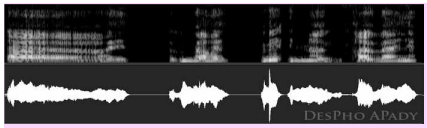


http://despho-apady.univ-avignon.fr/

A new approach for an acoustic-phonetic description of dysarthria

DesPho-APaDy project

L. Crevier-Buchman [1], C. Fougeron [1], C. Fredouille[2], A. Ghio[3], C. Meunier[3], C. Delooze[3], D. Duez[3], C. Gendrot [1], T. Legou[3], N. Lévêque [1], N. Audibert[2], A. Colazo Simon [1], C. Pillot-Loiseau [1], S. Pinto[3], G. Pouchoulin[2], D. Robert[3], F. Viallet[3], C. Vincent [1], B. Teston[3], C. Chevrie-Muller[4]

1 Lab. de Phonétique et Phonologie, UMR 7018 CNRS-Paris3/Sorbonne Nouvelle, Paris, France

2 University of Avignon, CERI/LIA, Avignon, France

3 Lab. Parole et Langage, UMR 6057 CNRS Aix-Marseille Univ., Aix-en-Provence, France

4 Lab. MoDyCo, UMR 7114, CNRS- Université Paris 10, Paris, France

A multidisciplinary team

Composed of phoneticians, clinicians, computer science engineers and automatic speech processing

3 partners with 27 participants :

- **Laboratoire de Phonétique et Phonologie, Paris (UMR 7018)**
- **Laboratoire Parole et Langage, Aix en Provence (UMR 6057)**
- **Laboratoire d'Informatique d'Avignon, Avignon (UPRES 4128)**

Rationale and Objectives of the Project

- **What do we want?**

  Identify and quantify reliable markers which are characteristic of different types of dysarthric speech profile, and could be followed in time.

  - Severity evaluation
  - Disease progression
  - Treatment efficacy

  Select reliable and robust French acoustic phonetic criteria able to distinguish

  - Normal and dysarthric speech
  - Different dysarthria types

Rationale and Objectives of the Project

- **Issues?**
  - Which technique & nature of the markers?
    - Perceptual
    - Acoustic
    - Articulatory
  - Which speech dimension to focus on?
    - Presumed altered speech dimension
    - All dimensions
  - What resources are available?
    - Single vs. multiple judges (expert-naive)
    - Manual vs. automatic processing devices

Rationale and Objectives of the Project

- **Approaches**
  - Combine semi-automated procedures from tools developed for automatic speech processing or scripts used for acoustic measurements **and**
  - Manual phonetic analysis at different level:
    - Temporal
    - Segmental
    - Suprasegmental
  - ➢ **Permanent back and forth between manual and automatic procedures**
  - Processing of a large amount of speech files
    - over 100 patients (SLA, Parkinson, Cerebellar)
    - with 1-2 min. of text reading per patient
  - ➢ **With minimal cost of time and human expertise**

The Multiple-Field Query Database

The CCM (Claude Chevrie-Muller)
Corpus recorded over 30 years
(1965 – 1997)

~1000 hours of disordered speech, 5000 patients – adults and children

* 860 patients **classified** according to their **neurological diagnosis**

***** 60 control speakers

Data recorded

Sound, EGG (ElectroGlottoGraph)

6

The Multiple-Field Query Database:

- The ANH (Aix-Neurology-Hospital)

Corpus recorded for the past 15 years:

* 990 patients,
* 160 control speakers

Data recorded

Sound, aerodynamic.

Selection of Patients for the Acoustic Phonetic Study)

- Focus on neurophysiologic alterations of 3 neurologic systems:
  - Pyramidal system: ALS dysarthria (30 patients)
  - Cerebellar system: ataxic dysarthria (30 patients)
  - Extrapyramidal system: Parkinsonian dysarthria (30 patients from the ANH corpus)
- Selection based on:
  - Severity of the dysarthria (clinical information, the certainty of the diagnosis, the ongoing treatment, patient demographics)
    - ➢ The relatively intelligible speech

8

Type of Acoustic parameters

- **Voice quality** and **fundamental frequency** measurements
  - to characterise laryngeal control
    - *F0* (Hz) (mean+stdev) measured in the middle of vowels +
    - F0 contours on selected sentences
- **Amplitude** measurements
  - for respiration & laryngeal control
- Measurements related to **nasalisation**
  - for velopharyngeal functions
- **Spectral** and **durational** correlates for vowels and consonants
  - for kinematics of supralaryngeal articulatory movements and
  - movement coordination
  - *F1, F2, F3* (mean+stdev) measured at different points in the vowels

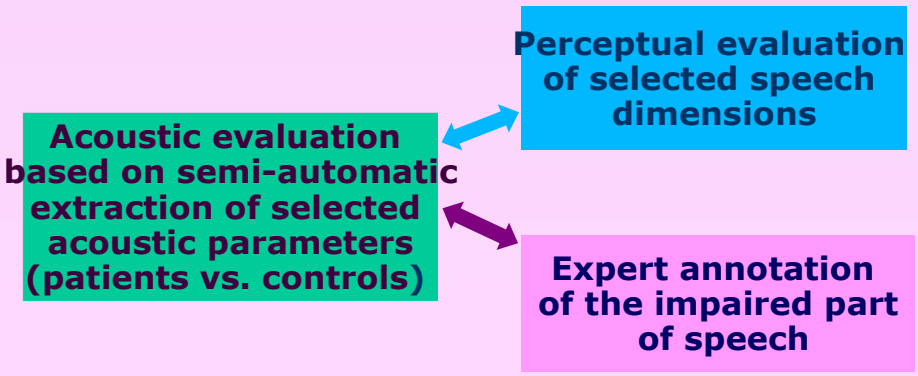
Type of Acoustic parameters

- Measurements linked to the **temporal** organisation at a suprasegmental level,
  - *Speech rate* (phoneme/sec) (mean+stdev)
  - *Phoneme + pause duration* (ms.) (mean+stdev)
- Measurements **correlated** with dysfluencies
  - *False start*
  - *Inappropriate pauses within phrases*

For suprasegmental features
⇒ **automatic tracking difficult to implement**
⇒ **ongoing studies to determine the best parameters**

We intent to screen all the population with all the parameters

Comparison of 3 approaches



Methods tested on a subset of our population

Comparison of 3 approaches

- 10 expert judges
- 35 items judged on a 5 points scale (0=ok, 4=very impaired)
- 1 min of speech per speakers
- 39 speech files: 35+4 for test-retest

Perceptual evaluation of selected speech dimensions

⇒ **About 3h of testing**

BECD *(Auzou & Rolland-Monnoury, 2006):*

Voice quality (12 items)

Phonetic realization (6 items)

Prosody (12 items)

Respiratory (3 items)

Intelligibility (1 item)

Naturalness (1 item)

	QUALITÉ VOCALE
1	Qualité vocale
2	Hauteur (+/-)
3	Intensité (+/-)
4	Voix rauque, enrouée
5	Voix éraillée
6	Voix soufflée, voilée
7	Voix mouillée, humide, gargouillante
8	Voix étranglée, hyperfonctionnelle
9	Tremblements de la voix
10	Voix avec fry
11	Dégradation de l'intensité vocale
12	Périodes d'aphonie
	RÉALISATION PHONÉTIQUE
13	Réalisation phonétique
14	Modification de la durée des phonèmes
15	Imprécision des consonnes
16	Imprécision des voyelles
17	Trouble de la résonance nasale
18	Souffle nasal

	PROSODIE
19	Altération de la prosodie
20	Perte des modulations de la hauteur
21	Fluctuations excessives de la hauteur
22	Perte des modulations de l'intensité vocale
23	Fluctuations excessives de l'intensité vocale
24	Débit (+/-)
25	Fluctuations du débit
26	Accélérations paroxystiques du débit
27	Ruptures de la fluence
28	Silences inappropriés
29	Paillalie
30	Difficultés d'initiation motrice
	RESPIRATION
31	Inspiration-expiration forcées
32	Bruit inspiratoire
33	Contrôle respiratoire
	INTELLIGIBILITÉ
34	Intelligibilité
	CARACTÈRE NATUREL
35	Caractère naturel

Agreement (+/- 1pt) : 93% intra judge; 83% inter judges

Comparison of 3 approaches

- 1 expert (speech pathologist)
- Evaluation of the speech files by ear and eyes

very costly in time!

- Classification of the impaired productions:
  - Inappropriate pauses
  - Voicing/devoicing
  - Phonation problems for vowels
  - Nasalization/denasalization
  - Fortition of fricatives
  - Fricatization of stops
  - Other 'spectral' problems (diphtongs, …)

Expert annotation of the impaired part of speech

Comparison of 3 approaches

Acoustic evaluation based on semi-automatic extraction of selected acoustic parameters (patients vs. controls)

1) Orthographic transcription of the production
2) Automatic alignment of the speech files
3) Semi-automatic extraction of a set of acoustic parameters (w/ Praat scripts)

- Nothing very new at this point (!): most of the parameters implemented have been described in the literature

- In process to select parameters applicable & meaningful for a continuous speech corpora

Pre-Processing of the Audio Files

- In order to be able to perform both the manual and automatic acoustic analyses,

  Need of:

  – Orthographic transcriptions

  – Automatic text-constrained phonetic alignment using the LIA tool box

15

Pre-Processing of the Audio Files:
Orthographic Transcriptions – Automatic alignment

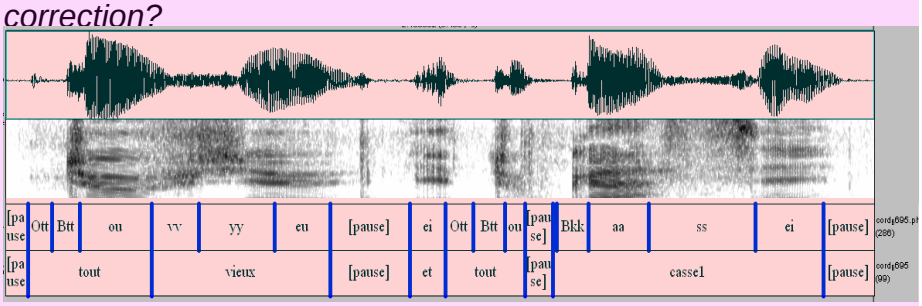
1. Orthographic transcriptions

Each speech production which differs from the original text was annotated using the SAMPA phonetic alphabet

? what level of details, what error to translate, how to translate?

2. Automatic text-constrained phonetic alignment

? What precision for the boundaries, necessity for manual correction?



Automatic Phonetic Alignment

When comparing segments produced manually and automatically…

- disagreement rate = proportion of segments with midpoint shift over 20 ms (our reference value)

	Male	Female
Control speaker	14%	17%
Moderate dysarthria	21%	26%
Severe dysarthria	66%	52%

- but **strong agreement** (manual/automatic) for **spectral** measurements (Formants frequencies…)

17

Where are we now ?

Pre-test on 35 recordings of pathological speech

– 1st manual transcription & automatic alignment

⇒ comparison between manual alignment done by 2 phoneticians on a subset of data

⇒add liaison consonants

⇒add pauses;

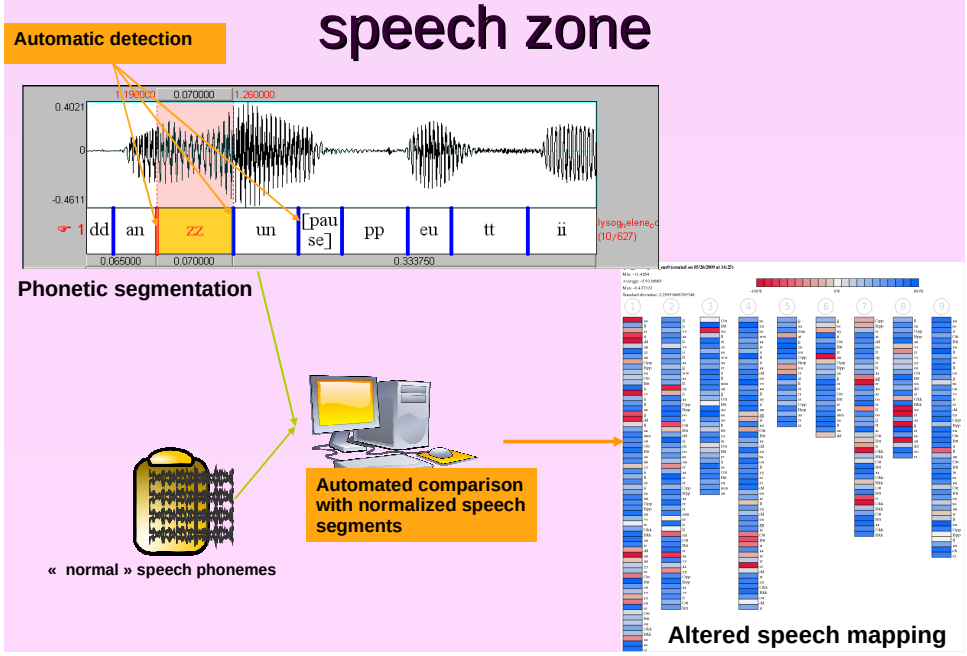
⇒errors on segment length (stop consonants, glides, vowels)

⇒**We have to Improve the procedure**

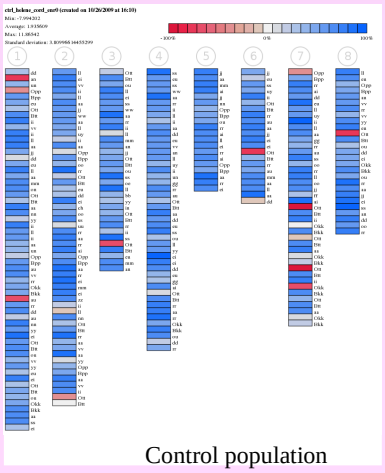
- adapt **transcription conventions**
- adapt the **aligner** LIA:
  - reduce the lexical system we should take into account,
  - remove pronunciation variations

18

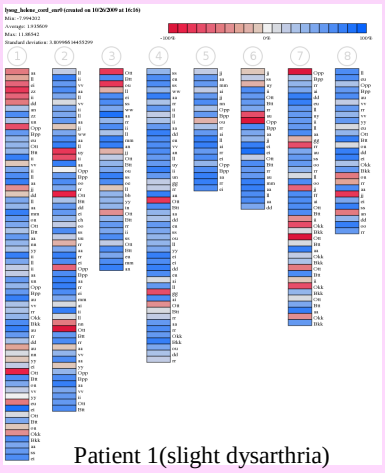
Automatic detection of altered speech zone



Mapping Exemples



Control population



Patient 1(slight dysarthria)

Conclusions/Perspectives

Assumption that an acoustico-phonetic description of dysarthric speech has to be apprehended in a

- **comprehensive fashion according to the multiple dimensions of speech production that can be impaired,**
- **but also by a systematic testing of all potentially informative acoustic criteria for every dysarthric speakers.**

Observing type and range of variations due to motor control deficit can help for a comprehensive model of speech variation for normal speakers.

A better understanding of the variations that permit to consider a dysarthric speech as deviant could provide insights to the boundary between normal and pathological speech.

21

Acknowledgments

Funded by the ANR BLAN08-0125 of the French National Research Agency

We also thank P. Clément, A. Nurenberg and O. Panseri who are also collaborating to this project

For further information, please visit:
http://despho-apady.univ-avignon.fr

22