (MOdelling MELody) is an algorithm developed by Daniel Hirst and Robert Espesser for the analysis and synthesis of intonation patterns. Momel needs a .hz files with pitch values: ASCII file with one value each 10ms. This Momel implementation can be used directly, with a large set of options: `> python $SPPAS/scripts/lib/momel.py -h`  
or in SPPAS, using the GUI, or in text-mode (only with default options): `> $SPPAS/sppas.py -i INPUT -momel`
- add the first version of INTSINT, proposed by Daniel Hirst. INTSINT is an acronym for INternational Transcription System for INTonation. INTSINT codes the intonation of an utterance by means of an alphabet of 8 discrete symbols constituting a surface phonological representation of the intonation.

**Packaging:**

- merge AUTHORS, VERSION and README files in the README file

- create packages in the lib directory (2012-03-20)
- ipu-segmentation in python (2012-02-16)
- improve ipu segmentation algorithm (2012-05-24)
- phonetization in python (2012-02-27), with an unknown word phonetization.
- alignment entirely in python (2012-03-29)
- syllabification implemented with python, and tool/syllabify.py
- add a simple terminal controller (not needed for Unix, just for “DOS”)
- Momel can read PitchTier (from Praat) files
- Momel can deal with long sounds (not only one IPU)
- manage source files, comments, exceptions...
- SPPAS works as a single instance by mean of a lock file.
- SPPAS can be installed in a directory containing spaces and can deal with file names with spaces (bug fixed: 2012-06-11)

**GUI:**

- GUI with the wxpython library (2012-03-27)
- Manage a list of selected files
- Add a “File information”
- Add a “Wav player”
- Fix options to each annotation step with the menu
- Open the log file after each process

**Resources:**

- French dictionary is using UTF-8 (instead of iso8859-1 in previous versions)
- French dictionary is based on X-SAMPA phone set
- Chinese: work with chinese characters instead of pinyin.

**Known Bugs:**

- The wav player can not play wav files if the filename contains ‘#’.
- The first line of the Italian dictionary must be changed to “# [#] #”.

**8.2.2 SPPAS 1.4.1**

(2012, 13th July)

**Resources:**

- Updated English acoustic model (from voxforge, 2012-06-25)
- English acoustic models converted to X-SAMPA

**Automatic annotations:**

- IPU's Segmentation annotation performs a simple silence detection if no transcription is available (the volume is automatically adjusted)

- A specific language can be selected for each annotation depending on available resources
- Updated transcription conventions:
  - truncated words: a ‘-’ at the end of the token string (an ex- example)
  - liaisons: the ‘letter’ between ‘=’ (an =n= example)
  - noises: \* (only for FR and IT)
  - short pauses: + (a + example)
  - silences: # (a # example)

**GUI:**

- Create (systematically) a merged annotations TextGrid file.

**Development:**

- Package’s management

**8.2.3 SPPAS 1.4.2**

(2012, 2nd August)

**GUI:**

- add a panel with a Transcription editor
- IPU segmentation can split a wav into tracks
- IPU segmentation can fix a shift value to boundaries
- IPU segmentation: min-volume option is removed (because the min-volume value is automatically adjusted)
- “File Information” button adds some tier manipulation tools: cut/copy/paste/rename/duplicate/move/preview/filter

**Known bug:**

- The filter frame is not working under Windows XP.

**8.2.4 SPPAS 1.4.3**

(2012, 10th October)

This is primarily a bug-fix release. The author is addressing many thanks to all users who send their comments!

**GUI:**

- Frames and Dialog design is more uniform.
- Users preferences changed. Themes and colors introduced.
- Help is available.

**Automatic annotations:**

- Bug fixed for phonetization/alignement if the input transcription contains series of silence intervals or series of speech intervals. Previous versions was supposing a **strict IPU**s input transcription.
- Tokenization is done.

**Development:**

- Code cleaning in the package wxGUI.
- Debug...

**8.2.5 SPPAS 1.4.4**

(2012, 6th December)

**GUI:**

- add information/options when “Request” a file
- debug Request/Filter (for Windows systems)

**Automatic annotations:**

- New Italian acoustic model
- New Chinese acoustic model, and some minor changes in the dictionary

**Development:**

- “.trs” files support (transcriptions from Transcriber)
- debug (alignment, tokenization)
- add “.lab” files export (HTK format)

**Known bugs:**

- Alignment: it fails under Windows, if `julius` is not installed properly.
- Syllabification: the last syllable of each file is “broken”.
- Alignment error if unknown words during phonetization.

**8.2.6 SPPAS 1.4.5**

(2013, 15th January)

**Development:**

- Correct a few bugs of the previous version (phonetization, alignment, syllabification)
- “.eaf” files support (transcriptions from Elan software)
- add script `tierfilter.py`
- add script `tiercombine.py`

**Automatic annotations:**

- Experimental version of Vietnamese
- add Tokenization as a specific annotation level
- add Phonetization of PinYin

**GUI:**

- Request/Filter a tier: multiple patterns filtering
- Request: add a “New File” button

**8.2.7 SPPAS 1.4.6**

(2013, 12th February)

**GUI:**

- Improved Request/Filter a tier:
  - add new modes: not contains, not starts with, not ends with
  - add time constraints: minimum duration, maximum duration, meets
  - multiple modes selection (replace radio buttons by check buttons)
- Add Requests/Stats to obtain basic statistics
- Requests Copy/Cut/Paste/Duplicate/Save debugged

**Automatic annotations:**

- IPU segmentation: can take a “Name” tier to fix names of tracks (if the “Split into tracks” option is checked)

**8.2.8 SPPAS 1.4.7**

(2013, 25th March)

**Development:**

- re-organization of lib, except for the wxGUI library.
- Cancel the sppas.lock file creation when SPPAS is running.
- Requests/Filter: “Starts search at...” and “Ends search at...”
- Requests/Filter: remove negative search (because of a bug...). Replaced by a “reverse” option.

**Resources:**

- add experimental version of Taiwanese automatic segmentation (from romanized transcriptions)

**Annotations:**

- New version of INTSINT, based on the algorithm proposed in (Hirst 2011) and implemented in the last version of Momel/INTSINT Praat plugin!

**8.2.9 SPPAS 1.4.8**

(2013, 30th May)

**Development:**

- reorganization of the GUI code. SPPAS is made of:
  - automatic annotations;
  - 3 components (wavplayer, transcriber, requests).
- sppas.py removed. sppas.bat created (for Windows).
- Input/Output library entirely changed

**GUI:**

- components (wavplayer, transcriber, requests) in separate frames
- new design
- Help and documentation changed, expanded and improved

**8.2.10 SPPAS 1.4.9**

(2013, 3rd July)

SPPAS has a new and more colored logo!

**Development:**

- bug correction:
  - Bug fixed in IPU segmentation with option “Split into tracks”
  - Bug fixed in Momel with some rare files
  - Bug fixed to create a merged file
  - Bug fixed in align and IPU seg.
  - Bug fixed in Transcriber
- library organization revised

**GUI:**

- Add an export button to the File List Panel
- Migrate wav infos from the Requests component to the Wav player component
- Volume control removed in the Wav Player component
- Save As improved in the Requests component

### 8.2.11 SPPAS 1.5.0

(2013, 2nd August)

#### Development:

- bug correction

#### Annotation:

- Tokenization: TOE support finalized

#### GUI:

- Transcribe: debug of IPU player, for MacOS

#### Resources:

- New French acoustic model. Attention: phoneset changed.
- New French dictionary: use the new phoneset!
- New Taiwanese acoustic model. Attention: phoneset changed: accept some chinese...
- New Taiwanese dictionary: use the new phoneset + some chinese syllables added.
- Vietnamese removed: due to the lack of data, the model can't be improved.

### 8.2.12 SPPAS 1.5.1

(2013, 29th August)

#### Development:

- bug correction in annotations
- help system debugged

#### Annotation:

- Phonetization of unknown words improved

#### Resources:

- French dict modified

### 8.2.13 SPPAS 1.5.2

(2013, 27th September)

All resources are moved into the “resources” directory.

**Development:**

- bug correction in annotations and GUI

**Resources:**

- French dict modified
- New resources for the tokenization

**Components:**

- “Statistics”: an advanced component to estimate (and save!) descriptive statistics on multiple annotated files

**8.2.14 SPPAS 1.5.3**

(2013, 25th October)

**Resources:**

- Add Spanish support

**Components:**

- improved Statistics component
- Add a “Filter” component, first version with a basic GUI

**GUI:**

- Help updated.

**8.2.15 SPPAS 1.5.4**

(2013, 3rd December)

**Components:**

- Wav Player changed.
- Information and Requests removed. Replaced by Data Roamer.
- Transcriber removed. Replaced by IPU Transcriber.
- Statistics updated.
- Filter updated.

**Annotations:**

- Add Repetitions (detection of sources only)

Notice that this is the first stable release.

**8.2.16 SPPAS 1.5.5**

(2013, 23th December)

**Development:**

- improved annotationdata (add methods: Find, Index; add uncertain label; debug radius).

**Components:**

- debug

**GUI:**

- Tips at start-up
- New theme: Christmast

**8.2.17 SPPAS 1.5.6**

(2014, 28th January)

**Development:**

- improved annotationdata (add methods: Near, Search).
- Package cleaning!

**Components:**

- debug

**Resources:**

- New Italian dictionary (including gemination)

**8.2.18 SPPAS 1.5.7**

(2014, 18th February)

**Development:**

- Plugins manager added.

**Resources:**

- Japanese support, thanks to the resources available on the Julius website.

**8.2.19 SPPAS 1.5.8**

(2014, 18th March)

**Development:**

- annotationdata: more flexibility while adding annotations, add sub-divisions, export csv modified, doc-strings, import/export in a native format (xra, version 1.0).

**Documentation:**

- add a PDF file of slides: “SPPAS for dummies”

**8.2.20 SPPAS 1.5.9**

(2014, 15th April)

**Components:**

- new: DataViewer, experimental version

**Annotations:**

- syllabification rules: accept 2 types of vowels (V and W)
- syllabification: faster!

**Development:**

- Export Elan
- Import/Export xra. XRA is the native format of SPPAS annotation files. See etc/xra for details.

**Resources:**

- New French acoustic model

### 8.2.21 SPPAS 1.6.0

(2014, 22th May)

#### **Package:**

- Rename folder “Doc” to “documentation”
- A documentation is created
- SPPAS-for-dummies updated

#### **Development:**

- helpsystem removed
- GUI: package re-organization, re-implementation.
- Alignment: choice of the aligner (julius, hvite or basic)

#### **Resources:**

- new acoustic model: FR-Read and FR
- new acoustic model: ZH
- new acoustic model: TW
- new acoustic model: SP

### 8.2.22 SPPAS 1.6.1

(2014, 26th September)

#### **Development:**

- bug correction (export, syllabification, alignment)
- DataViewer, major bugs corrected.

#### **Resources:**

- Italian: new pronunciation dictionary and new acoustic model
- add resources for Catalan

#### **GUI:**

- add a new Export button

### 8.2.23 SPPAS 1.6.2

(2014, 21th October)

**Resources:**

- language names changed. They are now corresponding to the international standard ISO639-3. See <http://www-01.sil.org/iso639-3/>
- Mandarin Chinese dictionary changed.
- Catalan dictionary changed.

**8.2.24 SPPAS 1.6.3**

(2014, 2nd November)

**Resources:**

- Add resources for Cantonese

**Documentation:**

- It's now (more or less) finished!

**Development:**

- Support of Elan improved (add Controlled vocabulary)

**GUI:**

- Change the organization of the main frame: annotations above components

This version is known to be a stable release.

## 8.3 The beginnings of SPPAS

**8.3.1 SPPAS 1.6.4**

(2014, 5th December)

From this version, SPPAS requires wxpython to be updated to version 3.0, particularly for MacOS users, and they need to install the 64 bits version. It is recommended to Windows users to install Python 2.7 and wxpython in 32bits.

**Development:**

- Package re-organized!
- Phonetization of unknown words improved
- Support of upper/lower of the extension of speech files (wav, WAV)
- Tokenization of languages with dictionaries in upper case (eng, ita): bug fixed.

- Creates systematically a dump file of resources for a faster load at the next use
- Read TextGrid files exported by Elan: bug fixed.
- `sppas.command` checks the system and run either in 32 or 64 bits (MacOS)

**Components:**

- IPUscribe replaces IPUTranscriber mainly for the support of large files: tested with a file of 1 hour speech (143 Go) and 800 IPUs.
- SndRoamer replaces WavPlayer
- Dataroamer has also a new version

**8.3.2 SPPAS 1.6.5**

(2014, 17th December)

This is primarily a bug-fix release.

**Development:**

- all programs in “bin” and “scripts” sub-directories were revised and tested, or removed.

**Annotation:**

- Tokenization: code cleaning and re-organisation.

**GUI:**

- Procedure outcome report: print a warning message in the log file if no file is matching the expected extension

**8.3.3 SPPAS 1.6.6**

(2015, 19th January)

Web site host has changed: <http://sldr.org/sldr00800/preview/>

Documentation completed and updated. Now, only the documentation of all the components is missing.

**Annotations:**

- log messages more explicit and status printed with intuitive colors.
- management of input/output file format re-done: now easier for the user.

### Development:

- package architecture revised: mainly “sppasgui” and “components” merged in “wxgui”, and many other changes.
- thread removed in automatic annotation process.
- debug of alignment: if too short units.
- radius value properly fixed in most of the automatic annotations.

### GUI:

- GUI is more homogeneous and pretty (hope!)
- Show the date in the status bar
- New Settings frame:
  - 4 icon themes available
  - Choice of foreground and background colours
  - Choice of the font
  - Choice of the input/output file format of annotations
- New in Help menu:
  - access to the project homepage
  - access to the online documentation
- New Feedback window
- New Help browser
- Add Keyboard shortcuts:
  - ALT+F4 to exit,
  - CTRL+A to add files in FLP
  - SHIFT+CTRL+A to add a directory in FLP
  - Del to remove selected files of the FLP
  - SHIFT+Del to erase selected files of the FLP
  - CTRL+C to copy files
  - CTRL+E to export files
  - F5 to refresh the FLP
  - F1 to open the help browser
  - F2 to open the “About” frame

### Components:

- GUI design unified for DataRoamer, SndPlayer, IPUscribe and SppasEdit
- New Tier Preview frame (still under development)
- SndPlayer print information about a sound file with colors:
  - Green: the information corresponds to the SPPAS requirements

- Orange: the information does not exactly corresponds to the requirements however SPPAS is able to deal with (after conversion)
- Red: SPPAS does not support. It must be converted before using it!

### 8.3.4 SPPAS-1.6.7

(2015, 16th February)

#### **Automatic Annotations:**

- By default, tokenization produces only one tier. Check an option to get TokensStd and TokensFaked, in case of EOT.
- radius value properly fixed in most of the automatic annotations.

#### **GUI:**

- Tested and debugged on MacOS (version 10.9.5, with wxpython 3.0.2)

#### **Development:**

- Tier hierarchy partly implemented: TimeAlignement and TimeAssociation are two “links” that can be fixed between tiers.

#### **Annotations:**

- Add Polish support

### 8.3.5 SPPAS-1.6.8

(2015, 9th April)

#### **Resources:**

- new French acoustic model
- new English acoustic model (VoxForge nightly build of March, 15th, 2015)
- add phoneset mapping tables

#### **Development:**

- Add a phoneme mapping in models, to allow both the dictionary to include real X-SAMPA symbols and the acoustic model to be compatible with Hvite requirements (only ASCII).
- annotationdata bug correction with min and max values
- IPU's Segmentation:

- bug correction when split into tracks with a tier “Name”
- add the ipu number in speech segments if silence/speech segmentation
- Self-repetitions debug in finding the repetition interval

**GUI:**

- DataRoamer: “New” button debugged
- DataRoamer: Add a button “Radius” to adjust manually the vagueness of each boundary of a tier

## 8.4 The development phase

### 8.4.1 SPPAS-1.6.9

(2015, 14th May)

The installation of dependencies is simplified: `sox` is unnecessary. Python 2.7.x, WxPython and Julius are the only remaining dependencies.

**Development:**

- package `annotationdata.filter` updated to support last changes in `annotationdata`: multiple labels and numerical labels.
- package `annotationdata.io`:
  - `praat.py` newly created. Support of `TextGrid`, `PitchTier` and `IntensityTier` files completely re-written
  - `htk.py` newly created to support `.lab` and `.mlf` files
  - `sclite.py` newly created to support `.stm` and `.ctm` files
  - `signaix.py` newly created to support `.hz` files
  - SPPAS native format XRA upgraded to version 1.1 to support improvements of the library.
- package `annotationdata`:
  - updated object `Transcription` to support more/different input data
  - updated hierarchy
  - updated meta-data
- package `signal`: partially re-written. As a consequence, `sox` won't be neither used, but the file conversion (if required) is slower.

**GUI:**

- `DataFilter` component partially re-written to facilitate its use
- Preview frame modified

### 8.4.2 SPPAS-1.7.0

(2015, 3th July)

#### Development:

- package annotationdata.io:
  - add support of subtitles: sub, srt
  - elan.py created to replace eaf.py: allows a full support of Elan annotations
  - transcriber.py created to replace trs.py for a better support of Transcriber files
  - anvil.py allows to import anvil files
- package annotationdata:
  - updated hierarchy (simplified)
  - updated meta-data
- package signal:
  - support of audio files re-written: can open/save wav, aiff and au files
  - add a lot of possibilities to manage one channel of an audio file

#### GUI:

- DataFilter component finalized: can deal with alternative labels, and typed labels (string, number, boolean)
- Statistics component fully re-written to facilitate its use
- SndRoamer displays more properties of audio files (min, mean and max added)

#### Annotations:

- IPUs Segmentation produces 2 tiers if a transcription is given: the IPUs segmentation itself, and the Transcription time-aligned at the IPUs level.

### 8.4.3 SPPAS-1.7.1

(2015, 5th August)

#### Development:

- package re-organization:
  - package signal is now standalone
  - package resources is now standalone
  - package presenters created
- updated statistics estimations
- package annotationdata:

- add Media object
  - add CtrlVocab object
  - add inheritance of MetaObject for Annotation
- package annotationdata.io:
  - full debug of all file formats
  - add comments and documentation
  - add also some tests
  - add Media/CtrlVocab in some file formats
  - add Annotation Pro support of antx files (in API)

**Components:**

- add Time Group Analyser (TGA) in Statistics
- add Kappa estimation on labels of 2 tiers with the same number of intervals

**Annotations:**

- New version of XRA: 1.2
- Make XRA the default input/output file format

**8.4.4 SPPAS-1.7.2**

(2015, 3th September)

**Development:**

- updated XRA reader/writer to solve problems with upper/lower cases of elements
- updated Elan reader to be compatible with format 2.8
- updated Praat reader/writer for single/double quotes
- support of AnnotationPro antx files in the GUI
- add the julius score in PhonAlign annotation (can be seen only if XRA)
- remove tk dependency

**8.4.5 SPPAS-1.7.3**

(2015, 9th October)

**Resources:**

- New word-based vocab for Cantonese
- New word-based dict for Cantonese
- New phoneme-based acoustic model for Cantonese

**Development:**

- Descriptive statistics debugged for detailed panels.

**8.4.6 SPPAS-1.7.4**

(2015, 6th November)

**Resources:**

- Add a vocab for Portuguese
- Add a dict for Portuguese
- Add an acoustic model for Portuguese. It has to be noticed that it was constructed from French/Spanish/Italian models, not trained from data.

**Samples:**

- Add Portuguese
- Change some samples in various languages

**Development:**

- Debug of the Tokenizer.

**8.4.7 SPPAS-1.7.5**

(2015, 11th December)

**Development:**

- Add vagueness to the annotation duration estimation
- Bug correction while saving the procedure outcome report in the GUI

**GUI:**

- Changed all pictures to remove the problem with colorset of some of them
- Add a christmast theme for the pictures of tips
- IPUscribe improvements:
  - add keyboard shortcuts to play/pause/stop the sound
  - add a new button to auto-play the sound

#### 8.4.8 SPPAS-1.7.6

(2016, 28th January)

Web site host has changed: <http://www.sppas.org/>

##### Development:

- IPU segmentation:
  - bug correction with option “split into tracks”
  - new option to add or not add the IPU index of each IPU into the transcription tier
- Tokenization:
  - bug correction in input tier selection
  - bug correction for replacements

##### Resources:

- add a vocabulary of Korean
- add an acoustic model for Korean (made from the English and the Taiwanese ones).

##### GUI:

- DataRoamer: bug correction of “Duplicate”

##### Others:

- Add a sample in Korean
- Updated references to cite SPPAS in publications

#### 8.4.9 SPPAS-1.7.7

(2016, 30th March)

##### Resources:

- Correction of errors in most of the acoustics models, for /@@/ and /dummy/ models
- New vocabulary and pronunciation dictionary for Mandarin Chinese.

##### GUI:

- Dialogs are customized
- DataRoamer debug and changes: save buttons moved in the main toolbar.
- HelpSystem is (welcome) back!

**Development:**

- Add the API and a script to train acoustic models with HTK.
- Add Kullback-Leibler distance estimator.
- Add a script to compare segmentation of 2 tiers.
- Classes of the package resources are debugged, improved and extended.
- Alignment: bug correction if input starts by an empty interval.
- `sppas.command` modified for MacOS-X (use `python` if `python2` not available)
- GUIs in `bin` directory are updated (test of `python`, `wxpython`, and so on).

**8.4.10 SPPAS-1.7.8**

(2016, 5th May)

**Resources:**

- Add a pronunciation mapping table `eng-fra.map`: to be used for the speech segmentation for French speakers reading an English text.

**Development:**

- Phonetization is extended: it can take into account a mapping table of phones to generate new pronunciation variants.
- Phonetization: use of minus instead of dots to separate phones as recommended in X-SAMPA standard.
- Alignment is restructured and extended: it can take into account two acoustics models and mix them.
- Filter is extended: Relations can be based on a delay between the intervals
- `annotationdata`: add Xtrans reader (extension: `.tdf`)
- Code cleaning in several packages.

**GUI:**

- Automatic annotations: more explicit log messages.

**8.4.11 SPPAS-1.7.9**

(2016, 3th June)

**GUI:**

- Some debug in `SppasEdit`
- `SndPlayer` renamed `AudioRoamer` because new functionalities were added:
  - See detailed information about each channel of an audio file
  - Can extract/save a channel or a fragment of channel
  - Can modify the frame rate and the sample width of a channel

- Can add a bias on amplitude values of a channel
- Can multiply amplitude values of a channel
- Can remove the offset of amplitude values of a channel

**Automatic annotations:**

- IPUs Segmentation fully re-implemented.
  - Silence/speech segmentation improved for both quality and fastness
  - Package code cleaning and re-organization.

## 8.5 The stabilization phase

### 8.5.1 SPPAS-1.8.0

(2016, 30th August)

**GUI:**

- Design fully revisited and tested under Linux Mint, Windows 10 and MacOS 10.9

**Development:**

- SLM package created: can estimate a statistical language model (without smooth method) on a small corpus

**Automatic annotations:**

- Add a diagnosis of files
- Tokenize extended: applied also on alternative labels
- Phonetize extended: applied also on alternative labels
- Alignment code cleaning and partly re-implemented
- Add “Chunk alignment”
- Use of a .ini file to configure each annotation instead of a sppas.conf file

### 8.5.2 SPPAS-1.8.1

(2016, 28th November)

A few tutorials are available on the web site.

**Automatic annotations:**

- Align: an ActivityDuration tier can be optionally added.
- Support of 3-columns tab-delimited files with .txt extension. It allows the compatibility with Audacity Label track files.
- Acoustic models training validated.

**Resources:**

- Catalan: new pronunciation dictionary and new acoustic model.

**8.5.3 SPPAS-1.8.2**

(2017, 18th January)

**Analysis:**

- debug of DataFilter

**Resources:**

- French vocabulary and dictionary updated

**Development:**

- new plugins package with a new plugin manager
- GUI integration of this new plugins system
- some unittest appended and all existing ones updated
- annotationdata.io renamed annotationdata.aio
- docstrings of some packages converted from epytext to reST syntax

**GUI:**

- DataStats, DataFilter and DataRoamer toolbars don't scroll anymore
- Themes management changed.
- Main font is managed by the Themes.

**8.5.4 SPPAS-1.8.3**

(2017, 10th March)

**Development:**

- Elan reader highly improved (faster reader).
- updated plugins

**8.5.5 SPPAS 1.8.4**

(2017, 10th April)

**Development:**

- Elan writer modified: create a time slot for each localization.

### 8.5.6 SPPAS 1.8.5

(2017, 20th April)

#### Development:

- Vizualizer renamed Visualizer
- AudioRoamer: bug with an icon corrected
- New Phonedit mrk format support.
- Updated AnnotationPro Antx reader/writer

### 8.5.7 SPPAS 1.8.6

(2017, 19th June)

#### Resources:

- Polish dictionary and acoustic model updated.

### 8.5.8 SPPAS 1.9.0

(2017, 28th July)

#### Programming:

- Relative imports used in the standard way for Python
- PEP 8 code style (except for wxgui/annotationdata)
- PEP 257 reST code documentation style (except for wxqui/annotationdata)
- Unittests:
  - existing tests verified, improved, extended
  - new tests added
  - tests migrated into the packages
- Compatibility for both Python 2.7 and Python > 3.2:
  - makeunicode.py contains functions and classes to deal with strings
  - utils, term, structs, resources, presenters, plugins, calculus packages: migration is done
- Exceptions are separately managed
- Introduction of a system for the internationalization of the messages. Done in English and French for the packages: audiodata, calculus, plugins, resources, structs, term, utils
- Package re-organization:

- new package “models”, with acm and slm
  - utils
  - resources
  - ...
- meta.py is replacing sp\_glob.py
- new scripts: tieralightophon.py, dictmerge.py
- new bin: pluginbuild.py
- Robustness to read malformed HTK-ASCII pronunciation dictionaries.
- Bug corrected in the management of pronunciation variants
- re-structured package TextNormalization
- re-structured package Repetitions
- new version of the plugins: updated and debugged.

**Resources:**

- Add support of Naija language (pcm)
- English-French mapping table updated

**Annotations:**

- Tokenizer renamed into Text Normalization:
  - a lot of debug mainly for English language, and punctuation managements
  - new option: can output a customized tier.
- Repetitions:
  - some debug

**Communication:**

- Add a description document of the orthographic transcription convention in the package.
- Web page updated, new tutorials available
- Documentation updated
- Better information about the licenses

**8.5.9 SPPAS 1.9.1**

(2017, 1st September)

**Programming:**

- Bug correction in the ELAN reader.
- Bug correction with quotation marks of the Praat writer.

**Resources:**

- Add an acoustic model for English, including laughter and noises.

**8.5.10 SPPAS 1.9.2**

(2017, 6th October)

**Programming:**

- Bug correction in the diagnosis of audio files.
- Package “anndata” continued, in the scope of replacing “annotationdata”.
- Acoustic model training procedure debugged and improved, code cleaned, scripts updated, etc.

**8.5.11 SPPAS 1.9.3**

(2017, 18th October)

**Programming:**

- Critical bug correction in Alignment: error when loading the tiedlist.
- Correction of TextNormalization of broken words: “-” is now not removed.

**Resources:**

- Italian pronunciation dictionary updated.
- French vocabulary updated: list of compound words revised.

**Known bugs:**

- Spanish alignment does not work: corrupted acoustic model

**8.5.12 SPPAS 1.9.4**

(2018, 15th January)

**Programming:**

- Script to train acoustic models updated
- Script to evaluate alignments updated
- Bug corrections of the search tier system

**Resources:**

- Spanish:
  - new pronunciation dictionary
  - new acoustic model

**8.5.13 SPPAS 1.9.5**

(2018, 26th April)

**Programming:**

- I/O development of the package “anndata”.
- Reading and writing of annotated files is now based on the new ‘anndata’ package instead of the old ‘annotationdata’ package. This was a big issue and it allows a better support of annotated files (less bugs and the amount of lost information is drastically reduced).

**8.5.14 SPPAS 1.9.6**

(2018, 25th May)

**Resources:**

- New: support of German language for Text Normalization (except num2letter), Phonetization and Alignment.
- French: new acoustic model with an updated phoneset and a better accuracy.
- Laugh and noise are uniformly represented in all languages.
- English: context-independent model is the default, instead of context-dependent model, for Alignment.

**Development**

- debug of the Transcriber file reader.
- debug in the management of the hierarchy in file readers/writers.
- Resources package: add compatibility with PLS (W3C xml) pronunciation dictionaries

**Known bugs:**

- Syllabification of French: the configuration file is not correct.

**8.5.15 SPPAS 1.9.7**

(2018, 23th July)

## Development

- new filter system in `anndata`
- IPUs segmentation: debug in fixing automatically the threshold for the search of silences (can't be negative!)
- Annotations: improved messages for the procedure outcome report
- Annotations: Text normalization is based on `anndata` API instead of `annotationdata`
  - the result is now a sequence of individual tokens (each token is a `sppasLabel()`), instead of a single string separating tokens with space.
- Annotations: Phonetization is based on `anndata` API instead of `annotationdata`
  - the result is now a sequence of `sppasLabel()` with alternative tags: each tag corresponds to a pronunciation variant
- Annotations: Syllabification re-programmed and now based on `anndata` API
  - only sequences of phonemes are syllabified (i.e. no silence, no laugh nor noise in the result)
  - `O` class name changed into `P` class name
  - do not generate the tier “structures” tier anymore
  - generating the “classes” tier is optional
- Annotations: INTSINT is based on `anndata` API instead of `annotationdata`
- IPUScriber is based on `anndata` API instead of `annotationdata`

## Resources:

- French syllabification rules updated to be compliant with the new phoneset (used since version 1.9.6)

## Documentation

- updated chapter 6: scripting with Python and SPPAS. It is now based on `anndata` API instead of `annotationdata`.

### 8.5.16 SPPAS 1.9.8

(2018, 06th September)

e-mail contact is changed to:

- *contact@sppas.org* for general questions of users
- *develop@sppas.org* for developer questions or for a bug alert

## Development

- bug correction for GMT with a negative value, like GMT-4.
- most of the scripts are based on `anndata` API instead of `annotationdata`
- dependency to markdown removed
- new solution to work with global variables with the new package `config` and the new classes like `sppasGlobalSettings`, `sppasPathSettings`, ...

- translation management moved in package `config`
- better way to work with imports
- increased pep8 and pep257 compatibility (should be continued)
- `anndata` is used in `DataRoamer`, `Visualizer`, `DataFilter`
- new script `trsshift.py` to shift the transcription of a delay in time

## Documentation

The API documentation is based on Sphinx <http://www.sphinx-doc.org>.

### 8.5.17 SPPAS 1.9.9

(2018, 23th October)

## Development

- package for the ‘alignment’ is fully re-structured (re-implemented partially), new unittests are added, and compatibility with python 3, pep8 and pep257.
- new GUI based on python3 + phoenix: it prints an information message in case SPPAS is launched with these versions.
- package ‘presenters’ updated: unittests, python 3, pep8 and pep257.

## Plugins

- `sampa2ipa` included to the SPPAS package
- `audiosegmenter` included to the SPPAS package

## Annotations

- IPU Segmentation removed. It is replaced by:
  1. Search for IPUs: to find silences/tracks from an audio file
  2. Fill in IPUs: to find silences/tracks from an audio file and its transcription
  3. AudioSegmenter plugin: to segment audio files into several tracks
- `annotation.py`: bug correction with the default output extension
- bug correction when the phonetization was unknown

## Resources

- add a file ‘monophones’ into each model, for the compatibility with HVite.
- errors of the acoustic models of “spa” and “nan” corrected.

### 8.5.18 SPPAS 2.0

(2019, 4th January)

The main change of this release is that the package `annotationdata` has been removed, so that all packages are using `anndata` instead. One of the most important consequence of this change is that all packages, except the `wxGUI`, are compatible with both Python 2.7 and Python 3.4+.

#### Annotations

- Alignment: bug correction if the last unit is empty.
- Self-Repetitions and Other-Repetitions split into two different packages
- Un-used package ‘Chunks’ removed
- Better management of annotations: `sppasBaseAnnot`, `sppasParam` and `sppasManager` revised.

#### UI

- All programs in `bin` folder updated. A ‘manual mode’ and an ‘auto mode’ are available in the annotations to fix input/output, and a `-log` option is proposed to save a report.
- Plugins: `marsatagplugin` added and debugged.

### 8.5.19 SPPAS 2.1

(2019, 28th February)

#### Development

- Praat TextGrid: it’s now allowed to use the symbol ‘=’ in annotation labels.
- Use of ‘json’ files instead of ‘ini’ files to configure SPPAS, the annotations and the plugins.
- New package ‘analysis’ with the filter and statistic systems.

#### Annotations

- Search for IPU: the program has been improved and evaluated.
- Alignment: priority is given to standard tokens (if EOT).
- Alignment: problem of whitespace in filenames solved.

#### UI

- Plugins: new plugin to remove un-translated IPU and to re-index the IPU
- extended internationalization of the messages (for English and French)

### 8.5.20 SPPAS 2.2

(2019, 09th May)

sppas.bat and sppas.command, the two main ways to launch the GUI, were fully re-written in order to search for 'pythonw' command first. It results in the following advantages: - it increases the compatibility with MacOS systems; - it allows to not display the dark frame of 'python' under Windows.

#### Development

- New package 'files' for a further use.
- Increased compatibility of file formats: can read alternative tags with {} system, whitespace vs CR, etc.

#### Plugins

- Debug of the plugin to clean IPUs.

#### Annotations

- Search for IPUs: a special threshold value is set when the audio recording has a very low median value (ie very low volume values)
- Normalization/Phonetization/Alignment/Syllabification can work without audio
- Search for IPUs and Fill in IPUs: problem with upper/lower extension solved

#### GUI

- "Delete" button of the "File Explorer": files are no longer definitively deleted. Instead, they are moved into an hidden folder with name '.trash' in the package of SPPAS.
- wxFrame to display log messages.
- Annotations in GUI: better compatibility to annotate written texts

#### Resources

- fra.dict: corrected 6 errors with a non-existing 'eu' sound

### 8.5.21 SPPAS 2.3

(2019, 25th June)

#### Development

- JSON configuration files of sppas modified.
- package "files" is used: the annotations manager is using a workspace instead of a list of files.

- the GUI based on wx4 is managing properly the workspaces and the annotations: a page with “Files” is displayed and allows to “add/remove/delete” files in the list, “import from/export to/pin&save/rename” workspaces, to “create/edit/delete” references and to “check with filters/check all/associate/dissociate” files and references.
- Workspaces are saved in “JSON” format in the “workspaces” folder.

### Annotations

- New annotation “Activity”. This annotation was previously available as an option of “Alignment”.
- Text Normalization: the Num2Letter module has been fully re-programmed.
- Text Normalization: new option “occurrence & duration” to estimate the number of tokens of each IPU and the duration of this latter.
- Annotations are categorized as: STANDALONE, SPEAKER or INTERACTION. The GUI based on wx3 only shows the STANDALONE ones, but the GUI based on wx4 can deal with all of them.

### Resources

- New folder “num” with dictionaries of numbers for cmn, eng, fra, ita, jpn, khm, pol, por, spa, vie.

### 8.5.22 SPPAS 2.4

(2019, 26th June)

### Development

- the GUI based on wx4 is managing the plugins
- bug correction in ‘files’ package

### Annotations

- New annotation “Re-Occurrences”, of type INTERACTION. Can only be used with the CLI or the GUI based on wx4.
- Either one or two new options in all the annotations:
  - inputpattern: to choose the pattern of the input file (like “-token” for Phonetization)
  - outputpattern: to choose the pattern of the output file (like “-token” for Tokenization)
- CLI: the script annotation.py is fully re-implemented

### Resources

- New pronunciation dictionary for Iranian Persian (pes), for Phonetization
- New acoustic model for Iranian Persian (pes), for Alignment

### 8.5.23 SPPAS 2.5

(2019, 30th July)

### Development

- the GUI based on wx4 can convert files.

### Annotations

- New annotation “RMS” to estimate the Root-mean square of an audio file in given intervals

### 8.5.24 SPPAS 2.6

(2019, 1st October)

### Various

- New script ‘dictsampa.py’ to convert a pronunciation dictionary in IPA into SAMPA encoding
- New script ‘clippingrate.py’ to estimate the clipping rate at several factors in sub-parts of the audio file
- New plugin ‘StatGroups’ to estimate distributional statistics on sequences of numbers in annotations.

### Development

- the GUI based on wx4 can convert files (debugged and extended)
- the GUI based on wx4 can display the content of a file (ListView) in the page “Analyze”
- GUI: “Save all” button debugged in DataFilter (remaining bug in DataRoamer)

### 8.5.25 SPPAS 2.7

(2019, 2nd December)

### Various

- Add the possibility to remove un-labelled annotations in tiers (in DataFilter)
- Plugin “StatGroups” updated: it can optionnally create a tier with the not-selected intervals

### Resources

- Modified French lexicon: include words of Quebec French
- New Quebec French pronunciation dictionary
- New Quebec French acoustic model
- Modified French syllabification file: added vowels of Quebec French
- Tiedlist of Italien model re-introduced... Alignment of “ita” is once again completely functional

### Development

- the GUI based on wx4: in “Files”, the list of files uses collapsible panels instead of a table

### 8.5.26 SPPAS 2.8

(2020, 17th January)

Two new scripts (sppas-py3.bat and sppas-py3.command) added to run directly the new GUI based on py3+wx4.

#### Development

- the GUI based on wx4: some debug, improved compatibility with the 3 OS and DataFilter and DataStats implemented in the page “Analyze”

#### Annotations

- New annotation: Stop Tags: estimate if tags of annotations are relevant or not. Relevance is estimated like for Other-Repetitions detections.

### 8.5.27 SPPAS 2.9

(2020, 03th April)

This is the last version for which both Python 2.7 and Python 3.5+ are supported. Running SPPAS with Python 2.7 will not be maintained past 2020, May when SPPAS 3.0 will be released.

#### Development

- the GUI based in wx4 is extended and debugged: audio roamer and ipuscriber integrated in the page ‘Analyze’, a multi-player is added - it can play several media at a time.

#### Resources

- updated dictionary and acoustic model of Polish language

#### Annotations

- New annotation LexMetric to estimate occurrences and rank

#### Known bugs:

In the page ‘Analyze’, the view “Multi-Player” does not properly displays the annotation boundaries and the scrollbar is not updated when needed.

## 8.6 Migrate to Python 3

**Python 2.7 reached the end of its life.** This version expects Python to 3.x as Python 2.7 is no longer maintained. No new bug reports, fixes, or changes will be made to SPPAS when used with Python 2.

A new setup program allows to install external programs to enable some features of SPPAS. It includes 'wxpython' to enable the Graphical User Interface.

### 8.6.1 SPPAS-3.0

(2020, 20th May)

#### Known bugs

- Does not support Julius 4.5+

#### GUI:

- The version based on wxPython 3 is deprecated. Instead, the one based on wxPython 4 is the default but some features are still not implemented.
- A large amount of debug of the GUI. Remaining bug: In the page 'Analyze', the view "Multi-Player" does not properly update the scrollbar when needed.
- Page annotate: can display the content of any report of the list, can delete existing reports.

#### Development:

- new package "preinstall" and new script preinstall.py to install external programs
- Package "files" renamed to "wkps". Updated classes and methods for a better support of reading and writing wjson files; a version 2.0 is then introduced. Add of a "MISSING" state allowing to properly manage missing entries of a workspace, relative paths added, etc.
- config package, classes and methods fully verified/modified including a new sppasAppConfig() class
- Modified **main** to enable to launch SPPAS GUI with the command: "python3 sppas"
- new bin "juliusdownload.py" to automatize the installation of julius under Windows.

#### Resources:

- Updated lexicon, dictionary and acoustic model for Naija Language.

#### Automatic annotations:

- Automatic alignment, when EOT, is rescuing a failed interval alignment with the "Tokens" tier instead of the default one.

### 8.6.2 SPPAS-3.1

(2020, 10th July)

#### Annotations:

- new annotation FaceDetection to detect faces in an image
- new annotation FaceLandmark to find coordinates of 68 points on an image representing a face, and only one
- new annotation SpkLexRep to find what are the tokens or phases repeated from one file to another one of the same speaker
- linguistic resources are downloaded and installed with the setup or directly from Ortolang web site

#### Others:

- renamed .deps~ into .app~

### 8.6.3 SPPAS-3.2

(2020, 4th September)

#### Known bugs

- Under Windows, when a ComboBox is clicked, the Log Window is focused.

#### Annotations

- better support of antx files
- new FaceTracking annotation to detect faces of a video, still under development, and not in the GUI.
- transfer the metadata of the input transcription to the output transcription. Reliable only for annotation formats that support metadata.
- face detection: sort detected faces by confidence score

#### Development

- bug correction and tests added
- debug of the features reading with UTF-8 encoding allowing the setup to work properly
- new videodata package to read and manipulate video files.

#### GUI

- The Home page includes links to the SPPAS web site
- some wx.ComboBox replaced by our sppasComboBox to look the same on all platforms
- Keyboard shortcuts added to message dialogs:
  - ESC: Cancel
  - Return: OK
  - n: No
  - y: Yes
- New keyboard shortcuts in the “Files” page of the main frame:

- alt + a: add files
- alt + e: export workspace
- alt + i: import workspace
- alt + s: pin&save workspace
- alt + n: rename the workspace
- alt + r: create a new reference
- alt + f: check some files
- alt + g: check all files
- alt + l: link checked files to checked references

### 8.6.4 SPPAS-3.3

(2020, 20th October)

#### Package

- “.logs” and “.trash” folders are no longer hidden. Now, they are named “logs” and “trash”.

#### GUI

- The page Analyze is simplified: there’s no longer a notebook to open files into tabs, and there’s no longer different “views”. Instead, 1/ to edit the text content of a file, the page “Files” has a new button “Edit checked” in the top toolbar; and 2/ to view files in a time-line style, a new page “Editor” was created. The latter is still under development.

#### Annotations

- Support of versions 4.5+ of the Julius CSR Engine, used by the ”Alignment automatic annotation.
- new FaceTrack annotation to detect faces into a video. Currently not fully tested.

#### Development

- new package “videodata” to support video files with opencv.
- preinstall: a different way to search for the python command. It should solve problems when the Python3 executable is “python” command instead of “python3”.

### 8.6.5 SPPAS-3.4

(2020, 4th December)

#### GUI

- web links to the reference publications are indicated for each automatic annotations, either in PDF or URL.
- the main app is no longer using the option “useBestVisual” - a user reported that the app crashes with anaconda/ubuntu if enabled.

- some debug and improvement of the page Analyze. Now, it looks better under Windows.
- the page Editor is continued. It allows to:
  1. open audio, video and transcription files (no limit on the number of files);
  2. play all audio and video files really synchronously: the max delay between media players is 10ms;
  3. a slider allows to select 5 different periods to play: either the whole duration, or the visible part, or the selected part, or the visible part until the selected part or the visible part after the selected part.
  4. scroll/zoom the files to change the visible part;
  5. display annotations both in a spreadsheet and in a timeline;
  6. select an annotation;
  7. edit the annotation content: a text entry allows to edit the labels of an annotation either in a serialized string, in XML or in JSON.
  8. add/delete/split annotations.
  9. save.

### Development

- setup.bat was modified, expecting it can install wxpython even if python was not installed from the Windows store. not tested.
- analyze: added a new option in the filter system when using relations. It allows to crop annotations of X when they are overlapping Y (overlaps, overlapped by and contains).

### 8.6.6 SPPAS-3.5

(2021, 20th January)

### GUI

- Page Annotate: the installation of new languages and new annotations is enabled.
- Page Analyze:
  1. view of the waveform of audio files
  2. the files of the timeline can be sorted
  3. in the timeline, right-click on a transcription file to check-uncheck tiers

### Annotations

- LexMetric: bug correction of the supported files. No remaining unjustified error message.
- FaceDetection: can detect faces on both an image or a video.
- FaceMark renamed FaceSights: can fix 68 sights of detected faces on both an image or a video.

### Development

- two new scripts: 1/ to convert a folder of images into a video and 2/ to export a video into a folder of images

### 8.6.7 SPPAS-3.6

(2021, 23th February)

#### Plugins

- cleanipus updated: solved import error and version updated to 1.1

#### Annotations

- INFO1220 migrated from report to logging
- replaced by dummy for unknown phonetizations
- FaceDetection: more options, more models
- New annotation FaceIdentity to assign an identity on detected faces

#### Development

- bug correction in audioframes with an exception not properly raised
- anndata: sppasWEKA renamed to sppasTable and enhanced with export of TRA files – Table Rich Annotations
- installer: added numpy to video feature with version > 1.20 due to the Windows 2004 update and its consequences...
- workspaces: when reading wjson, the references are properly assigned to files
- sppasLPC renamed to sppasCuedSpeech

#### CLI

- the script trstotable.py is replacing trstoweka.py
- other-repetition script can take a workspace as input

### 8.6.8 SPPAS-3.7

(2021, 14th April)

#### Annotations

- creating a merged file is enabled by default
- Overlaps: new INTERACTION annotation to create a tier with overlapping activities.
- Alignment: a better support of non us-ascii characters
- Stopwords: no remaining error about file extension.
- FaceDetection: the default is to load both a DNN and a HAAR model instead of two DNNs. Enlarged portrait size (if option is enabled).
- FaceIdent: can export a smoothed portrait video.
- CuedSpeech: can export the video with the position of the 5 vowels and a number to represent the consonant during the audio key. The resulting tier with the key is made of two labels: the consonant number and the vowel number.

- OtherRepetitions: 3 more tiers about the OR source are created (word strain, nb of labels, source type). Bug correction in the position of the first found echo.
- SelfRepetitions: 3 more tiers about the SR source are created (word strain, nb of labels, source type)
- Linguistic resources and annotation resources are no longer downloaded from ortolang but from gandi instead. Manual download is still on ortolang (when it works!)

### Analysis

- filter system: New filter 'nlab' on the number of labels of annotations

### Development

- bug correction in video coords reader
- videodata: the FPS is not converted to integer anymore, now it can be a float. Support of .mov files with H264 codec.
- updated sppasui.json to re-rank annotations of faces
- images: extended support with transparency and overlay added

### GUI

- page files: bug correction when adding a folder, automatically adjust the column width of the filename when adding new files
- page analyze: new filter added to the SingleFilter, modified files can be saved
- page editor: annotation list view and labels edit re-organized, a new rise panel for the timeline view of file allows to see action buttons when collapsed instead of when expanded. No remaining infinite loop (macos, linux) when displaying tiers. A proportional frame when reading a video.
- Problem with foreground color of lists solved (macos, linux)

### 8.6.9 SPPAS-3.8

(2021, 26th May)

### Annotations

- New annotation IVA - Interval Values Analysis.
- Lexical Metric: Add the number of annotations and the number of labels in segments. Segments are fixed from a list of separators.
- Fill in IPUs: more options to fix shift start and shift end.
- CuedSpeech: bug correction in key generation at the end of an IPU with a consonant.
- CuedSpeech: return 2 different tiers with the key, either with a single combined label or with two different labels (one for C then one for V).
- Alignment: a better way to convert point to intervals in case of alignment without audio.
- Other Repetitions: optionally can add a tier with all echo candidates.
- Extended search for syllables tier name: it allows to apply TGA to any tier with 'Syll' in its name.

## GUI

- buttons have new custom events: it allows to customize their look when mouse is entering, leaving, ...
- page editor: annotation boundaries can be moved. The selected bound is in green if several annotations share the same time value, or in red if the bound corresponds to only one time value. Both the midpoint and the radius can be modified.
- page editor: can create an annotation in the timeline by dragging with the mouse.
- page editor: Some minor changes in the look of the page.
- page editor: bug correction with threads in video player under MACOS. Allows to play several videos at a time, like on the other systems.
- page editor: bug correction in the list of annotations when adding before.
- page files: bug correction when adding a non-valid file.
- page files: adding the parent folder '.' is forbidden.

## Development

- preinstall: added linux+microsoft in the list of supported distrib for those who installed Ubuntu from the Windows Store
- a global sppasLogSetup() instance is declared in the config.
- the script slmtrain.py can take sentences from stdin and then write the probas on stdout.
- anndata: refactor of sppasRW into sppasTrsRW.
- anndata: added a trs\_type in all reader/writer classes and in sppasTrsRW file properties. Allows to classify annotated files into: ANNOT, MEASURE or TABLE.
- annotations: by default, automatic annotations can only create files of the type ANNOT (except if overwrite the method).
- config: add a log error message if no write access to sppas directory.

## Scripts

audioinfo: new options to get clipping values in details or no clipping values at all

### 8.6.10 SPPAS-3.9

(2021, 6th July)

## Annotations

- More logging messages for a better understanding of the automatic annotations procedures
- bugs corrections in RMS and SpkLexRep
- The “merge” can now be checked/unchecked in the page to annotate. If checked, it merges only checked files of each root.

## GUI

- Under MacOS, the GUI application is not crashing anymore on exit - previously it was randomly either not crashing, or bus error or segmentation fault
- Make the GUI work correctly with the MacOS Dock

- Page File: Edit files button is moved to the vertical toolbar and two buttons are added to remove missing files and to add all files sharing checked roots
- Page Editor: An elaborated search dialog is available in the editor in order to find label patterns in annotations of one or several tiers

### Development

- setup for Windows is simplified
- a new class `autils.SppasFiles()` allows a better control of the files to manage in the automatic annotations
- annotations: No longer required vs optional inputs. A more generic solution is used instead in all the annotations

## 8.7 Stable

**Python 2.7 is no longer supported.**

### 8.7.1 SPPAS-4.0

(2021, 5th October)

**\*\* One of Python 3.6 or 3.7 or 3.8 is required.\*\***

### GUI

- `sppas.bat`: do not launch `pythonw3` anymore but `python3` in order to display error messages if any
- GUI for `wx3+py2` is definitively removed
- A new “Colored” icons theme
- Translation enabled under MacOS
- Bug correction: the settings are not modified anymore when cancel is clicked
- Fade in and fade out of dialog windows can be fixed into the settings
- The splash delay can be fixed into the settings
- The Page Editor can play synchronously audio and video files
- Page Editor: Bug correction in displaying the waveform. It is now properly synchronized on the timeline.
- Page Editor: Bug correction in the annotation label editor which can now write special chars - very useful for an enriched transcription.
- Page Files: Remove missing files also remove remaining missing roots and paths.

### Development

- config: simplified use of locale. We now use language but not country (`en_US => en` ; `fr_FR => fr`)
- gui: the application has a new custom splash with a gauge to indicate the loading progress
- gui: the install app is updated – code cleaning, better macOS support, etc
- annotations: bug correction in Face Detection because in some videos an image of the buffer can be `None`
- annotations: bug correction of input pattern management in LPC, Syllabify, FaceIdent, Repet
- annotations: JSON config file modified for LPC and Syllabify

- tips moved to structs package (they are now unused)
- ui: only one gettext message definition of \_ in init. No longer a “msg” definition in each class...
- ui.players: code re-organized, class diagram simplified, package moved into ui.phoenix
- videodata: a faster solution to seek at a pos when the pos is closely later
- preinstall: the installer uses sys.executable to get the appropriate python

### Annotations

- CuedSpeech: use a different code to represent vowels and consonants. The new configuration file is not compatible with previous versions of SPPAS.

### 8.7.2 SPPAS-4.1

(2021, 25th November)

### Known problems

- ‘audioplay’ feature is not available for python 3.9 or 3.10 because simpleaudio pypi package is not available. Another solution has to be found...
- opencv-contrib 4.5.4+ is not supported. The authors changed something in the library. Some tests have to be performed to find what exactly changed and how SPPAS can deal with. opencv-contrib is required by the ‘video’ feature.

### Package

- The XRA 1.5 scheme is described in a PDF documentation file (16 pages)
- The package includes the PDF of the documentation of resources
- The package includes a PDF with the workflow of 19 auto. annotations

### GUI

- page edit: the waveform is not drawn in an over-period anymore.
- page edit: new keyboard shortcuts in the timeline to:
  - zoom in (ctrl+i),
  - zoom out (ctrl+o),
  - scroll before (ctrl+left)
  - scroll later (ctrl+right)
  - select the next annotation (Alt+RIGHT)
  - select the previous annotation (Alt+LEFT)

### Annotations

- Search for IPUs: the min sil value is taking into account the shift values. In the new results, none of the silences will have a duration less than the given one.

- By default, the Face Detection, Face Landmark and Person Face Identity are creating an XRA file instead of a CSV one.
- Person Face Identity is creating a video and an XRA of the portrait of each identified person. It also has a better selection of identities, based on both coordinates overlap and image similarity.
- The “do merge” option is turned off by default when using either the GUI or the CLI
- An annotation development progress indicator is introduced. This is a percentage value defined into the annotation JSON files and displayed in the GUI if no reference is available.
- Person Face Identity: two different videos are created when ‘tag’ option is enabled, one without the tagged face and one with.

## Resources

- Support of Bengali language.

## Development

- an exception is raised at the **init** stage if python is deprecated.
- `anndata.annlocation`: bug correction in filtering system with overlaps intervals
- `anndata`: Two new types are introduced for a `sppasTag`: it can be a “fuzzy point” or a “rectangle”
- `anndata`: A `sppasLabel` has a new member ‘key’ (string).
- new XRA format. Version 1.5 can take into account the 2 new types of a `sppasTag` and the ‘key’ attribute of a label.
- bug correction in the workspaces: the application captures the `ValueError` that is raised when using “relpath” on Windows OS if the path is on a different drive.
- bug correction in the workspaces when adding brothers in roots
- Less restrictions in the definition of a pattern filename (now it can have a single char).
- “sppas” is not longer added in the python path by the **init**
- `models` package migrated into ‘Align’ to provide circular imports. The training of ACM models is updated.
- Bengali language added to the `num2letter` module
- `fileutils.py` migrated into `utils` package; `makeunicode.py` migrated into `config` package
- `annotations`: bug correction with exception `EmptyOutput` and an added message in `TextNorm` and `Phon` (name of the tier).
- `structs`: bug correction in `lang list`. It isn’t ignoring the resource name anymore.
- `videodata`: bug correction in the `videowriter.set_size()` in order to write videos of a custom size.
- `sppas.json` renamed `sppasmeta.json` with a few meta added
- `install`: takes the version into account when installing a pypi package
- `setup.bat`: add a 5 seconds delay after `wx` is installed in order to let Windows OS realize that something changed...

### 8.7.3 SPPAS-4.2

(2022, 28th January)

## Known problems

- This version (and the previous ones) does not work with Python 3.10.x

- ‘audioplay’ feature is not available for python 3.9 or 3.10 because simpleaudio pypi package is not available. Another solution has to be found...
- **The use of Python 3.6, 3.7 or 3.8 is then recommended.**

## Package

**Important** The setup is now creating a Python Virtual Environment in the hidden folder `.sppaspyenv~`. Then, the setup has to be launch each time a new SPPAS package is installed. The sppas commands are launching the GUI with Python of this virtual environment.

## Annotations

- Syllabification: a new option to output one more tier with syllable structures
- CuedSpeech: code cleaning and re-organization, two more tiers are created: HandShapes and HandPositions. They are predicting respectively the moments D1-D2 (change of shape) and M1-M2 (change of position). The prediction model is a set of rules based on the work described in (Attina, 2005), page 117 and page 136. It can export a video with badges moving and standing depending on this model, or with pictures of a hand when the debug logging level is set (in the `.app~` file).
- FaceDetection: can choose the model(s) to enable or disable in the options. MediaPipe Face Detection is added to one of the possible recognizers - it’s very fast but results are worse than other models.
- FaceSights: MediaPipe Face Mesh is used to predict the 68 face sights (weight=6) interpolated with OpenCV LBF (weight=1). No longer use of the DAT model nor the empirical sights.
- PersonFaceIdentity: A new filter is added to remove un-relevant known identified persons, when number of stored images is less than 1/3 of the expected ones. Added options: mincoords, minrefcoords, mindistcoords, mindistings, nbfrimg. And now, the person is identified with the automatic face recognizer first, then with the coordinates instead of the contrary.

## GUI

- Page Edit: new buttons in the player control panel to play one frame of a video in both ways (next or previous frame)
- Page Files: a better view of the filename in the tree, particularly under Windows, and the pattern is now in a specific column.

## Development

- the “video” feature is now also requiring MediaPipe - [https://mediapipe.dev/#!](https://mediapipe.dev/)
- workspaces: extended max length of the key of an attribute. It was 12, it’s 20.
- features.ini and object detection: can support all versions of opencv-contrib. They changed the output of the detection data structure at version 4.5.4.
- imgdata: the objects detection can load all models then enable or disable any of them.
- Sights and SightsWriter migrated into imgdata package.
- sppas.command is ‘cd’ to \$SPPAS in order to make sppas.command working properly with the Finder on MacOS and is creating a ‘symlinks’ virtual environment on MacOS because it does not support the ‘copies’ option
- setup.bat: search for python.exe first.

### 8.7.4 SPPAS-4.3

(2022, 22th March)

#### Known problems

**This version has an “import” bug: it does not start at all on some computers.**

#### Resources

- Updated French acoustic model, added French pronunciation dictionary for standard French with name `fre.dict`, and updated documentation of French resources.

#### Annotations

- MediaPipe (<https://mediapipe.dev/>) can be used by Face Detection and is used by default in the Face Landmarks automatic detection. It is installed in order to enable the video feature.
- Face landmarks: smooth sights trajectory
- Search for IPUs: the default output pattern is `-ipus`
- RMS: can optionally add 2 tiers with min and max value of each interval
- CuedSpeech: a new tier is created with the coordinates of the vowels (if sights are given)
- If any, add properly the language in the metadata of the output transcription file, and add the linguistic resource filename in the metadata of the tier.

#### Development

- `anndata`: bug correction when reading and writing `.lab` files
- `plugins`: `audioseg` bug correction when searching for the annotated file and a new option is added.
- `plugins`: `classphones` bug correction in imports and new options.
- `GUI-Edit`: the player control can play frame-by-frame with a step value, not only 1.
- `facesights` & `faceidentity`: smooth points trajectory with another algorithm
- `cuedspeech`: what, when and where files separated in 3 packages

#### GUI

- updated icon: a more modern version with fewer colors
- updated splash image in order to add a representation of both a video and a hand
- `plugins`: solved a problem of page refresh under Windows when adding or deleting a plugin

#### Scripts

- `formants.py` creates a file with `f1`, `f2`, `f3`, `f4` of all phonemes. It requires Praat and to install the pypi of `ParseLMouth`.
- `saplay.py` plays an audio file or a part of it. It uses either `pyaudio` or `simpleaudio` - one of them must be installed with pypi.

- `img_to_cartoon.py` is able to cartoonize an image or a video.
- `sppasseg.py` allows to perform the full automatic segmentation process. It's a three automatic steps process:
  1. Prepare the data (if required):
    - Nothing is done if the ortho. transcription is already time-aligned into the IPU's;
    - Fill in IPU's if the ortho. transcription is already available but not time-aligned. Notice that it can't work if the recording protocol was not very very very strict: it works only if the number of sounding segments EXACTLY matches the given number of transcribed segments.
    - Search for IPU's and auto. transcribe with the STT DeepSpeech. It requires installing DeepSpeech in the currently used Python and to give the path of the models. Get it here (DeepSpeech + English model): <https://github.com/mozilla/DeepSpeech/releases/> or here for another language: <https://discourse.mozilla.org/t/links-to-pretrained-models/62688>
  2. Perform Speech Segmentation:
    - no option to configure the annotations. They are all fixed internally!
    - can estimate an intensity coefficient;
    - can predict silent pauses (the ones into the IPU's).
  3. Format the output (if asked):
    - can map the phoneme into another set (IPA, Praat, ...) if a mapping table is given. It's a file with two columns, both separated by a whitespace: 1st is SPPAS phoneme, 2nd is expected phoneme.
    - can create a data structure with all the results and write into a JSON file.

### 8.7.5 SPPAS-4.4

(2022, 5th April)

#### Known problems

The use of Python 3.6 or 3.7 is possible but **Python 3.8 is recommended** because: - the feature "audioplay" can't be enabled with a more recent version, - recent versions of MediaPipe have better results and the most recent one is available for python 3.8.

#### Annotations

- Hand&Pose: a new automatic annotation of an image or a video. SPPAS is a wrapper for MediaPipe Hand detection and Mediapipe Pose detection. It also proposes a custom solution in order to detect right-left hands of a person.

### 8.7.6 SPPAS-4.5

(2022, 15th April)

### Known problems

For a full experience of SPPAS **using of Python 3.8 or 3.9 is recommended**. The use of Python 3.6 or 3.7 is possible.

### Install

- Under Windows, pipwin is installed and used when pip failed to install a package. pipwin installs unofficial python package binaries for Windows provided by Christoph Gohlke here: <http://www.lfd.uci.edu/~gohlke/pythonlibs/>
- Both simpleaudio and PyAudio are installed - if available for the OS. The “audioplay” feature is enabled if one of them is successfully installed.

### Package

- a demo is added with all its annotated files. Only the orthographic transcription was performed manually. SPPAS did the rest.

### Resources

- updated French dictionary (fra.dict)

### GUI

- Page Edit: select the nearest annotation when clicking in a hole of a tier
- Page Analyze: when a tier is copied/paste, its media and controlled vocab are also copied/paste to the destination file.

### Annotations

- Alignment: add the aligner name in the metadata of the result, i.e. either julius, hvite or basic.
- FaceSights: bug correction when writing coordinates in a CSV file.

### 8.7.7 SPPAS-4.6

(2022, 22th August)

### Known problems

For a full experience of SPPAS **using of Python 3.8 is recommended** because some of the dependencies are no longer maintained and no new release is published. Other versions of Python are possible but without any warranty.

## Annotations

- CuedSpeech is continued. The Proof of Concept is turned into some stable and tested packages but it is not yet a version 1.
- FaceIdentity: new option to shift the portrait/selfie at left or right
- FaceIdentity and FaceSights: debug of the smoothing. The points are no longer shivering.
- HandPose: bug correction when writing hand sights for an image for XRA format
- Updated demo files with the results of the automatic annotations

## Plugins

- The plugins folder migrated from the base directory to the sppas directory; and each plugin is now a python package.
- audioseg: A version 3.0 is released, an option “video” was added in order to segment both the audio and the video into tracks. Track filenames start with an index instead of the interval number.

## Resources

- updated French dictionary (fre.dict) with new entries added and with “J” instead of “n j” in all words containing “gn”

## Development:

- anndata.tier: bug correction in index() method in order to search for the index of a point in a point-tier.
- preinstall: a new key “pip\_opt” in the features.ini file. The installer is now trying pip with option ‘-pre’ if the installation of the stable version failed’. The pip\_opt allows to set options to pip. It is used by wxPython to fix the pre-build host.
- imgdata.coordinates: added a xy ratio in parameters of portrait() in order to force a ratio between width and height. Default is 14:16.
- anndata.label: can optionally add the same tag into a label, without adding their scores
- anndata.aio.praat: an appropriate error message in case of chronological textgrid file.

## 8.7.8 SPPAS-4.7

(2022, 12th October)

## Annotations

- CuedSpeech is continued. The transition duration of the shape is reduced (d2 is before) and the hand picture is not blur on the video when the position is changing.
- Text Normalization: a new option is used by default to estimate the number of tokens in each IPU but only if the token is speech, events are ignored (i.e. dummy, noise, laugh, pause and silence).

## Development

- Python 3.10 compatibility.

- Bug correction with imports if video is not enabled.
- Bugs correction in workspaces: Can append a file with a whitespace before the pattern, like “toto - Copie.txt” and the relative path should work properly.
- Installation: bug correction with quotes when launching a command that solved the installation problem - setup.bat does not need to be launch twice anymore, once is enough!
- Installation: do not stop an installation if a pip package fails to be updated.
- Script mlf2labs.py: a new script to split an MLF file into several lab files.
- The custom event’s management changed in the Phoenix UI.

## **GUI**

- There’s no longer a LogWindow in the installation wizard.
- Bug correction in slider and progress dialog (mainly used by the Edit page).
- Revised toolbars.
- Page Files: 2 buttons moved from associate to files (remove missing, add bros) and the delete button migrated from files to associate.
- Page Editor: Search dialog debugged.

## **Resources**

- Updated list of phones for Bengali language, updated dictionary and acoustic model



## Appendix

### 9.1 List of error messages

SPPAS introduced a system for the internationalization of its messages. In the same time, a quality and a number was assigned progressively to each of them in the packages. This system is aimed at providing the users with appropriate error messages and a deeper understanding of the problems.

#### 9.1.1 Global errors

##### **:ERROR 0001:**

This is an un-classified error, with the generic name “sppasError”. A message after the error should indicate the problem.

##### **:ERROR 0513:**

This is the `sppasPermissionError`, with the message: access to {place} is denied. It occurs when SPPAS was not able to create a file or a directory in some place of the computer. The most frequent cause is that the user do not have permissions to write in the directory where the SPPAS package is installed, or the rights are configured incorrectly or not configured at all.

To work around this issue, use one of the following workarounds as appropriate for your situation:

1. install SPPAS package in a place you have appropriate rights;
2. ask your OS Administrator to modify the rights of the directory in which SPPAS package is installed;
3. launch SPPAS with Administrator rights (not recommended).

#### 9.1.2 From “annotations” package

##### **:ERROR 1010:**

Unknown option with key {key}.

**:ERROR 1020:**

This error has the message: Empty input tier {name}. It occurs when an automatic annotation is expecting an annotated tier but the given one does not contain annotations.

**:ERROR 1030:**

Missing input tier. Please read the documentation.

**:ERROR 1040:**

Bad input tier type.

**:ERROR 1050:**

Inconsistency between the number of intervals of the input tiers. Got: { :d } and { :d }.

**9.1.3 From “utils” package**

:ERROR 1210: The directory {dirname} does not exist.

:ERROR 1220: The directory {dirname} does not contain relevant data.

**9.1.4 From “audiodata” package**

:ERROR 2000: No audio file is defined.

:ERROR 2005: Audio type error: not supported file format {extension}.

:ERROR 2010: Opening, reading or writing error.

:ERROR 2015: No data or corrupted data in the audio file {filename}.

:ERROR 2020: {number} is not a right index of channel.

:ERROR 2025: From {value1} to {value2} is not a proper interval.

:ERROR 2050: No channel defined.

:ERROR 2060: Channels have not the same sample width.

:ERROR 2061: Channels have not the same frame rate.

:ERROR 2062: Channels have not the same number of frames.

:ERROR 2070: Invalid sample width {value}.

:ERROR 2080: Invalid frame rate {value}.

### 9.1.5 From “calculus” package

:ERROR 3010: Both vectors p and q must have the same length and must contain probabilities.

:ERROR 3015: Value must range between 0 and 1. Got {f}.

:ERROR 3016: Probabilities must sum to 1. Got {f}.

:ERROR 3025: Error while estimating Euclidian distances of rows and columns.

:ERROR 3030: The given data must be defined or must not be empty.

:ERROR 3040: Value {value} is out of range: expected value in range [{min\_value},{max\_value}].

### 9.1.6 From “plugins” package

:ERROR 4010: Missing plugin configuration file.

:ERROR 4014: Missing section {section\_name} in the configuration file.

:ERROR 4016: Missing option {s} in section {s} of the configuration file.

:ERROR 4020: Unsupported plugin file type.

:ERROR 4024: Unsupported plugin file type.

:ERROR 4030: A plugin with the same name is already existing in the plugins folder.

:ERROR 4040: No plugin with identifier {plugin\_id} is available.

:ERROR 4050: No such plugin folder: {s}.

:ERROR 4060: A plugin with the same key is already existing or plugin already loaded.

:ERROR 4070: {command\_name} is not a valid command on your operating system.

:ERROR 4075: No command was defined for the system: {s}. Supported systems of this plugin are: {s}.””””

:ERROR 4080: No option with key {s}.

### 9.1.7 From “resources” package

#### **:ERROR 5005:**

Encoding error while trying to read the file: {name}.

#### **:ERROR 5010:**

Error while trying to open and read the file: {name}.

#### **:ERROR 5015:**

Read file failed at line number {number}: {string}.

#### **:ERROR 5020:**

The n value of n-grams pattern matching must range [1;{maximum}]. Got {observed}.

**Others:**

:ERROR 5022: The gap value of pattern matching must range [0;{maximum}]. Got {observed}.

:ERROR 5024: The score value of unigrams pattern matching must range [0;1]. Got {observed}.

:ERROR 5030: The dump file can't have the same extension as the ASCII file ({extension}).

:ERROR 5040: The count value must be positive. Got ({count}).

**9.1.8 From “structs” package**

:ERROR 6010: {meta} is not a known meta information.

:ERROR 6020: Unknown resource type: expected file or directory. Got: {string}.

:ERROR 6024: The resource folder {dirname} does not exists.

:ERROR 6028: The language must be “und” or one of the language list. Unknown language {lang}.

**9.1.9 From “models” package**

:ERROR 7010: Expected a {data\_name} of type {expected\_type}. Got {data\_type} instead.

:ERROR 7500: The file {!s:s} contains non UTF-8 characters: {!s:s}.

:ERROR 7505: Fail formats: unrecognized file format {!s:s}.

:ERROR 7510: Fail formats: the folder {!s:s} does not contain a known model.

:ERROR 7515: No model found or empty model in {!s:s}.

**9.2 GNU Free Documentation License**

Version 1.3, 3 November 2008

Copyright (C) 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc. <http://fsf.org/>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

**0. PREAMBLE**

The purpose of this License is to make a manual, textbook, or other functional and useful document “free” in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of “copyleft”, which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software

does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

## 1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a worldwide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The “Document”, below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as “you”. You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A “Modified Version” of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A “Secondary Section” is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document’s overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The “Invariant Sections” are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The “Cover Texts” are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A “Transparent” copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not “Transparent” is called “Opaque”.

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The “Title Page” means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not

have any title page as such, “Title Page” means the text near the most prominent appearance of the work’s title, preceding the beginning of the body of the text.

The “publisher” means any person or entity that distributes copies of the Document to the public.

A section “Entitled XYZ” means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as “Acknowledgements”, “Dedications”, “Endorsements”, or “History”.) To “Preserve the Title” of such a section when you modify the Document means that it remains a section “Entitled XYZ” according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

## **2. VERBATIM COPYING**

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

## **3. COPYING IN QUANTITY**

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document’s license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

## 4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled “Endorsements”, provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

## **5. COMBINING DOCUMENTS**

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled “History” in the various original documents, forming one section Entitled “History”; likewise combine any sections Entitled “Acknowledgements”, and any sections Entitled “Dedications”. You must delete all sections Entitled “Endorsements”.

## **6. COLLECTIONS OF DOCUMENTS**

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

## **7. AGGREGATION WITH INDEPENDENT WORKS**

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an “aggregate” if the copyright resulting from the compilation is not used to limit the legal rights of the compilation’s users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document’s Cover Texts may be placed on covers that bracket

the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

## 8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled “Acknowledgements”, “Dedications”, or “History”, the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

## 9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

## 10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License “or any later version” applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy’s public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

## 11. RELICENSING

“Massive Multiauthor Collaboration Site” (or “MMC Site”) means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A “Massive Multiauthor Collaboration” (or “MMC”) contained in the site means any set of copyrightable works thus published on the MMC site.

“CC-BY-SA” means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

“Incorporate” means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is “eligible for relicensing” if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.