Automatic detection of syllable boundaries in spontaneous speech

Brigitte Bigi
Christine Meunier, Irina Nesterenko, Roxane Bertrand

Laboratoire Parole et Langage
LREC - May 2010

Main goal

- Automatic detection of syllable boundaries
- Development of a Rule-Based System (RBS) :
  - for automatic syllabification of phonemes' strings
  - of the size greater than a graphic word
  - Example :

Phonemes	i	l	e	k	s	p	l	i	k	e	p	a	v	r	e	m	ã	s	k	i	i	a	v	e	d	ã

- Application to conversational speech
  - 8 French informal dialogues

Main RBS Principles

- The syllabification process is based on 2 main principles :
  1. a syllable contains a vowel, and only one.
  2. a pause is a syllable boundary.
- These two principles bring the problem to find the boundaries between two vowels

Phonemes	**i** l **e** k s p l **i** k **e** p **A** v r **e** m **Ã** s k **i** j **A** v **e** d **Ã**

Related works (1)

- *syllabation.awk*, made by C. Pallier
  - GPL
  - To segment phonetized words into syllables
  - Phonemes are grouped into 4 classes :
    - vowels, glides, liquids and other consonants
  - 9 segmentation rules are established
    - to find the boundary between 2 vowels
    - by using the classes or
    - by using the phonemes directly in cases a class is not relevant
  - Successfully applied to lexical databases : Brulex and Lexique

Related works (2)

- M. Adda-Decker, P. Boula de Mareüil, G. Adda, L. Lamel

  "Investigating syllabic structures and their variation in spontaneous French", SpeechCom, 2005
  - Software not available
  - Part of *graphon+*
  - To segment phonetized spoken French
  - Phonemes are grouped into 4 classes :
    - vowels, glides, liquids, other consonants
  - 13 segmentation rules are established
    - to find the boundary between 2 vowels
    - by using the classes or
    - by using the phonemes directly in cases a class is not relevant

Related works (3)

- *syllabify2.praat*, made by J.P. Goldman
  - GPL
  - Part of *EasyAlign* software
  - To segment phonetized spoken French
  - Phonemes are grouped into 6 classes :
    - silence, vowels, glides, liquids, [p t k b d g f v] and [s ʃ z ʒ m n ŋ ɲ]
  - About 60 segmentation rules are established
    - to find the boundary using classes
    - not specifically between vowels

Group phonemes into 6 classes

V - Vowels : i e ɛ a ɑ ɔ o u y ø œ ə ɛ̃ ɑ̃ ɔ̃ œ̃

G - Glides : j ɥ w

L - Liquids : l ʀ

O - Occlusives : p t k b d g

F - Fricatives : s z ʃ ʒ f v

N - Nasals : m n ŋ ɲ

- Unlike other systems, we divide consonants into 3 classes : O, F, N.

Phonemes	i	l	e	k	s	p	l	i	k	e	p	ɑ	v	ʀ	e	m	ɑ̃	s	k	i	j	ɑ	v	e	d	ɑ̃			
Classes	V	G	V	O	F	O	L	V	O	V	O	V	F	L	V	N	V	F	O	V	G	V	F	V	O	V			

General rules

The letter X to mention one of G, L, O, N or F.

	Observed sequence	Segmentation rule	Examples (French)
1	VV	V.V	poëte : po.ɛt, il y a un : i.ɑ.œ̃
2	VXV	V.XV	limité : li.mi.te, et donc on : e.dɔ̃.kɔ̃
3	VXXV	VX.XV	jardin : ʒɑʀ.dɛ̃, comme ça : kom.sɑ
4	VXXXV	VX.XXV	avec moi : ɑ.vek.mwa
5	VXXXXV	VX.XXXV	il se présentait : il.spʀe.zɑ̃.te
6	VXXXXXV	VXX.XXXV	alors je crois : ɑ.lorʒ.krwɑ

Exception rules

	Observed sequence	Segmentation rule	Examples
1	VXGV	V.XGV	baignoire : be.nwɑʀ, spéciaux : spe.sjo
2	VFLV	V.FLV	découvre : de.ku.vʀə,
3	VOLV	V.OLV	il trouve : i.tʀuv, mais de la : me.dlɑ
4	VFLGV	V.FLGV	effroyable : ef.ʀwɑ.jɑbl
5	VOLGV	V.OLGV	incroyable : ɛ̃.kʀwɑ.jɑbl
6	VOLOV	VOL.OV	connaître tu : ko.netʀ.ty

Discussion

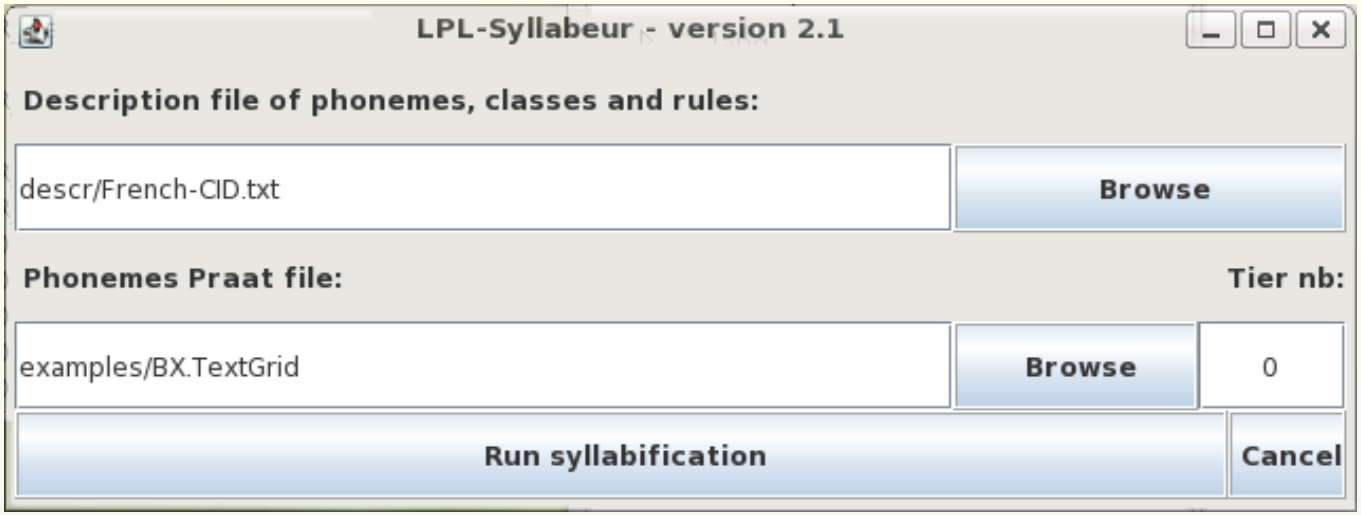
The rules we propose follow usual phonological statements for most of the corpus. Our aim is not to propose a true set of syllabification rules for French, but to provide an acceptable syllabification for the most part of spontaneous speech corpus.

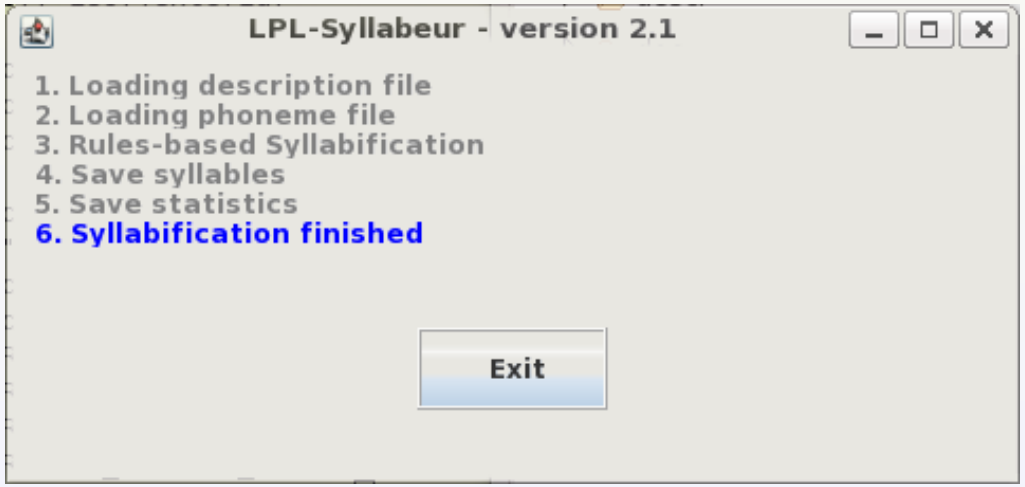
Transcription	il expliquait pas vraiment ce qu'il y avait dedans
Phonemes	i l e k s p l i k e p a v r e m ɑ̃ s k i j a v e d ɑ̃
Classes	V G V O F O L V O V O V F L V N V F O V V V F V O V
Syllables Auto	i . lek . spli . ke . pa . vre . mɑ̃ . ski . ja . ve . dɑ̃
Syllables Expert1	i . lek . spli . ke . pa . vre . mɑ̃ . ski . ja . ve . dɑ̃
Syllables Expert2	i . leks . pli . ke . pa . vre . mɑ̃ . ski . ja . ve . dɑ̃

The LPL-Syllabeur Tool

- Implemented in java 1.6 and tested under linux and windows®
- GPL
- Input and Output in a TextGrid Praat file
- **A configuration file** that the user can change as needed to specify
  - phonemes and classes :
    ```
    PHONCLASS e V
    PHONCLASS p O
    ```
  - general and exception rules :
    ```
    GENRULE VXXV 1
    EXCRULE VFLV 0
    ```
  - some phoneme sequences and a boundary shift to apply :
    ```
    OTHRULE ANY p s k -2
    ```

The LPL-Syllabeur : French or English GUI





Corpus description

- The CID - Corpus of Interactional Data
- Audio-video recording of 8 hours of spontaneous French dialogues
- Each dialogue involves two participants (spoke very freely)
- Phonetization from the transcription

- An Enriched Orthographic Transcription, which includes :
  - Elision, the omission of one or more sounds

    j'ai on a j'ai p- (en)fin j'ai trouvé l(e) meilleur moyen
    c'(é)tait d(e) loger chez des amis
  - Particular phonetic realisations

    [elle,] dormait
    faire des [stats, stateu]
  - The rate of elision and particular realisation is about 17%

Syllabification of the CID

- 139751 vowels = syllables
- Syllables structures are obtained only after applying segmentation rules between 2 vowels :
  - CV 60.70%
  - V 12.95%
  - CVC 11.46%
  - CCV 10.67%
  - CCVC 1.83%
  - VC 1.37%
  - others are less than 1%

$\Rightarrow$ a correct distribution for French

Syllabification Evaluation

- The test corpus is 1.6% of the CID
  - about 7 minutes of a dialogue
  - 2068 syllables
- The test corpus was manually segmented by two experts
  - a syllable agreement rate of 97.77% (23 boundary mismatches)
- Number of boundary mismatches and syllable difference rate :

	syllabation.awk (1)	*graphon+* (2)	*syllabify2.praat* (3)	**LPL-Syllabeur**
Expert 1	74 7.16%	80 7.74%	67 6.48%	**43** 4.16%
Expert 2	84 8.12%	85 8.22%	75 7.25%	**53** 5.13%

Examples

Transcription	et donc on mange sur la baignoire donc c'est c'est ça
Phonemes	e d ɔ̃ k ɔ̃ m ɑ̃ ʒ s y ʀ l a b e n w a ʀ d ɔ̃ k s e s e s a
Classes	V O V O V N V F F V L L V O V N G V L O V O F V F V F V
Syllables (Auto & Experts)	e . dɔ̃ . kɔ̃ . mɑ̃ʒ . syʀ . la . be . nwaʀ . d ɔ̃ k . se . se . sa
Transcription	non dans les parcs c'est un peu limité
Phonemes	n ɔ̃ d ɑ̃ l e p a ʀ k s e t œ̃ p ə l i m i t e
Classes	N V O V L V O V L O F V O V O V L V N V O V
Syllables Auto	nɔ̃ . dɑ̃ . le . paʀ . kse . tœ̃ . pə . li . mi . te
Syllables Experts	nɔ̃ . dɑ̃ . le . paʀk . se . tœ̃ . pə . li . mi . te

Mismatches Experts/Automatic examples

- Transcription : *offre le*
  - Syllables expert1 and expert2 : zof . lə
  - Syllables auto : zo . flə
- Transcription : *comme une*
  - Syllables expert1 and expert2 : kom . yn
  - Syllables auto : ko . myn
  - Syllables expert1 and expert2 : ʀeks . me
  - Syllables auto : ʀek . sme
- Transcription : *glaces comme*
  - Syllables expert1 and expert2 : glɑs . kom
  - Syllables auto : glɑ . skom

Experts are influenced by lexical boundaries

Mismatches Expert1/Expert2/Automatic Examples

In most of the cases, mismatches between the automatic syllabification and the experts' syllabification concern ambiguous boundaries for which experts propose variable syllabification

- Transcription : *retrouver les*
  - Syllables expert1 : ɑʀt . ʀu
  - Syllables expert2 : ɑ . ʀtʀu
  - Syllables auto : ɑʀ . tʀu
- Transcription : *va se faire*
  - Syllables expert1 and auto : vɑs . fɛʀ
  - Syllables expert2 : vɑ . sfɛʀ
- Transcription : *pas le truc*
  - Syllables expert1 and auto : pɑl . tʀyk
  - Syllables expert2 : pɑ . ltʀyk

When there are sequences of more than 2 consonants (not quite frequent in French, though often resulting from reduction phenomena), expert2 pays more attention to the lexical material, while expert1 and our algorithm favours a more balanced structure and respects the sonority principle.

Conclusion

- A rule-based phoneme to syllable segmentation system
- Compared with current state-of-the-art systems, the advantages of the *LPL-Syllabeur* are that :
  1. it is made with a small number of simple rules for the syllabification of a spontaneous French corpus in a friendly dialogue context
  2. the tool uses an object-oriented language, under GPL license
  3. it is very easy to adapt to a specific corpus by adding or modifying rules, phoneme encoding or phoneme classes, by the way of a new configuration file