



Un analyseur morphologique pour l'akkadien

François Barthélemy

`Francois.Barthelemy@inria.fr`

CNAM (Cédric) et INRIA (Atoll)

Introduction

- Travail à but descriptif: description de la morphologie du verbe fort akkadien.
- Réflexion sur le formalisme utilisé.
- Réflexion sur la mécanique sous-jacente (machines finies).
- L'akkadien est un exemple intéressant.
- Dans cet exposé: focalisation sur la notion de partition (Kiraz).

L'akkadien

- langue de l'ancienne mésopotamie (royaume d'Akkad)
- utilisée de -2500 à 0 environ
- deux grands dialectes: babylonien et assyrien
- écriture cunéiforme (syllabique et idéographique)
- nombreux documents (ex: 50.000 au British Museum)

La langue akkadienne

- langue sémitique, branche sémitique oriental

La langue akkadienne

- langue sémitique, branche sémitique oriental
- l'écriture note toutes les voyelles (et même davantage!)

La langue akkadienne

- langue sémitique, branche sémitique oriental
- l'écriture note toutes les voyelles (et même davantage!)
- morphologie verbale riche

Le verbe akkadien

- Forme verbale:
préfixe—base verbale—suffixe
- Base verbale:
racine + consonnes + voyelles

Le verbe akkadien

schème	sans infi xe	infi xe t	infi xe tn
I ou G	forme de base iprus	reciproque réflexive séparative iptaras	habitative iterative iptarras
II ou D	factitive fréquentative uparris	passive uptarris	habitative iterative uptarris
III ou Š	causative factitive ušapris	passive uštapis	iterative uštapis
III/II ou ŠD	identique à III ušparris		
IV ou N	passive pour I réflexive ipparis	 iptaras	itérative ittapas

Le verbe akkadien

	préf.	suff.	vocal. base	cons. préfi x.	cons. infi x.	red. rad. 2
genre	XXX	XXX				
nombre	XXX	XXX				
personne	XXX	XXX				
aspect	XXX	XXX	XXX			XXX
schème	XXX		XXX	XXX		XXX
sous-schème					XXX	
mode		XXX				

Écriture cunéiforme



Écriture cunéiforme

- système mixte: syllabe + idéogrammes + déterminatifs
- système ambigu
 - les mêmes signes utilisés pour les 3 usages
 - un signe note plusieurs syllabe
 - une syllabe notée par plusieurs signes

Écriture cunéiforme

- système mixte: syllabe + idéogrammes + déterminatifs
- système ambigu
 - les mêmes signes utilisés pour les 3 usages
 - un signe note plusieurs syllabe
 - une syllabe notée par plusieurs signes
- contraintes: seules des syllabes V, CV, VC ou CVC sont notées.

Écriture cunéiforme

- système mixte: syllabe + idéogrammes + déterminatifs
- système ambigu
 - les mêmes signes utilisés pour les 3 usages
 - un signe note plusieurs syllabe
 - une syllabe notée par plusieurs signes
- contraintes: seules des syllabes V, CV, VC ou CVC sont notées.

Exemple de caractère

	man	≠ ^u	NIŠ <u>ešrā</u> vingt ;
	mīn	≠	MĪN, MAN <u>šinā</u> deux ; <u>šanū</u> second, autre ^{v*} ;
	niš	≠	<u>šanū</u> faire pour la deuxième fois, changer ^{v*} ;
	nis	+	<u>kilallān</u> les deux ; <u>kašādu</u> atteindre ^v ;
◀	mīm	⊥	gišū (šim) MĪN-DU (ou IM-MĪN-DU) <u>suādu</u> arbruste a-
	mam	≠	-romatique ; (li) MAN-DI-DI (pseudo-id.) <u>mandidu/mādidu</u>
	šar _u	⊥	mesureur [*] ;
	wan	⊥	ŠUŠANA <u>šuššānu</u> un tiers ;
	naš	+	BUZUR ₂ <u>puzru</u> secret [*] (en ass.) ;
	šaruru	⊥	(d) 20 le dieu <u>Šamaš</u> ^{v*} ; <u>šaruru</u> roi ^{v*} (abu MAN-MĒŠ-ni, cf. n° 151)
◀	274	240	(d) 21 le dieu <u>Anu</u> ^v (cf. pa-lāh 21, 50 u 40 qui honore <u>Anu</u> , <u>Enlil</u> , <u>Ea</u>). 471

Travaux de Kiraz (1)

- George Anton Kiraz, travaux sur le syriaque

Travaux de Kiraz (1)

- George Anton Kiraz, travaux sur le syriaque
- livre: *Computational Nonlinear Morphology*, Cambridge University Press (2001)

Travaux de Kiraz (1)

- George Anton Kiraz, travaux sur le syriaque
- livre: *Computational Nonlinear Morphology*, Cambridge University Press (2001)
- extensions de la morphologie à deux niveaux (Koskenniemi)

Travaux de Kiraz (1)

- George Anton Kiraz, travaux sur le syriaque
- livre: *Computational Nonlinear Morphology*, Cambridge University Press (2001)
- extensions de la morphologie à deux niveaux (Koskenniemi)
 - description lexicale utilisant plusieurs dimensions indépendantes se traduisant en des transducteurs multi-rubans.

Travaux de Kiraz (1)

- George Anton Kiraz, travaux sur le syriaque
- livre: *Computational Nonlinear Morphology*, Cambridge University Press (2001)
- extensions de la morphologie à deux niveaux (Koskenniemi)
 - description lexicale utilisant plusieurs dimensions indépendantes se traduisant en des transducteurs multi-rubans.
 - notion de partition pour décrire des correspondances terme à terme de séquences de caractères.

Travaux de Kiraz (2)

Description lexicale à trois dimensions

- la racine consonantique
- le vocalisme
- le schéma consonne/voyelle

Travaux de Kiraz (2)

Description lexicale à trois dimensions

- la racine consonantique
- le vocalisme
- le schéma consonne/voyelle

Apparemment, les partitions à un seul caractère sauf gémination (deux occurrences de la même consonne dans une partition).

Travaux de Kiraz (2)

Exemple:

racine	k	0	t	0	b
vocalisme	0	a	0	e	0
schéma	C1	V	C2 C2	V	C3
surface	k	a	tt	e	b

Travaux de Kiraz (2)

Exemple:

racine	k	0	t	0	b
vocalisme	0	a	0	e	0
schéma	C1	V	C2 C2	V	C3
surface	k	a	tt	e	b

$\langle X, 0, C2 \ C2, X \ X \rangle \Leftrightarrow \text{---}$
[cat=verb, measure=pa"el]

Mécanique opérationnelle

- Transducteurs multi-rubans (description de relations n-aires et non binaires)

Mécanique opérationnelle

- Transducteurs multi-rubans (description de relations n-aires et non binaires)
- Notion de partitionnement: les chaînes mises en relation sont découpées en un nombre égal de segments.

Mécanique opérationnelle

- Transducteurs multi-rubans (description de relations n-aires et non binaires)
- Notion de partitionnement: les chaînes mises en relation sont découpées en un nombre égal de segments.
- pas de synchronisation entre rubans dans une partition.
Synchronisation sur les limites de partition.

Opérations pour transducteurs

- opérations internes sur les transducteurs avec partitions:
 - jointure sur n rubans.
 - projection (oubli de rubans)

Opérations pour transducteurs

- opérations internes sur les transducteurs avec partitions:
 - jointure sur n rubans.
 - projection (oubli de rubans)
- opération externe: composition avec une expression régulière non partitionnée sur un ruban.

Intérêt du *partitionnement* en *akkadien*

Les phénomènes morphologiques sont localisés:

- couleur des voyelles d'appui:
šprus > šuprus, *iptrus > iptarus, *pars >
paris

Intérêt du partitionnement en akkadien

Les phénomènes morphologiques sont localisés:

- couleur des voyelles d'appui:
šprus > šuprus, *iptrus > iptarus, *pars > paris
- place des voyelles d'appui

Intérêt du *partitionnement* en *akkadien*

Les phénomènes morphologiques sont localisés:

- couleur des voyelles d'appui:
šprus > šuprus, *iptrus > iptarus, *pars > paris
- place des voyelles d'appui
- traces des radicales faibles:
elles dépendent de la position de la radicale et de la nature des voyelles plus que de leur couleur.

Intérêt du partitionnement en akkadien

Les phénomènes morphologiques sont localisés:

- couleur des voyelles d'appui:
šprus > šuprus, *iptrus > iptarus, *pars > paris
- place des voyelles d'appui
- traces des radicales faibles:
elles dépendent de la position de la radicale et de la nature des voyelles plus que de leur couleur.
- le redoublement ne concerne que la deuxième radicale

Partitionnement choisi

Sept segments.

- préfixe personnel (figés)
- préfixe de schème (+ infixes et voyelles)
- première radicale (+ infixes et voyelles)
- deuxième radicale (+ gémination et voyelle)
- la troisième radicale
- suffixe personnel (figé)
- suffixe de mode

Partitionnement choisi

Exemples:

segment	1	2	3	4	5	6	7
iprus	i		p	ru	s		
ittanaprasū	i	ntana	p	ra	s	ū	
taptarrasam	ta		pta	rra	s		am

Partitionnement choisi (2)

Partitionnements alternatifs

1. séparer les infixes de la consonne support

Partitionnement choisi (2)

Partitionnements alternatifs

1. séparer les infixes de la consonne support
2. une partition par consonne de la base verbale

Partitionnement choisi (2)

Partitionnements alternatifs

1. séparer les infixes de la consonne support
2. une partition par consonne de la base verbale
3. nombre de partition variant selon les formes

Partitionnement choisi (2)

Partitionnements alternatifs

1. séparer les infixes de la consonne support
2. une partition par consonne de la base verbale
3. nombre de partition variant selon les formes

Caractère hétérogène des partitions.

Partitionnement choisi (2)

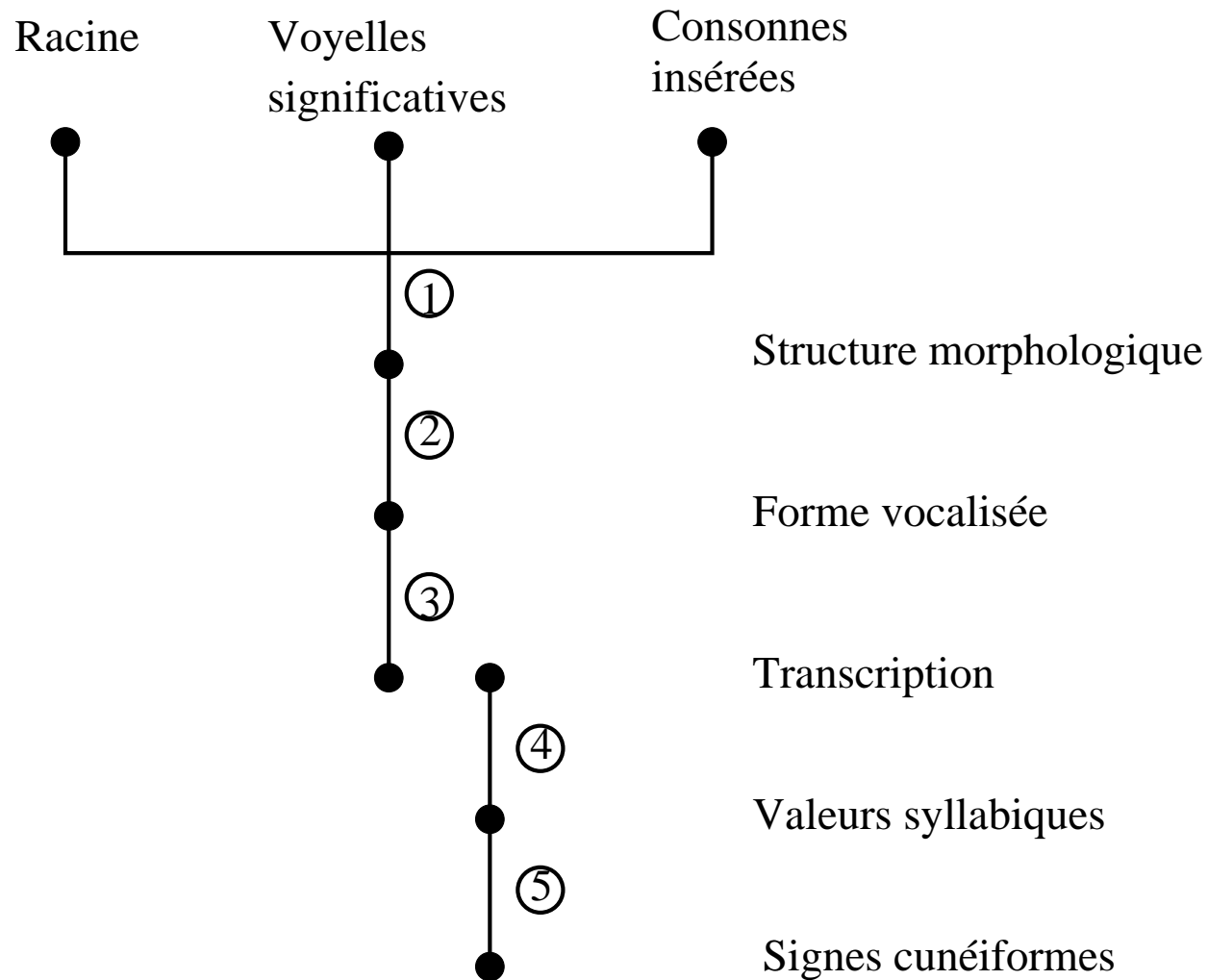
Partitionnements alternatifs

1. séparer les infixes de la consonne support
2. une partition par consonne de la base verbale
3. nombre de partition variant selon les formes

Caractère hétérogène des partitions.

Partitionnement plus ambitieux que celui de Kiraz.

Architecture de l'analyseur



Exemple de forme

segment	1	2	3	4	5	6	7
racine			p	r	s		
voyelle				a			
insertions		nt					
	ta	nt	p	ra	s		am
	ta	nta	p	ra	s		am
	ta	tta	p	ra	s		am
	ta	t	tap	ra	sam		
syllabe	tá	at	tap	rà	sam		
caractère	355	10	124	381	318		

Description des machines

- Expressions régulières locales
- composées par union ou intersection
- environ une trentaine pour la machine 1

Exemple régulières

d'expressions

```
%%% redoublement
```

```
let local C in [Cons]
```

```
<_><_><_><[formeI],_,C,_,CC[Voy]><_><[inaccompli
```

```
<_><_><_><[formeI],_,C,_,C([Voy]?)><_><_><_>
```

```
...
```

```
%%% vocalisation finale
```

```
<_><_><_><[Cons]([Cons]?),[Cons]([Cons]?)i><_><0
```

```
%%% end
```

Partitionnements différents

Il y a deux partitionnements différents:

- un de nature morphologique
- un de nature graphique

Partitionnements différents

Il y a deux partitionnements différents:

- un de nature morphologique
- un de nature graphique

Pas d'opération pour combiner deux transducteurs à partitionnement.

Partitionnements différents

Il y a deux partitionnements différents:

- un de nature morphologique
- un de nature graphique

Pas d'opération pour combiner deux transducteurs à partitionnement.

Pas de machine pour représenter deux partitionnements de la même relation.

Insertion des voyelles

- Machine 2 à deux rubans pour insérer les voyelles d'appui.

Insertion des voyelles

- Machine 2 à deux rubans pour insérer les voyelles d'appui.
- Variation autour de l'identité: les deux rubans sont identiques sauf quelques voyelles.

Insertion des voyelles

- Machine 2 à deux rubans pour insérer les voyelles d'appui.
- Variation autour de l'identité: les deux rubans sont identiques sauf quelques voyelles.
- Difficile à exprimer dans le formalisme actuel: il ne doit pas y avoir de synchronisation au sein d'une partition.

Insertion des voyelles

- Machine 2 à deux rubans pour insérer les voyelles d'appui.
- Variation autour de l'identité: les deux rubans sont identiques sauf quelques voyelles.
- Difficile à exprimer dans le formalisme actuel: il ne doit pas y avoir de synchronisation au sein d'une partition.
- Bricolage: utiliser le fait que la longueur des partitions est bornée (et connue).



Phénomènes de surface

Certains phénomènes de surface (assimilations, par exemple) sont peut-être indépendants de la notion de partition.

On les exprimerait (peut-être) plus facilement si on partait d'une correspondance caractère par caractère, type règle de morphologie à deux niveaux traditionnelle.

Conclusion sur le partitionnement

- Les expressions régulières multi-rubans à partition sont très pratiques et agréables pour exprimer la structure morphologique de l'akkadien.

Conclusion sur le partitionnement

- Les expressions régulières multi-rubans à partition sont très pratiques et agréables pour exprimer la structure morphologique de l'akkadien.
- Il faut pouvoir représenter en même temps plusieurs partitionnements différents, dont un au niveau du caractère.

Conclusion sur le partitionnement

- Les expressions régulières multi-rubans à partition sont très pratiques et agréables pour exprimer la structure morphologique de l'akkadien.
- Il faut pouvoir représenter en même temps plusieurs partitionnements différents, dont un au niveau du caractère.
- L'utilisation des règles à deux niveaux n'est pas le point clé.

Conclusion sur l'analyseur

- Le développement de l'analyseur est subordonné à celui des outils opérationnels.

Conclusion sur l'analyseur

- Le développement de l'analyseur est subordonné à celui des outils opérationnels.
- Actuellement, prototype en perl compilant vers FSM (Mohry, ATT).

Conclusion sur l'analyseur

- Le développement de l'analyseur est subordonné à celui des outils opérationnels.
- Actuellement, prototype en perl compilant vers FSM (Mohry, ATT).
- Actuellement, la machine à 5 rubans fonctionne de façon satisfaisante.

Conclusion sur l'analyseur

- Le développement de l'analyseur est subordonné à celui des outils opérationnels.
- Actuellement, prototype en perl compilant vers FSM (Mohry, ATT).
- Actuellement, la machine à 5 rubans fonctionne de façon satisfaisante.
- Bon espoir de pouvoir décrire les verbes faibles.

Conclusion sur l'analyseur

- Le développement de l'analyseur est subordonné à celui des outils opérationnels.
- Actuellement, prototype en perl compilant vers FSM (Mohry, ATT).
- Actuellement, la machine à 5 rubans fonctionne de façon satisfaisante.
- Bon espoir de pouvoir décrire les verbes faibles.
- Limitation: problème de ressources (lexique des racines et des caractères).