

Quels systèmes de transcription pour représenter les données d'apprenants ?

Elisabeth DELAIS-ROUSSARIE
UMR 7110 / LLF
Laboratoire de Linguistique formelle
Université de Paris-Diderot



Introduction (1)

Pour travailler sur l'acquisition de la dimension orale d'une langue, il est fréquent de s'appuyer sur des enregistrements audio/vidéo.

Mais cela nécessite, préalablement à toute analyse, de représenter le contenu de ces données.

- sous la forme d'une représentation physique,
- ou même sous la forme d'une représentation symbolique abstraite ou transcription.



Introduction (2)

Parmi les représentations abstraites, on peut citer

- Les transcriptions orthographiques;
- Les transcriptions à l'aide de l'API (ou d'une extension comme SAMPA);
- les représentations prosodiques (TOBI, etc.)

Objectif de la communication :

quelle représentation / quels systèmes de transcription pour les faits prosodiques, d'autant qu'il n'y a pas de standard comparable à l'API.



Introduction et plan (3)

L'étude et l'évaluation des systèmes de transcription de la prosodie nécessitent cependant de s'interroger sur la tâche de transcription et sur ce qu'elle suppose.

Plan

- Transcrire : présupposés et problèmes
- les modèles de transcription de la prosodie: description, avantages et limitations
- vers une nouvelle approche



Qu'est-ce que transcrire ? (1)

La pratique et les présupposées

- Définition

La tâche de transcription consiste à représenter symboliquement les éléments linguistiquement pertinents présents dans le signal sonore. Cela se fait généralement par l'assignation d'une étiquette à une unité.

- Cette tâche implique :

- une segmentation du flux en unités;
- La définition d'un ensemble d'étiquettes;



Cela doit être fait de façon à intégrer ce qui peut être linguistiquement pertinent.

Qu'est-ce que transcrire ? (2)

Pour la prosodie, les événements à décrire ont différentes caractéristiques et relèvent de plusieurs domaines.

- Les domaines à représenter:
 - Intonation
 - Accentuation
 - Phrasing
- Les paramètres physiques à discrétiser et à encoder:
 - Durée
 - Intensité
 - Fréquence fondamentale



Qu'est-ce que transcrire ? (3)


- La segmentation:

La segmentation revient à représenter le signal comme une succession d'éléments discrets.

- pour le niveau segmental : mots, segments
- Mais cela est plus délicat pour le niveau prosodique:



Qu'est-ce que transcrire ? (4)

- Les variations des paramètres prosodiques sont continues, donc il est difficile de trouver une unité
 - Les unités retenues doivent permettre d'encoder les faits prosodiques, or, les candidats possibles varient
 - La syllabe pour l'accentuation 
 - Les constituants prosodiques pour le phrasing
 - Les constituants prosodiques ou les points cibles pour l'intonation
- > absence de consensus. La définition de bon nombre de candidats est ou dépendante d'une théorie ou dépendante d'une langue

Qu'est-ce que transcrire (5)

Les étiquettes prennent aussi des formes différentes. Elles peuvent être :




- Des primitives comme les primitives tonales H, B, M ;
- Les variantes allophoniques des primitives ;
- Des étiquettes liées aux catégories fonctionnelles (syllabe accentuée, accent primaire, etc.)

Sources de difficultés :

- En prosodie, les variations graduelles et les changements catégoriels peuvent être liées;
- L'assignation et la définition des étiquettes peut se faire relativement à la forme ou à la fonction.


Qu'est-ce que transcrire ? (6)

Le troisième élément est lié à l'idée d'encoder/ d'annoter ce qui est linguistiquement pertinent :

- Une transcription peut-elle être indépendante de la langue ? 
- Peut-elle être neutre théoriquement.

Pour le segmental, la réponse est non. Dans tous les cas, la transcription se fait au niveau phonologique, ou au moins, suppose une connaissance du système sous-jacent.

Qu'est-ce que transcrire ? (7)

Pour la prosodie, et en tous cas pour les variations intonatives, il est possible de proposer une représentation symbolique sans connaître les catégories phonologiques, ou le fonctionnement phonologique de la langue à transcrire/décrire. 

Avant de montrer comment procéder, on va déjà regarder comment les systèmes les plus utilisés fonctionnent.

Les systèmes existants (1)

Parmi les systèmes qui existent, on va s'arrêter sur quelques systèmes :

- ToBi
- API
- IViE/IVTS
- INTSINT



Plusieurs points vont être vus rapidement:

- Informations/ paramètres encodés
- Niveaux d'encodage : phonologique ou plutôt phonétique


Cela permettra de mettre en avant leurs avantages et inconvénients.

Les systèmes existants (2)

	Intonation	Accentuation	Phrasing	Fo	Durée	Intensité
TOBI	++	-	+	+	-	?
API	+	+	+	+	+	?
IVIE	+	+	-	+	+	?
INTSINT	+	-	-	+	-	-



Les systèmes existants (3)


- Parmi ces systèmes, la plupart suppose généralement une connaissance du système phonologique à transcrire :
 - API, ToBI
- Certains sont liés à un cadre théorique particulier
 - ToBi,
- Certains ont recours à des unités de segmentation dont la définition fait débat : 
 - Sur son statut phonologique (Intsint)
 - Sur son extension (Unité tonale, IP, etc.)

Vers une nouvelle approche

- Les systèmes existants connaissent des limitations
 - Difficultés d'encoder les faits prosodiques dans leur ensemble
 - De nombreux systèmes sont dépendants de la langue à transcrire ou d'une théorie
- Tentative de développer un système automatique qui soit indépendant de la langue et qui soit assez neutre théoriquement



PROSOTRAN (1)

- Choix fondamentaux
 - Système automatique 
 - Encodage au niveau syllabique
 - Encodage symbolique mais visant à représenter l'évolution des paramètres prosodiques dans le temps
- Modalité de fonctionnement
 - Part d'une transcription et d'un alignement
 - Génère un fichier PRAAT pour l'évolution des paramètres

PROSOTRAN (2)

L'assignation des symboles se fait à partir du calcul suivant :



- Pour la fréquence fondamentale
 - Hauteur en tenant compte du registre, après avoir écrêté les 3% de valeurs les plus hautes et les plus basses, et avoir segmenté le registre résultant en 6 zones: N1, N2, N3, N4, N5 et N6
 - Importance de la pente en calculant l'indice de glissando sur le mouvement entre la syllabe cible et celle qui précède

PROSOTRAN (2)

- Pour la durée:
 - Calcul de la durée moyenne de chaque noyau en fonction de la forme ([a], [i], etc.). Assignment d'un indice allant de --- à ++++ en fonction de degré de réduction ou d'allongement. Les syllabes avant pause ne sont pas prises en compte
- Pour l'énergie
 - Calcul sur des bases analogues à ce qui est fait pour la durée

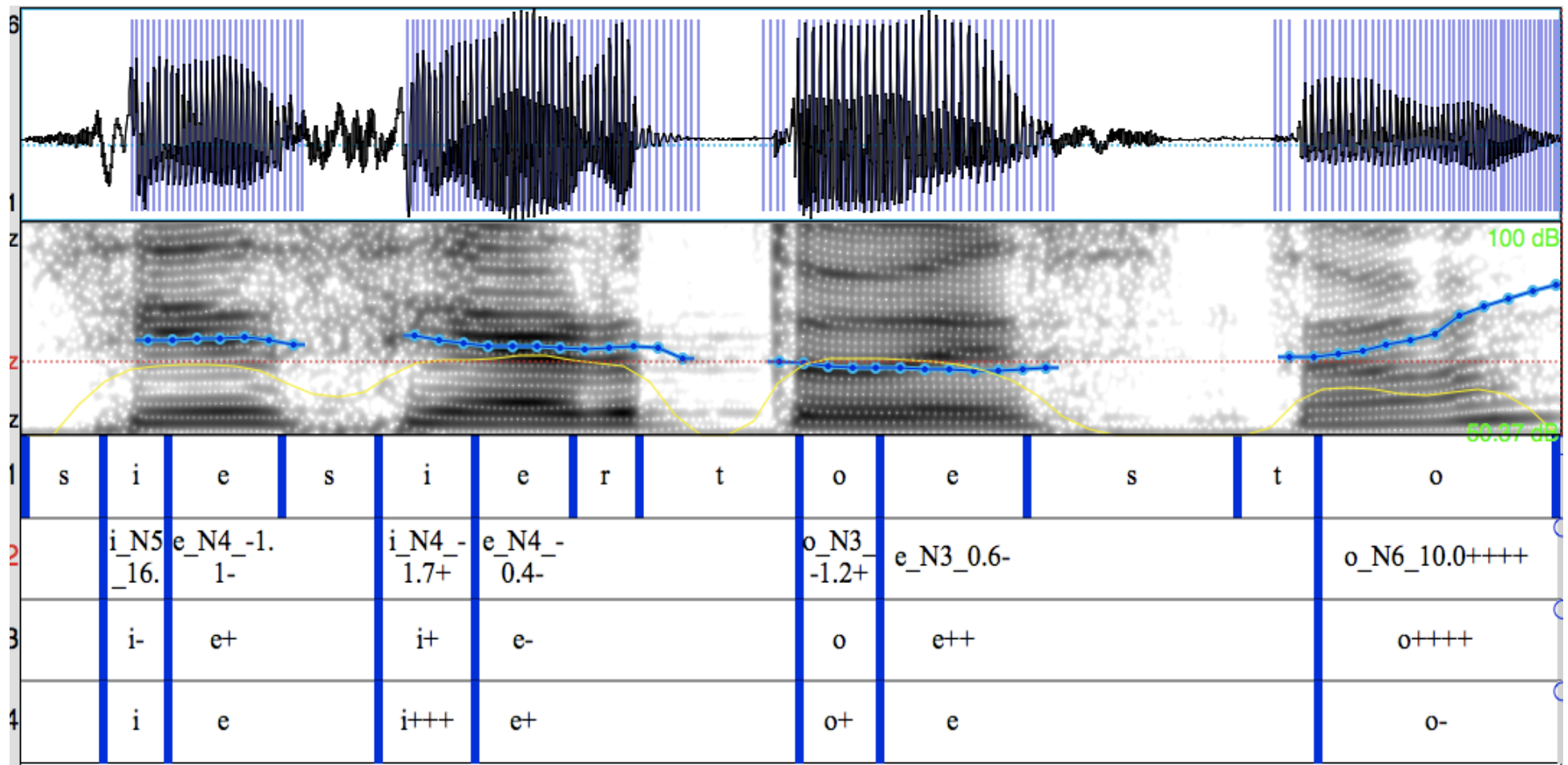


Quelques exemples et quelques résultats


- Le système a été utilisé pour annoter un ensemble de données comprenant des textes lus en français par des locuteurs français, en espagnol par des hispanophones, mais aussi en français par des hispanophones qui apprennent le français
- L'encodage est comparable dans les différents cas



Quelques exemples et quelques résultats (2)



Quelques exemples et quelques résultats (3)

- Permet de faire émerger des régularités qui peuvent affecter
 - Un paramètre 
 - Plusieurs paramètres
- A permis de dégager des traits caractéristiques de l'intonation des questions en français et en espagnol, et de là, de dégager les points problématiques pour les non-natifs

CONCLUSION

- La transcription, bien que nécessaire, doit autant que possible éviter les interprétations, mais elle doit aussi :
 - Être théoriquement neutre,
 - Générer un codage symbolique
 - Permettre de construire les catégories
- Pour la prosodie, il est possible de s'appuyer sur les valeurs associées aux paramètres prosodiques, tout en les croisant avec les valeurs des seuils perceptifs (glissando, etc.)
- Essayer maintenant d'affiner l'outil et de vérifier la validité des hypothèses phonologiques par des tests de perception.

