

the Aix-Marsec database

Daniel Hirst, Cyril Auran & Caroline Bouzon

2002-2004

version: 2009 December 09

*Equipe Prosodie et Représentation Formelle du Langage
Laboratoire CNRS UMR6057, Parole et Langage
Université de Provence
Aix-en-Provence, FRANCE*

daniel.hirst@lpl-aix.fr

This database may be freely distributed and used without any restriction except that it should always be accompanied by this notice.

Our only request is that the providers of the database (us) should be informed of any enrichments you or others may make to it and that these enrichments should be made freely available for future distributions.

This CD-Rom contains the following folders besides this *Read Me* file.

- **Aix-Marsec.praat** – a Praat script which allows you to view the Sounds and TextGrids of the Aix-Marsec database in Praat.

- **documents** - a number of reference documents etc. used by the scripts

- **labels** - the original label files from the Marsec corpus. When the original word label files from the Marsec corpus have been manually corrected (by Caroline Bouzon) this is indicated by the suffix `_mod`

- **Praat plug-ins** – a set of plugins adding Momel, Intsint and other functions to the Praat software.

- **publications** - a selection of articles in *pdf* format either concerning the database itself or the techniques used to label it (in particular Momel and Intsint).

- **scripts** - a collection of Perl and Praat scripts (mostly by Cyril Auran) used in generating the database from the original Marsec corpus.

- **tabular data**

A tabular file containing data extracted from the complete Aix-Marsec database. This is described in detail below as well as in the header of the file.

- **TextGrid** - Praat label files for each sound file containing the following tiers :

Phonemes, Syllables, Abercrombie, Jassem, Text, UI, Intsint, Valeurs de F0

The content of these tiers is described in detail on the next page.

- **txt** - ascii texts corresponding to each sound file

- **wav** - 408 *wav* format sound files from the Marsec corpus, each corresponding to about one minute of sound.

Tiers used in the TextGrid label files

- **Phonemes** - SAMPA labels for the phonemes obtained by dictionary and aligned automatically (cf Auran et al. 2004, Auran, Bouzon & Hirst 2004).

- **Syllables** - defined according to the Maximal Onset Hypothesis (an intervocalic cluster is segmented attributing as many consonants as possible to the onset, consistent with permissible syllable onset constraints).

"city" = [SYLLSI][SYLLtI]

- **Abercrombie** – Following Abercrombie (1964), feet (labelled F, pauses are labelled P) begin after an intonation boundary or with a stressed syllable and continue until before the next boundary or stressed syllable.

"it's 'almost im'possible" = [F-it's][F-almost im-][F-possible]

- **Jassem** – Following Wiktor Jassem (1952) Rhythm Units are of two kinds: Narrow rhythm units (labelled NRU), which begin with a stressed syllable and end at the following word boundary, and Anacruses (labelled ANA) containing any syllables which are not in a NRU.

"it's 'almost im'possible" = [ANA-it's][NRU-almost][ANA-im-][NRU-possible].

- **Text** - word labels including tonetic stress marks (TSM) as used by GOK or BJW in the original SEC corpus [words and TSMs are to be put onto two separate tiers]

- **UI** (this should be IU – UI is left over from French versions) Intonation Units as delimited by minor and major intonation boundaries. The units are annotated (I) when ended by a minor boundary and (II) when ended by a major intonation boundary. Pauses are included in the following intonation unit. [to be modified]

- **Intsint** - the intonation pattern is represented as a sequence of tonal labels using the INTSINT alphabet (cf Hirst & Di Cristo 1998, Hirst et al. 2000, Hirst 2000, 2001, 2005a). The labels were obtained automatically from the *Valeurs de F0* tier using an algorithm described in Hirst 2005.

- **Valeurs de F0** (another name left over from French versions)- These are the "target pitch" values output by the automatic algorithm **Momel** (cf Hirst & Espesser 1993, Hirst et al 2000, Hirst 2005).

Variables used in the tabular data file.

The tabular data file contains the following heading:

```
# Aix-Marsec dat.txt
#
#   Produced between 2002 and 2004 by members of the research team:
#   Prosody and the Formal Representation of Language
#   (Prosodie et Représentation Formelle du Langage)
#   Daniel Hirst (leader), Cyril Auran, Caroline Bouzon
#   Laboratoire CNRS UMR6527, Parole et Langage,
#   Université de Provence
#   Aix-en-Provence, FRANCE
#
#   daniel.hirst@lpl.univ-aix.fr
#
#   This file was produced from the Aix-Marsec database and is in appropriate format
#   as input for statistical software such as R.
#   Values are separated by tabs. Missing values are indicated by "NA"
#   The file contains the following variables:

#index a unique index for each phone or pause in the corpus
#file a five symbol code representing the block and passage
#word the lexical form of the word (pauses are indicated by "PAUSE")
#phon the corresponding phoneme coded using the SAMPA phonetic alphabet
#dp the duration of the phone or pause in milliseconds
#zp the z-score of the phone obtained from the mean (mp) and standard deviation (sp)
for each phoneme for the whole corpus (zp = (dp - mp)/sp)
#z_code a five value code for z-score: "--" [-∞...-1.5]; "-" [-1.5...-0.5];
# "0" [-0.5...0.5]; "+" [0.5...1.5]; "++" [1.5...∞]
#ds duration of syllable in milliseconds
#nps number of phonemes in current syllable as defined by Maximal Onset Principle. #
"city" = /(sI)(tI)/ : nps = 2 2 2 2
#npsa number of phonemes in current syllable as defined to allow for ambisyllabicity.
# "city" = /(si(t)I) : npsa = 3 3 3 2
#dsa duration in ms. of the current syllable - defined to allow for ambisyllabicity
#df duration in milliseconds of the current foot - which begins after an intonation
# boundary or with a stressed syllable, and finishes before the next boundary
# or stressed syllable
#npf number of phonemes in the current foot
#nsf number of syllables in the current foot
#dw duration in milliseconds of the current word
#npw number of phonemes in the current word
#nsw number of syllables in the current word
#di duration in milliseconds of the current intonation unit which begins and ends
# with an intonation boundary.
#npi number of phonemes in the current intonation unit
#nsi number of syllables in the current intonation unit
#nfi number of feet in the current intonation unit
#nwi number of words in the current intonation unit
#dr duration in milliseconds of rhythm units - these are of two types
# - narrow rhythm units - which begin with a stressed syllable and finish
```

at the next word boundary
- anacrusis which consist of all the syllables not in a narrow rhythm unit
#npr number of phonemes in the current rhythm unit
#nsr number of syllables in rhythm unit
#pho phoneme type "C" consonant, "V" short vowel, "V:" long vowel, "VD" diphthong, "NA" for PAUSE.
#ss syllable structure "O" onset, "N" nucleus, "C" coda,
assuming syllable as defined by Maximal Onset Principle.
#ssa Potentially ambisyllabic consonants coded "A" others "NA".
#str stress: "S" stressed, "U" unstressed, "NA" for pauses.
#tsm tonetic stress mark as coded by BJW or GOK.
#ru nature of rhythm unit "R": narrow rhythm unit (begins with stress), "A": anacrusis, "S": initial anacrusis after intonation boundary, "NA" for PAUSE.
#ips index of phoneme in syllable
#ipr index of phoneme in rhythm unit
#ipw index of phoneme in word
#ipi index of phoneme in intonation unit
#

References

- Abercrombie, D. 1964. Syllable quantity and enclitics in English. In Abercrombie, D., Fry, P., MacCarthy, N. & Trim, J. (eds): *In Honour of Daniel Jones*. London: Longman, 216-222.
- Auran, C., Bouzon, C. & Hirst, D.J. 2004. The Aix-MARSEC project: an evolutive database of spoken English. In Bel, B. & Marlien, I. (eds) *Proceedings of the Second International Conference on Speech Prosody*, Nara, Japan, 561-564.
- Auran, C., Bouzon, C., Hirst, D.J., Lévy, C., Nocéra, P. 2004. Algorithme de prédiction d'élisions de phonèmes et influence sur l'alignement automatique dans le cadre du projet Aix-MARSEC. in *Actes des XXVèmes Journées d'Etudes sur la Parole*, Fès, Maroc. 57-60.
- Hirose, K.; Hirst, D.J. & Sagisaka, Y. (eds.) 2005. *Quantitative prosody modeling for natural speech description and generation* (=Speech Communication, 46 (3-4)).
- Hirst, D.J. 2000. ProZed: a multilingual prosody editor for speech synthesis. *Proceedings of IEE Workshop on Speech Synthesis*. London, April 2000.
- Hirst, D.J. 2001. Automatic analysis of prosody for multi-lingual speech corpora. in E. Keller, G. Bailly, A. Monaghan, J. Terken & M.Huckvale (eds.) *Improvements in Speech Synthesis*. (London, John Wiley). 320-327.
- Hirst, D.J. 2005 Form and function in the representation of speech prosody. in K.Hirose, D.J.Hirst & Y.Sagisaka (eds) *Quantitative prosody modeling for natural speech description and generation* (=Speech Communication 46 (3-4)), 334-347.
- Hirst, D.J.; Auran, C. 2005. Analysis by synthesis of speech prosody: the ProZed environment. *Interspeech/Eurospeech 05. 9th European Conference on Speech Communication and Technology*, September 2005, Lisbon.
- Hirst, D.J.; Bouzon, C. 2005. The effect of stress and boundaries on segmental duration in a corpus of authentic speech (British English). *Interspeech/Eurospeech 05. 9th European Conference on Speech Communication and Technology*, September 2005, Lisbon.
- Hirst, D.J. & Di Cristo, A. 1998. A survey of intonation systems. in Hirst & Di Cristo (eds) 1998. *Intonation Systems. A survey of Twenty Languages*. (Cambridge, Cambridge University Press), 1-44.
- Hirst, D.J. & Di Cristo, A. (eds) 1998. *Intonation Systems. A survey of Twenty Languages*. (Cambridge, Cambridge University Press). [ISBN 0 521 39513 S (Hardback); 0 521 39550 X (Paperback)].
- Hirst, D.J. & Espesser, R. 1993. Automatic modelling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de Phonétique d'Aix* 15, 71-85
- Hirst, D.J.; Di Cristo, A.; Espesser, R. 2000. Levels of representation and levels of analysis for intonation. in M. Horne (ed) *Prosody : Theory and Experiment*. Kluwer Academic Publishers, Dordrecht. 51-87
- Jassem, W. 1952. *Intonation in Conversational English*. Warsaw, Polish Academy of Science.

Version history

2009:12:09 added Aix-Marsec.praat script to view Sounds and TextGrids with Praat.

2007:09:29 added Momel-Intsint plugins

2005:10:26 added momel-intsint folder to scripts

2005:09:01 reorganised CD to include *_Read Me.doc* file, publications and tabular data