

# Projets AmbiDYS et XpliciT

Solen Quiniou

`solen.quiniou@univ-nantes.fr`

LS2N - Université de Nantes

Journées scientifiques de l'Université de Nantes - 21/06/2019



# Projet AmbiDYS (12/2016-06/2018)

- Type de projet

- ▶ Projet financé par le RFI OIC avec le LS2N (équipe TALN), le laboratoire READi de l'École de Design et la startup Mobidys

- Public visé

- ▶ Enfants dyslexiques

- Problématiques

- ▶ Détecter des ambiguïtés au niveau du texte dans des livres jeunesse
- ▶ Afficher des aides de lecture associées sans surcharger les livres

- Étapes du projet

- 1 Annotation d'un corpus de référence
- 2 Construction d'un système d'analyse des pronoms anaphoriques sujet (détection, résolution et indication de la difficulté)
- 3 Création d'un nouveau menu d'accès à des aides de lecture
- 4 Évaluation intrinsèque et extrinsèque du système complet

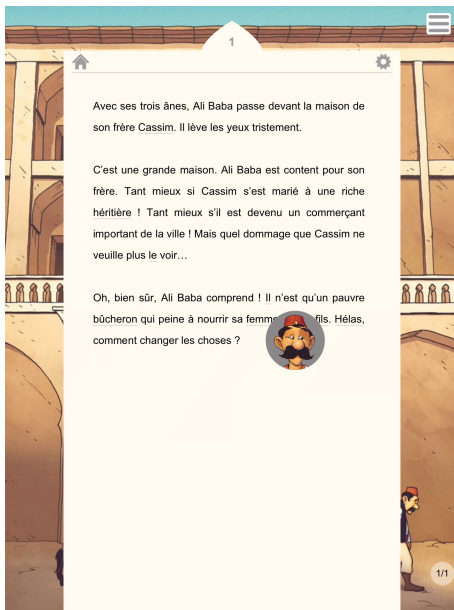
- Application

- ▶ Afficher la vignette du groupe de mots référent lors de l'appui sur un pronom

# Exemple d'utilisation dans les livres de Mobidys (1)



# Exemple d'utilisation dans les livres de Mobidys (2)



# Annotation des anaphores d'un corpus de référence

- Définition

- ▶ **Anaphore** : procédé par lequel un mot ou un groupe de mots rappelle un autre mot ou groupe de mots précédemment énoncé (l'antécédent)
- ▶ **Exemple** : *Oh, bien sûr, Ali Baba comprend ! Il n'est qu'un pauvre bûcheron*

- Annotations effectuées

- ▶ Annotation des pronoms *il* et *elle*
- ▶ Annotation des **anaphores difficiles** (exemples : plusieurs référents possibles, référent qui se trouvent plusieurs phrases avant. . . )
- ▶ Annotations réalisées par une étudiante en orthophonie

Texte	Nb mots	Nb pronoms impersonnels	Nb pronoms anaphoriques	Nb anaphores difficiles
<i>Ali Baba et les quarante voleurs</i>	2 458	7	102	10 (9,8%)
<i>L'arbre et le bûcheron</i>	1 517	9	35	9 (25,7%)
<i>Le buveur d'encre</i>	1 002	10	26	7 (26,9%)
<i>Dans le ventre du cheval de Troie</i>	2 623	20	41	2 (4,9%)
<i>Emporté par le vent</i>	3 005	8	164	18 (11,0%)
<i>Nos étoiles contraires</i>	35 484	17	85	7 (8,2%)

# Système d'analyse des anaphores pronominales sujet

## ● 1. Détection des pronoms *il* impersonnels

- ▶ **Exemple** : *il fait beau*
- ▶ Détection à l'aide de **15 règles** [Danlos, 2005] (expressions régulières)
  - ★ **Exemple de règle** : `word="il" word="ne"? lemme@"meteo"`
  - ★ **Exemple de règle** : `word="il" lemme="ne"? lemme="y" lemme="en"? lemme="avoir"`

## ● 2. Résolution des pronoms *il* et *elle* anaphoriques

- ▶ Utilisation de l'algorithme de [Mitkov, 2002]
- ▶ **Avantages** : nécessite uniquement un étiquetage grammaticales des mots et l'identification des *chunks* (groupes de mots formant une unité de sens)
- ▶ **Principe**
  - 1 Identification des **chunks nominaux** qui précèdent un pronom anaphorique, dans les 2 phrases précédentes
  - 2 Vérification de la **concordance de la flexion** du pronom et des chunks
  - 3 Utilisation de **8 indicateurs pour trier** les chunks (exemple : distance, titre du livre, personnages du livre. . . )
  - 4 Sélection du **chunk de score le plus élevé** comme référent

## ● 3. Détection des anaphores difficiles

- ▶ **Cas considérés**
  - ★ Pronoms avec plusieurs référents candidats ou aucun référent candidat
  - ★ Pronoms dans des dialogues

# Évaluation intrinsèque du système complet

- Calcul du rappel et de la précision par rapport au corpus de référence
  - Résultats dépendants des livres
    - ▶ Bonne détection des pronoms impersonnels
    - ▶ Résolution plutôt bonne des pronoms anaphoriques
      - ★ Erreurs surtout dues aux chaînes anaphoriques
      - ★ Erreurs également dues à une mauvaise détection des groupes nominaux
    - ▶ Détection plutôt bonne des anaphores difficiles
      - ★ Erreurs dues à la mauvaise résolution des anaphores
      - ★ Erreurs également dues à une mauvaise détection des dialogues

Texte	Pronoms impersonnels		Pronoms résolus		Anaphores difficiles	
	R	P	R	P	R	P
<i>Ali Baba. . .</i>	85,7%	100%	88,2%	89,1%	80,0%	71,0%
<i>L'Arbre et le bûcheron</i>	77,8%	100%	74,3%	74,3%	56,0%	56,0%
<i>Le buveur d'encre</i>	90,0%	100%	73,1%	73,1%	75,0%	75,0%
<i>Dans le ventre. . .</i>	90,0%	100%	68,3%	68,3%	56,0%	64,0%
<i>Emporté par le vent</i>	25,0%	100%	72,6%	71,3%	64,0%	64,0%
<i>Nos étoiles contraires</i>	52,9%	100%	49,4%	47,2%	—	—

# Évaluation extrinsèque du système complet

- Étude au niveau de la compréhension

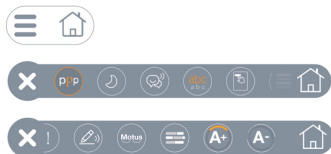
- ▶ Mesure de la difficulté de compréhension des questions sur les anaphores pronominales sujet

- Étude au niveau de l'interface : accessibilité et utilisabilité

- ▶ Menu des aides à la lecture utilisé dans les livres de Mobidys



- ▶ Menu des aides à la lecture revu par le laboratoire READi





# Déroulement de l'évaluation extrinsèque (1)

- 2 interfaces

- ▶ Mobidys
- ▶ Ambidys (version du laboratoire READi)

- 2 populations

- ▶ 8 enfants diagnostiqués DYS, suivis par l'APAJH 44 (Association Pour Adultes et Jeunes Handicapés de Loire-Atlantique) pour troubles multiples, de 9 à 12 ans
- ▶ 23 enfants d'une classe de CE1, de 7 à 8 ans

- Matériel utilisé

- ▶ Livre *Attention fragile* sur tablette numérique



# Déroulement de l'évaluation extrinsèque (2)

## ● Déroulement d'une passation

- ▶ Accueil : court questionnaire démographique et consignes
  - ▶ Lecture version 1 (interface Mobidys ou Ambidys) avec des questions de compréhension en cours de lecture
  - ▶ 6 questions
    - ★ As-tu aimé lire avec la tablette ?
    - ★ As-tu trouvé facilement le menu ?
    - ★ Est-ce que tu trouves bien le dessin des aides ? ...
  - ▶ Lecture version 2 (autre interface) avec les questions de compréhension en cours de lecture
  - ▶ 6 mêmes questions
- Les deux versions d'interface sont **contrebalancées** : la moitié des enfants commence avec l'interface Mobidys et l'autre moitié avec l'interface Ambidys

1. As-tu aimé lire avec la tablette ?



Non pas du tout



Non moyen



Bof...



Oui moyen



Oui beaucoup

2. As-tu trouvé facilement le menu ?



Non pas du tout



Non moyen



Bof...



Oui moyen



Oui beaucoup

# Exemples de questions de compréhension

Chapitre 4 : Q1 – page 1 (après la phrase : « Oh, non ! A-t-il dit d'une voix serrée. »)  
« Maintenant, est-ce que tu peux me dire **à qui ou à quoi correspond « il »** dans la phrase que tu viens de lire ? Choisis l'image qui correspond à « il ». »



Marcel



Cécile



Dylan



Chloé

Chapitre 4 : Q3 – page 3 (après la phrase : « J'ai aussitôt compris ce qu'il voulait faire. »)  
« Maintenant, est-ce que tu peux me dire **à qui ou à quoi correspond « il »** dans la phrase que tu viens de lire ? Choisis l'image qui correspond à « il ». »



Marcel



Frédéric Harmé



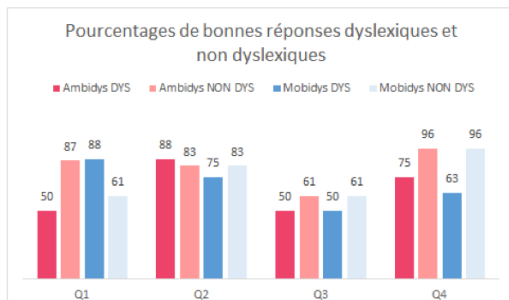
Dylan



Le bulldozer

# Résultats des questions de compréhension

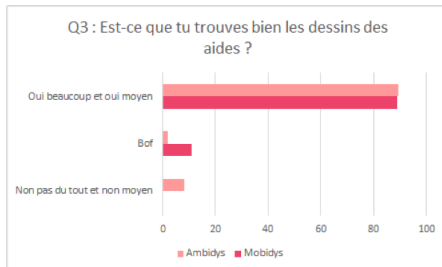
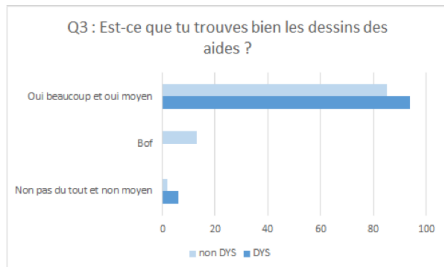
- Questions discriminantes
  - ▶ Pas d'effet plafond (tout le monde aurait répondu juste)
  - ▶ Pas d'effet plancher (tout le monde aurait répondu faux)
- Corrélation entre bons lecteurs et scores de compréhension (pour les élèves de CE1)
- Taux de bonnes réponses plus fiables que temps de réponse
- Interaction entre la population et les interfaces
  - ▶ Résultats un peu meilleurs pour les enfants DYS avec l'interface Mobidys
  - ▶ Résultats meilleurs pour les CE1 avec l'interface Ambidys



# Évaluation du menu des aides de lecture (1)

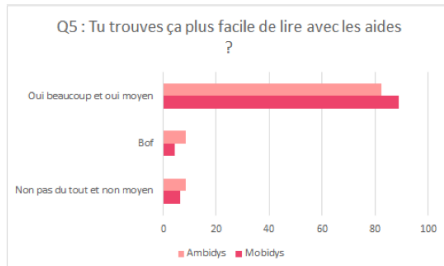
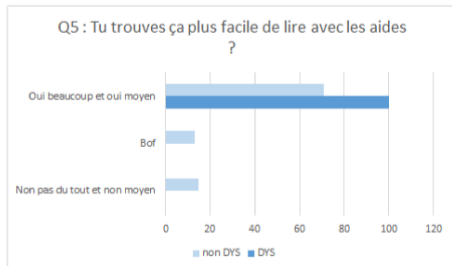
- Données subjectives : résultats des questionnaires

→ Globalement, les enfants apprécient les dessins des menus des aides



# Évaluation du menu des aides de lecture (2)

- Globalement, les enfants de CE1 pensent qu'il est plus facile de lire avec les aides et tous les enfants DYS le pensent aussi
- La version du menu d'aide n'a pas d'effet clairement visible



# Projet XpliciT (12/2018-06/2020)

- Type de projet

- ▶ Projet financé par le RFI OIC avec le LS2N (équipe TALN) et le laboratoire READi de l'École de Design

- Public visé

- ▶ Adultes apprenants du français, notamment demandeurs d'asile et réfugiés

- Problématiques

- ▶ Détecter des ambiguïtés au niveau du texte dans différents types de texte
- ▶ Afficher des aides de lecture associées sans surcharger les textes
- ▶ Travailler sur des smartphones

- Étapes du projet

- 1 Analyse des besoins des demandeurs d'asile et des réfugiés
- 2 Sélection et annotation d'un corpus de référence
- 3 Construction d'un système d'analyse des ambiguïtés (détection et explicitation des ambiguïtés)
- 4 Création d'une interface pour smartphone, pour afficher le texte et les explicitations correspondantes
- 5 Évaluation intrinsèque et extrinsèque du système complet