

La syllabe en français du midi à la lumière des réseaux calculatoires dynamiques

Julien Eychenne

Journées PFC
jeudi 8 décembre



rijksuniversiteit
 groningen

Objectifs

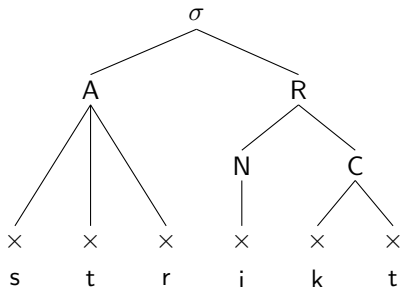
- défendre une conception ondulatoire de la syllababilité
- montrer que l'échelle de sonorité n'est pas nécessaire dans un réseau calculatoire dynamique
- amorcer une réflexion sur un traitement subsymbolique de la phonologie du français méridional (à partir d'une enquête PFC)

Plan

- 1 Introduction
- 2 Les réseaux calculatoires dynamiques
- 3 Corpus
- 4 Résultats

Syllabicit  : perspective syntagmatique

“As far as its internal structure is concerned, the syllable has traditionally been seen as containing an obligatory *nucleus* preceded by an optional consonantal *onset* and followed by an optional consonantal *coda*. . . traditional grammar recognizes an additional subconstituent called the *rhyme* (or *rime*) that includes the nucleus and the coda. The constituents of the traditional syllable are depicted in (2a)”. (Kenstowicz, 1994 : 252-3)



Syllabicit  : perspective ondulaire

“Any phoneme which is louder than the phoneme (or the silence) which precedes, and at the same time louder than the phoneme (or the silence) which follows, is a *crest of sonority* or a *syllabic*; the other phonemes are *non-syllabic*. . . An utterance is said to have as many *syllables* (or *natural syllables*) as it has syllabics. The ups and downs of *syllabification* play an important part in the phonetic structure of all languages”.

(Bloomfield, 1933 : 120-1)

Jack	caught	a	red	bird
dʒæk	kɔ:t	ə	red	bɜ:d
314	414	1	213	313

Critiques de la syllabe

- Angoujard (1997) : la syllabe est le résultat de l'interaction entre la sonorité des segments (calculée sur la base de primitives unaires) et une grille rythmique de type “creux pic (creux)”.
- la phonologie du gouvernement de type CVCV (Scheer, 2004)
- de nombreux phénomènes ont une explication perceptuelle : neutralisation laryngée Steriade (1999), effacement et épenthèse (Côté, 2000)
- Larson (1990) : syllabation en espagnol

La sonorité

Echelle de sonorité (Larson, 1993)

voyelles basses >> voyelles hautes >> glissantes >> liquides >>
nasales >> fricatives voisées >> fricatives non voisées >> occlusives
voisées >> occlusives non voisées

Principe d'intégration de la sonorité (*Sonority Sequencing Principle*)
calcul de la sonorité à partir des traits de classe majeure à la SPE
(Clements, 1990)

3 niveaux de prééminence syllabique dans une perspective ondulatoire :

- H (high) sommet local
- L (low) creux local
- O (other) autre

Ex : *sortir* [sɔʁ.tir] LHOLHL

Plan

- 1 Introduction
- 2 Les réseaux calculatoires dynamiques
- 3 Corpus
- 4 Résultats

Les réseaux calculatoires dynamiques

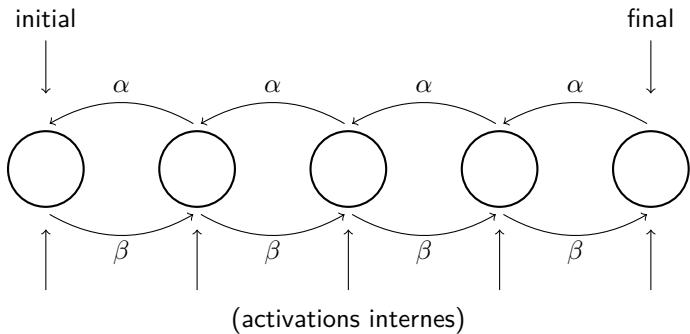
Dynamic Computational Networks (DCN) ou *Dynamic Linear Models (DLM)*

- perceptron unicouche
- excitation/inhibition bilatérale
- Larson (1993) : 3 niveaux
 - Autosegmental (A)
 - Syllabique (S)
 - Métrique (M)
- $A \rightarrow S \rightarrow M$

Références

Goldsmith (1992, 1993); Goldsmith and Larson (1990, 1993); Larson (1990, 1993); Laks (1995, 1997).

Réseau calculatoire dynamique



Mise à jour du réseau

La sonorité dérivée d'une unité à un instant donné est la somme de :

- la sonorité inhérente de l'unité
- la sonorité dérivée de l'unité suivante au temps précédent multipliée par α
- la sonorité dérivée de l'unité précédente au temps précédent multipliée par β

Intuition

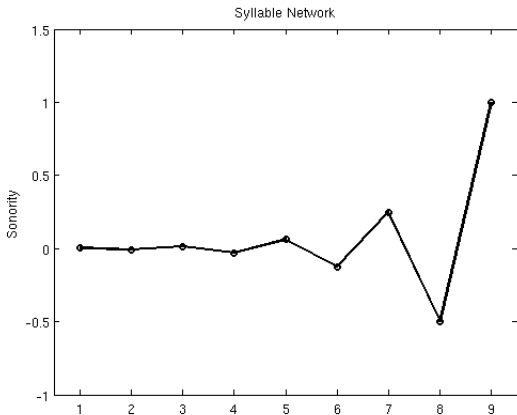
La sonorité dérivée d'une unité est déterminée par sa sonorité inhérente et par la sonorité dérivée de ses voisins, pondérée par les paramètres α et β .

Équation de mise à jour :

$$d_i^{t+1} = u_i + \alpha \cdot d_{i+1}^t + \beta \cdot d_{i-1}^t$$

Illustration

Sonorité dérivée pour l'input (0,0,0,0,0,0,0,1), avec $\alpha = -0,5$ et $\beta = 0$



DCN étendu

Laks (1995) : la syllabation en français

- corpus de 794 mots
- 832 formes syllabifiées
- 99,87% sur un corpus de test
- 6 classes : voyelles, glissantes, occlusives, fricatives, nasales et liquides

Paramètres du modèle :

- $\mathbf{a} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)$
- $\mathbf{b} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)$

Algorithme d'apprentissage

Recuit simulé (*simulated annealing*), voir Larson (1993)

Algorithme stochastique inspiré de la thermodynamique.

Hyper-paramètres du modèle :

- τ : température du système
- $\Delta\tau$: facteur de refroidissement
- θ : seuil de gel
- κ : contrôle l'amplitude de la modification de α et β en cas d'erreur
- λ : contrôle l'amplitude de la modification de la sonorité inhérente en cas d'erreur

Plan

- 1 Introduction
- 2 Les réseaux calculatoires dynamiques
- 3 Corpus**
- 4 Résultats

Le corpus

Enquête Douzens

- 185 minutes de conversation retranscrite (environ trois heures)
- ajout des liste et texte PFC
- 46682 tokens soit 3314 types
- nettoyage et réduction des homophones ayant le même lemme (*arrivé - arrivait*)
- 2530 mots
- phonétisation semi-automatique
- syllabation manuelle

Choix méthodologiques de transcription

- locuteur méridional avec schwa et loi de position
- “voyelles nasales” transcrites comme voyelles orales + N
- groupes r+j tautosyllabiques (ex : intérieur [ɛnterjœr])
- e graphique en syllabe initiale : [ɔ]
- schwa ambigu patelin : [*]

<i>génial</i>	ʒe.pal	CV.CGVC
<i>généalogiques</i>	ʒe.ne.a.lo.ʒi.kə	CV.CV.V.CV.CV.CV
<i>général</i>	ʒe.ne.ral	CV.CV.CVC
<i>générale</i>	ʒe.ne.ra.lə	CV.CV.CV.CV
<i>génération</i>	ʒe.ne.ra.sjɔN	CV.CV.CV.CGVN
<i>génétique</i>	ʒe.ne.ti.kə	CV.CV.CV.CV

Structures syllabiques

6440 syllabes, 27 types syllabiques

type	occurrences	%
CV	3619	56,2
CVN	572	8,9
CVC	552	8,6
CCV	425	6,6
V	383	5,9
CGV	242	3,8
VN	133	2,1
CGVN	131	2,0

Syllabes les plus fréquentes :

/a/, /tə/, /te/, /rə/, /e/, /de/, /maN/, /sə/, /re/, /rø/

Utilisation du corpus dans cette étude

- élimination des mots contenant des hiatus (Larson, 1993 : 99-102)
- 2486 formes
- corpus d'entraînement : 60% des données (1492 mots)
- corpus de validation : 20% (497 mots)
- corpus de test : 20% (497 mots)

Procédure

- conversion des formes lexicales en vecteurs numériques et séquences LOH
- DCN
- lecture de la sonorité dérivée en sortie du DCN
- conversion de la sonorité dérivée en séquence LOH
- comparaison de la séquence prédite et de la séquence attendue

Pour l'apprentissage :

- 1000 simulation
- initialisation aléatoire de α et β à chaque simulation
- multiples présentations du lexique (époques) jusqu'à convergence

Plan

- 1 Introduction
- 2 Les réseaux calculatoires dynamiques
- 3 Corpus
- 4 **Résultats**
 - Dispositif d'apprentissage biaisé
 - Dispositif d'apprentissage non biaisé

Sonorité *a priori*

occlusives non voisées	1
occlusives voisées	2
fricatives non voisées	3
fricatives voisées	4
occlusives nasales	5
liquides	6
glissantes	7
schwa	8
voyelles hautes	9
voyelles moyennes	11
voyelle basse	13

Performance sur 1000 simulations

Corpus de validation :

- converge après 4,8 époques en moyenne
- 98,98% de réussite en moyenne
- 96,78% minimum
- 100% maximum
- paramètres moyens : $\alpha = 0,210659$ et $\beta = -0,015455$

Performance corpus test : 98,59% (7 erreurs)

Performance d'un modèle avec $\alpha = 0$ et $\beta = 0$: 93,56%

Sonorité apprise par le dispositif d'apprentissage biaisé

a : 14,33	ɔ : 12,48	ɛ : 12,26	e : 12,15	o : 11,33
œ : 11,20	ø : 11,11	ə : 10,41	i : 10,33	u : 9,59
y : 9,41	w : 7,75	ɥ : 7,06	j : 7,05	ŋ : 6,36
r : 5,48	l : 5,24	ɲ : 4,90	n : 3,99	ʒ : 3,69
z : 3,47	m : 3,46	v : 3,06	ʃ : 2,79	f : 2,42
g : 1,92	k : 1,73	b : 1,50	s : 1,27	d : 1,24
p : 0,54	t : -0,40			

Dispositif d'apprentissage non biaisé

- converge en 7,8 époques en moyenne
- moyenne sur le corpus de validation de 99,13%
- avec un minimum de 97,38%
- un maximum de 100%
- **performance corpus de test : 98,79%** (6 erreurs)

a : 4,58	ɔ : 3,99	ɛ : 3,84	ə : 3,76	e : 3,62
i : 3,33	o : 2,71	u : 2,67	y : 2,45	œ : 2,44
ø : 2,05	w : 1,98	ʉ : 1,51	j : 1,39	ɳ : 1,12
r : 0,72	ɲ : 0,70	l : 0,60	ʃ : -0,01	ʒ : -0,06
z : -0,31	n : -0,57	g : -0,59	f : -0,81	v : -0,86
k : -0,87	m : -1,02	b : -1,15	s : -1,28	d : -1,37
p : -1,50	t : -2,13			

Comparaison des deux modèles

- biaisé 98,98%
- non biaisé 99,13%
- distributions asymétriques (test de Shapiro-Wilk : $p < 0,001$)
- test de Wilcoxon pour la médiane ($p < 0,001$)

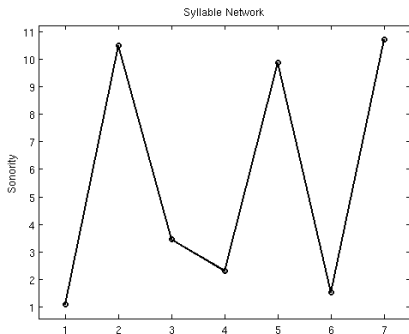
Exemple de syllabification dérivée

discute : /diskytə/

sonorité inhérente = (1.24, 10.33, **1.27**, **1.73**, 9.41, -0.40, 10.41)

sonorité dérivée = (1.08, 10.50, **3.45**, **2.30**, 9.87, 1.52, 10.73)

$\alpha = 0,210659$ et $\beta = -0,015455$



Analyse qualitative des erreurs

mot	syllabation réelle	syllabation prédite
islamique	is.la.mi.kə	i.sla.mi.kə
patchwork	patʃ.wœrk	pat.ʃwœrk
spécialité	spe.sja.li.te	s.pe.sja.li.te
spécialisé	spe.sja.li.ze	s.pe.sja.li.ze
station	sta.sjɔ̃N	s.ta.sjɔ̃N
stage	sta.ʒə	s.ta.ʒə
technologique	tɛk.no.lo.ʒi.kə	tɛ.kno.lo.ʒi.kə

Conclusion et perspectives

- les DCN classiques sont capables d'apprendre la syllabe du français
- une échelle de sonorité précâblée n'est pas un avantage
- la sonorité apprise fait émerger les classes consonnes, voyelles et semi-voyelles (+ appendice nasal)
- questions ouvertes :
 - hiatus
 - le segment comme nœud
- prochaine étape : traitement de l'interaction entre niveaux syllabique et métrique pour le traitement du schwa

Références I

- Angoujard, J.-P. (1997). *Théorie de la syllabe. Rythme et qualité*. CNRS Editions, Paris.
- Bloomfield, L. (1933). *Language*. Allen and Unwin, London. 12^eédition.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In Kingston, J. and Beckman, M. E., editors, *Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech*, pages 283–333. Cambridge University Press.
- Côté, M.-H. (2000). *Consonant Cluster Phonotactics : a perceptual approach*. PhD thesis, MIT.
- Goldsmith, J. (1992). *Connectionism : Theory and Practice*, chapter Local Modelling in Phonology, pages 229–246. Oxford University Press, New York.

Références II

- Goldsmith, J. (1993). Harmonic phonology. In Goldsmith, J., editor, *The Last Phonological Rule*, pages 21–60. The University of Chicago Press, Chicago and London.
- Goldsmith, J. and Larson, G. (1990). Local modeling and syllabification. In *Papers from the 26th Annual Meeting of the Chicago Linguistic Society : Parasession on the syllable in phonetics and phonology*.
- Goldsmith, J. and Larson, G. (1993). Using networks in a harmonic phonology. In *Papers from the 28th Annual Meeting of the Chicago Linguistic Society*, volume 2, page 94125, Chicago. Chicago Linguistics Society.
- Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Blackwell Publishing, Malden/Oxford.
- Laks, B. (1995). A connectionist account of french syllabification. *Lingua*, 95 :51–76.

Références III

- Laks, B. (1997). *Phonologie accentuelle. Métrique, autosegmentalité, constituance*. CNRS Éditions, Paris.
- Larson, G. (1990). Local computational networks and the distribution of segments in the spanish syllable. In *Papers from the 26th Annual Meeting of the Chicago Linguistic Society : Parasession on the syllable in phonetics and phonology*.
- Larson, G. (1993). *Dynamic Computational Networks and the Representation of Phonological Information*. PhD thesis, University of Chicago.
- Prince, A. (1993). In defense of the number *i*. anatomy of a linear dynamical model of linguistic generalizations. Technical report, Rutgers University, Center for Cognitive Science.
- Scheer, T. (2004). *A Lateral Theory of Phonology*, volume 1 : What is CVCV, and why should it be ? Mouton de Gruyter, Berlin.
- Steriade, D. (1999). Phonetics in phonology : the case of laryngeal neutralization. Ms.

Easter egg

Solution analytique d'un DCN

Équation de Prince (forme vectorisée, voir Prince (1993)) :

$$\mathbf{d} = (\mathbf{I} - \mathbf{W}_n)^{-1} \mathbf{u}$$

où \mathbf{d} est le vecteur de sonorité dérivée, \mathbf{u} est le vecteur de sonorité inhérente, \mathbf{I} la matrice d'identité et \mathbf{W} une matrice tridiagonale de dimension $n \times n$ contenant une diagonale α au dessus de la diagonale principale et une diagonale β au dessous.