



RObot de **S**ous-titrage **Et** **T**oute **T**raduction **A**daptés

Comité de Conseillers n°2
29 septembre 2021

Réunion #2

Consortium - Advisory board

Démonstrations des avancées techniques
de chaque partenaire



Locaux de SYSTRAN et Zoom Meeting - 29
Septembre 2021

Bienvenue / Tour de table des participants



Le comité consultatif
ATAA
Unanimes
BUCODES SurdiFrance
FICAM
FNSF

ROSETTA
accès multilingue

Partners and sponsors:

- france•tv access
- SYSTRAN beyond language
- MOCAPLAB
- ÉCOLE PRATIQUE des HAUTES ÉTUDES | PSL
- Lutin Userlab Cité des sciences et de l'industrie
- cnrs
- LISN LABORATOIRE INTERDISCIPLINAIRE DES SCIENCES DU NUMÉRIQUE
- bpi france Co-financier
- cap-digital Labellisation du projet
- technologies data - intelligence artificielle DaIA
- INVESTISSEMENTS D'AVENIR
- Holken Consultants & Partners Sous-traitant valorisation



ROSETTA

accès multilingue

A partir de données
de haute qualité

france•tv
access



Et d'alignements
et annotations



Des outils
automatiques
pour que les
professionnels



puissent rendre leur
contenu multimédia
accessible



Sous-titres
français

Sous-titres
multilingues

france•tv
access

Evalués par
les métiers

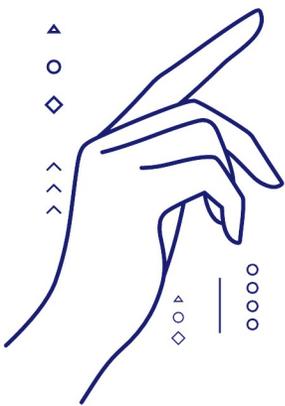


Evalués par un panel
varié d'utilisateurs



Holken Consultants
& Partners

Objectifs



ROSETTA
accès multilingue

Développement de solutions technologiques

O1



Une automatisation de la chaîne de production de sous-titres adaptés et multilingues de contenus audiovisuels

O2

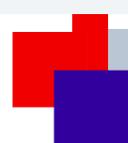


Un pas vers une automatisation de la traduction de contenus journalistiques du français vers la langue des signes française (LSF) via l'animation d'un avatar 3D

O3



Une intégration de ces technologies dans un démonstrateur évalué par les cibles du projet



Objectifs de la réunion / Ordre du jour



Démonstrations individuelles
ROSETTA
Feedback /recommandation
des conseillers

14:00	SYSTRAN HC	Bienvenue Tour de table Rappel des principaux objectifs de ROSETTA Bref retour sur la réunion #1
14:15	LISN	Génération automatique de sous-titres français multi-domaines : divertissement, apprentissage, information <ul style="list-style-type: none">• Démonstration• Feedback / recommandations des conseillers• Enjeux de recherche et d'application
14:45	SYSTRAN	Génération automatique de sous-titres multilingues (anglais, espagnol, italien, portugais) <ul style="list-style-type: none">• Démonstration• Feedback / recommandations des conseillers• Enjeux de recherche et d'application
15:15	SYSTRAN LISN <u>MocapLab</u>	Génération d'animations en LSF au moyen d'un avatar 3D, pour l'annonce de titres de journaux télévisés <ul style="list-style-type: none">• Démonstration• Feedback / recommandations des conseillers• Enjeux de recherche et d'application
15:45	<i><u>Pause Café & Photo de groupe</u></i>	
16:00	<u>france.tv access</u>	FTV Player <ul style="list-style-type: none">• Fonctionnalités actuelles• Développements à venir• Feedback / recommandations des conseillers
16:30	LUTIN	Méthodes d'évaluation <ul style="list-style-type: none">• Questionnaires en ligne• Focus groupes réalisés et à venir• Feedback / recommandations des conseillers
17:00	Discussions avec les conseillers et les observateurs	Réflexions sur la suite et la pérennité des modules du projet Possibilités d'applications concrètes Possibilités de dissémination
17:15	<u>Systran, HC</u>	Conclusions de la journée et prochaines étapes de l'AB
17:30	<i>Fin de la réunion</i>	

Bref retour sur la réunion #1



Perspectives des échanges et présence sur le site

- Rôle des conseillers et observateurs
- Etat de la situation
- Perceptions à propos de la présence de logos sur le site ROSETTA

- Être informés des avancées
- Donner des avis en fonction des intérêts des communautés
- Présence sur le site : valorisation des conseillers
 - Prévoir quelques lignes sur la raison d'être de chaque institution
 - Mettre un lien vers les sites respectif
 - Aucune obligation d'adhésion



1. Génération automatique de sous-titres français multi-domaines : divertissement, apprentissage, information



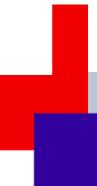
Annelies Braffort

(François Yvon, François Buet)

• 14h15 - 14h45



- Démonstration
- Feedback / recommandations des conseillers
- Enjeux de recherche et d'application



Les sous-titres doivent être faciles à comprendre



« *Il est 20 heures, à la une de l'actualité ce soir* »



- À la une, ce soir

↪ *Simplification de la parole*



Les sous-titres doivent être faciles à lire



tout en étant synchrones avec le son et l'image



CPS < 15



CPL < 36

↳ *Contraintes spatiales et temporelles*

Les genres d'émissions introduisent une variabilité dans le sous-titrage



Direct



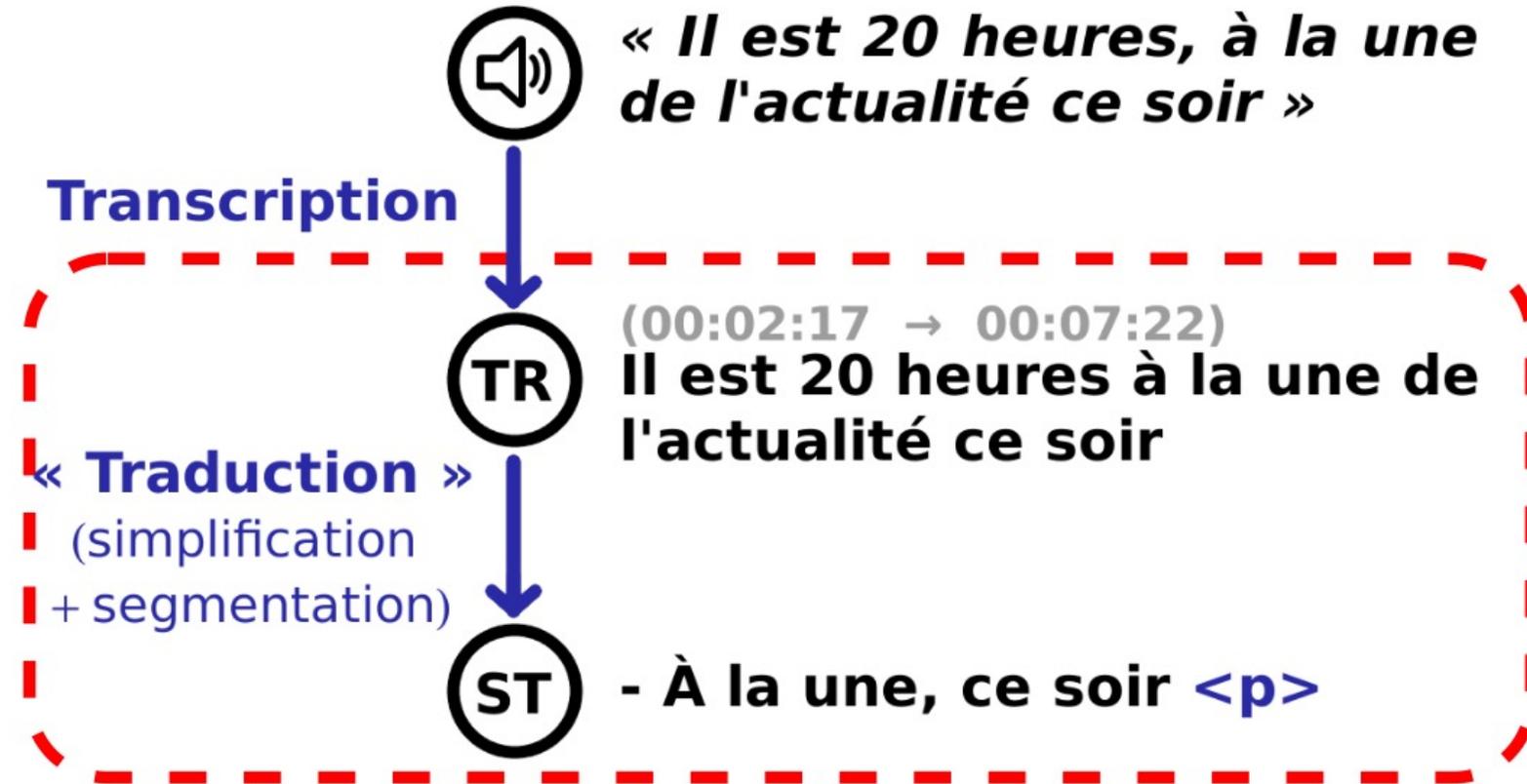
Stock



- sous-titres en direct / stock
- discours préparé / spontané
- format (magazine, jeux...)
- thèmes (cuisine, santé...)

↪ *Adaptation multi-genre*

Le sous-titrage comme traduction monolingue



Récolte et traitement des données France.tv access



API france.tv access



Transcription
automatique

ST

TR

Alignement
automatique

```
<span tts:textAlign="center" tts:color="white" style="background-color: yellow;">-A votre passage dans l'émission,</span>  
<br />  
<span tts:textAlign="center" tts:color="white" style="background-color: yellow;">comment les élèves avaient réagi ?</span>  
</p>  
<p xml:id="caption84" ttm:role="caption" begin="00:04:32:21" end="00:04:36:11">  
<span tts:textAlign="center" tts:color="white">-Ca m'amène à une anecdote.</span>  
</p>
```

```
ne="265.51" dur="0.24" conf="0.982"> plafonds </Word>  
ne="265.87" dur="0.24" conf="0.985"> parfois </Word>  
<Word stime="266.11" dur="0.00" conf="0.999"> . </Word>  
</SpeechSegment>  
<SpeechSegment ch="1" sconf="1.00" stime="266.78" etime="272.35" spkid="MS8">  
<Word stime="267.08" dur="0.84" conf="0.988"> Suite </Word>  
<Word stime="268.19" dur="0.15" conf="0.988"> à </Word>  
<Word stime="268.34" dur="0.30" conf="0.988"> votre </Word>  
<Word stime="268.64" dur="0.33" conf="0.978"> passage </Word>  
<Word stime="269.21" dur="0.18" conf="0.988"> dans </Word>  
<Word stime="269.45" dur="0.57" conf="0.988"> émission </Word>  
<Word stime="270.11" dur="0.27" conf="0.988"> comment </Word>  
<Word stime="270.38" dur="0.39" conf="0.988"> les </Word>  
<Word stime="271.08" dur="0.26" conf="0.973"> élèves </Word>  
<Word stime="271.38" dur="0.05" conf="0.927"> à </Word>  
<Word stime="271.43" dur="0.03" conf="0.988"> l' </Word>  
<Word stime="271.50" dur="0.16" conf="0.988"> époque </Word>  
<Word stime="271.73" dur="0.12" conf="0.582"> avait </Word>  
<Word stime="271.85" dur="0.31" conf="0.947"> réagi </Word>  
<Word stime="272.16" dur="0.00" conf="0.999"> . </Word>  
</SpeechSegment>  
<SpeechSegment ch="1" sconf="1.00" stime="272.35" etime="308.01" spkid="MS7">  
<Word stime="272.42" dur="0.22" conf="0.753"> Ah </Word>  
<Word stime="272.72" dur="0.31" conf="0.984"> bien </Word>
```

Un corpus pour le sous-titrage constitué à partir d'émissions variées



TR Suite à votre passage dans l'émission comment les élèves à l'époque avait réagi.
ST A votre passage dans l'émission, `
` comment les élèves avaient réagi ? `<p>`



Volume de données alignées (juillet 2021) :

- 2 300 heures de vidéo,
- 600 000 paires de phrases.

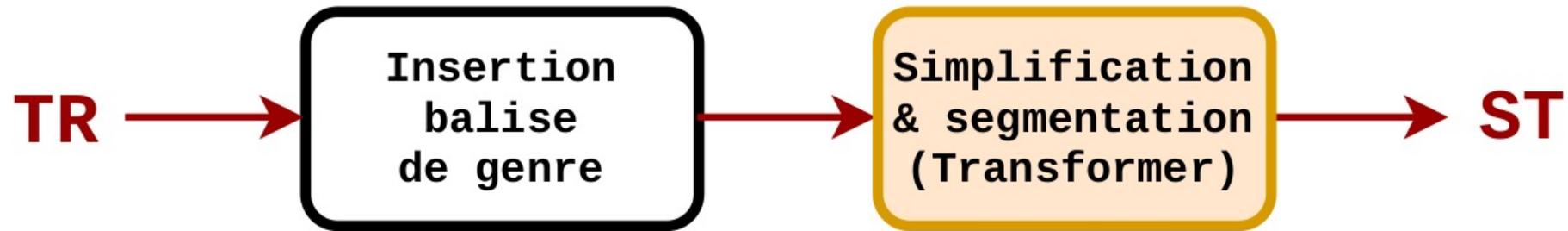
Volume de données non-alignées (pas de transcription) additionnelles :

- 1 300 heures de vidéo,
- 300 000 phrases.

Traduction vers ST : systèmes adaptés au genre



Transformer + frontières + balise de genre :



Transformer + frontières + contrôle de longueur :



Adaptation au genre via des balises ou des objectifs de longueur



Systemes indifferents au genre :



Balises :

∅

Consignes de longueur :

CpR = 0,75

CPS = 14,5

Systemes adaptes au genre :



Balises :

<direct>

<direct>

<stock>

<stock>

<journal>

<mag>

<jeu>

<serie>

Consignes de longueur :

CpR = 0,86

CpR = 0,70

CpR = 0,69

CpR = 0,86

CPS = 17,2

CPS = 14,5

CPS = 12,5

CPS = 12,2

Performances : point de vue de l'évaluation métier (en cours)



Systemes	↑orthographe	↑contresens	↑moyenne
<i>Systemes indifférents au genre (février 2021)</i>			
+ front.	2,18	2,63	2,84
+ front. + RT	2,33	2,57	2,81
<i>Systemes adaptés au genre (février 2021)</i>			
+ front. + balise	2,33	2,71	2,86
+ front. + RT + balise	2,44	2,70	2,89
<i>Systemes adaptés au genre (juillet 2021)</i>			
+ front. + balise	2,46	2,90	2,92

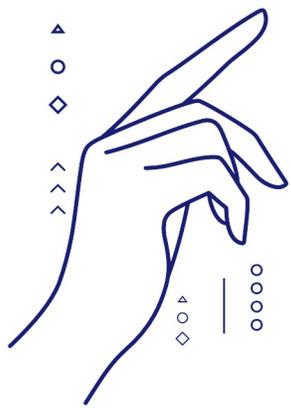
- 6 critères notés sur 5 par des sous-titres professionnels : orthographe, contresens, lisibilité, mots manquants, ponctuation, césure.
- Évaluation sur des émissions pour lesquelles TR était de meilleure qualité...
- ... malgré cela les sous-titres produits restent encore loin d'être acceptables.

Exemples de sous-titres produits



TR	Suite à votre passage dans l'émission comment les élèves à l'époque avait réagi .
Transf.	Suite à votre passage dans l'émission , comment les élèves <p> avaient réagi ? <p>
LRPE	Suite à votre passage dans l'émission . <p> Comment? <p>
ST	A votre passage dans l'émission , comment les élèves avaient réagi ? <p>
TR	Donc nous on monte les mâts oui mais bon, on va laisser finir .
Transf.	On monte les mâts. <p>
LRPE	On monte les mâts. <p> On va laisser finir. <p>
ST	Nous , on monte les mâts. <p> Ah oui . <p> Stop . On va les laisser finir. <p>
TR	En Alabama, les anti-avortement sont majoritaires, 70 pourcent des habitants se déclarent pour une interdiction de l'IVG .
Transf.	En Alabama, les anti-avortement sont majoritaires . <p> 70 % des habitants se déclarent pour une interdiction de l'IVG . <p>
LRPE	En Alabama, les anti-avortement sont majoritaires . <p> 70 % des habitants se déclarent pour une IVG . <p>
ST	En Alabama, les anti-avortement sont majoritaires . <p> 70% des habitants se déclarent pour une interdiction de l'IVG . <p>

2. Génération automatique de sous-titres multilingues (anglais, espagnol, italien, portugais)



ROSETTA
accès multilingue

Elise Michon

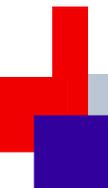
• 14h45 - 15h15



SYSTRAN
beyond language



- Démonstration
- Feedback / recommandations des conseillers
- Enjeux de recherche et d'application



Réalisations

V1 – Preuve de concept interne Modèles de traduction de ST Août 2020

- Système unique et multilingue, du français vers toutes les autres langues
- Apprentissage d'une **segmentation** non calquée sur la langue source mais **adaptée à la langue cible**

V2 - Modèles de génération de ST à partir de la reconnaissance vocale Mai 2021

- Système multilingue
- Apprentissage d'une segmentation en cible **en l'absence de segmentation en source**

V3 – Modèles de traduction de ST basée sur le contexte Juillet 2021

- Système monolingue, du français vers chaque langue séparément
- **Prise en compte du contexte** pour la traduction
- Amélioration de l'intégration (**lisibilité**)



AMELIORATION DES MODELES



Forces et faiblesses de la reco auto



- certaines phrases ont un volume trop faible pour être détectées
- reconnaissance des noms propres
- homophones

- fidélité au discours :

sound : - Bah, je te dépose avant d'aller à l'hôpital, non ?

référence : -Je te dépose.

v1 : - I'm dropping you.

v2 : I'll drop you before I go to the hospital.

v3 : -I'll challenge you.

Prise en compte du contexte



référence : Audrey va pouvoir avancer son enquête.

v1 : Audrey is going to be able to advance his investigation.

v3 : Audrey's going to be able to roll over her investigation.

référence : Il faut qu'on l'aide.

v1 : We need to help him.

v3 : We need to help her.

Contresens et traductions littérales



référence : Elle a bien fait de s'entêter :

v1 : She did the right thing to survive.

v2 : And I'm doing it, though, and I'm doing it.

v3 : She did well.

référence : Le coronavirus dégage le ciel de Chine

v1: The coronavirus puts the sky out of China.

v2: The coronavirus takes off the sky out of China.

v3: The coronavirus clearing the skies of China.

Evaluations

V1 – Preuve de concept interne
Modèles de traduction de ST
Août 2020

V2 - Modèles de génération de ST
à partir de la reconnaissance vocale
Mai 2021

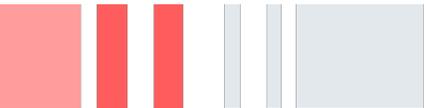
V3 – Modèles de traduction de ST
basée sur le contexte
Juillet 2021

3 DOMAINES / USAGES

1. Série - se divertir
2. Journal TV – s'informer
3. MOOC - apprendre

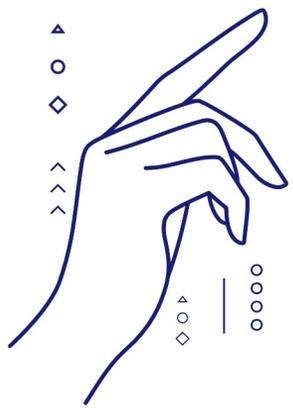
3 TYPES D'EVALUATION

1. **Evaluation interne** : calcul de métriques automatiques dédiées pour mesurer les progrès : **qualité de traduction, qualité de segmentation, longueur, lisibilité**
2. **Evaluation métier** : par l'équipe de sous-titres professionnels par échantillonnage
3. **Evaluation utilisateurs** : focus groupes et questionnaire en ligne en cours



RAPPORT D'EVALUATION &
INTEGRATION DANS LE DEMONSTRATEUR

3. Génération d'animations en LSF au moyen d'un avatar 3D, pour l'annonce de titres de journaux télévisés



ROSETTA
accès multilingue

Boris Dauriac

• 15h15 - 15h45

- Démonstration
- Feedback / recommandations des conseillers
- Enjeux de recherche et d'application

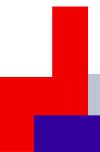


SYSTRAN
beyond language

LISN
LABORATOIRE INTERDISCIPLINAIRE
DES SCIENCES DU NUMÉRIQUE



MOCAPLAB



3. Génération d'animations en LSF

:Contexte



Entrée :

Texte français
(sous-titre?)

Structure de la
phrase (AZee)

Corpus
motion
capture



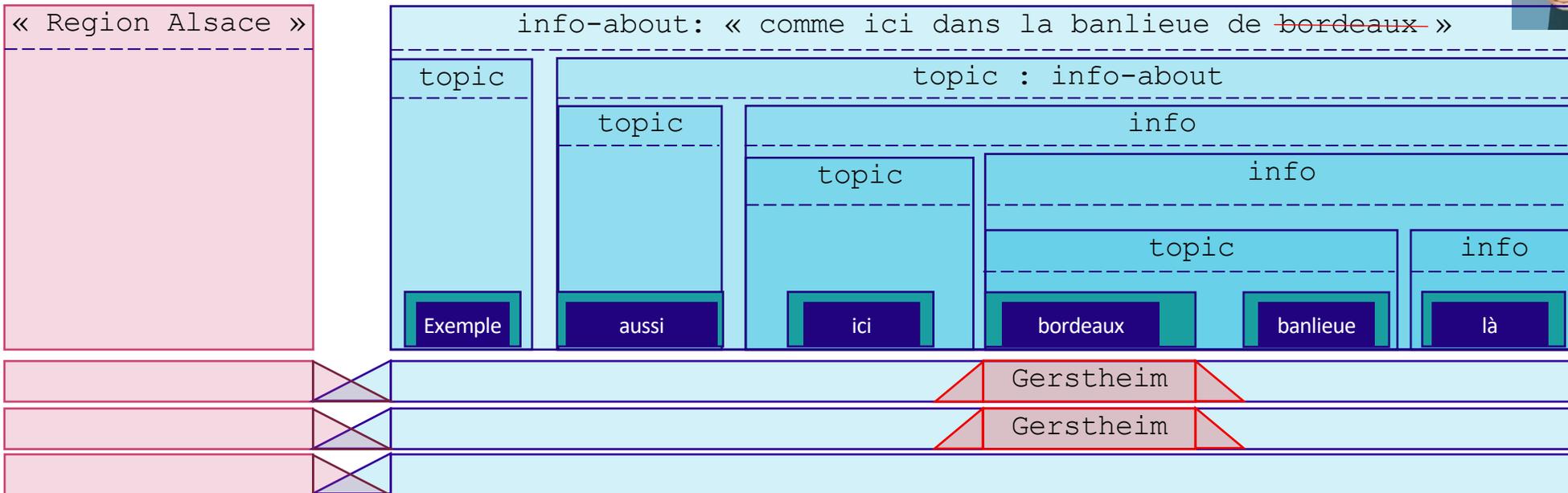
Méthodologie
de
combinaison

Sortie :

Vidéo d'avatar
signant en LSF



Gerstheim



3. Génération d'animations en LSF : Démonstration 1



3. Génération d'animations en LSF : Démonstration 2



François Bouchon / Le Figaro



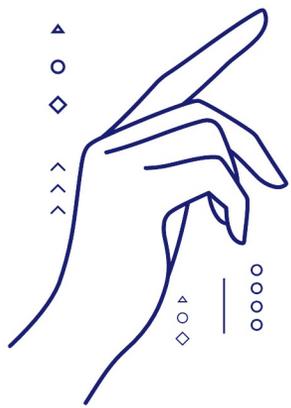


15h45 - 16h

Pause café & Photo de groupe



4. FTV Player

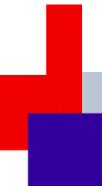


ROSETTA
accès multilingue

Rémy Dal Molin
france•tvaccess

• 16h – 16h30

- Fonctionnalités actuelles
- Développements à venir
- Feedback / recommandations des conseillers



Historique du projet et objectifs



Un premier player labélisé Accessiweb argent en 2014



- Développé en HTML 5
- Open source
- Ouvert aux contributions via [Github](#)
- Nos objectifs
 - *de montrer concrètement quel niveau d'accessibilité est atteignable avec un outil approprié*
 - *de traduire dans les faits notre engagement en faveur de l'accessibilité*
 - *De communiquer autour de l'accessibilité avec pour ambition la dissémination du player auprès d'institutions culturelles, éducatives, de savoir, etc.*

Fonctionnalités actuelles

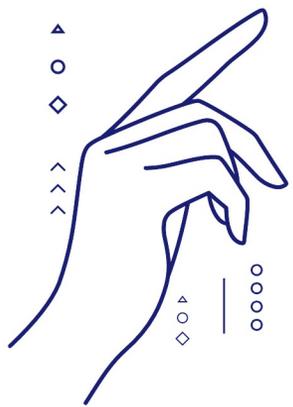


Fonctionnalités actuelles

- Sous-titres
- Audiodescription
- LSF
- Transcription

Cliquez pour ajouter du
texte

5. Méthodes d'évaluation



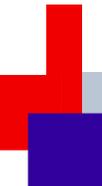
ROSETTA
accès multilingue

Charles Tijus
Taïna Victor
Léa Lachaud

Lutin Userlab
Cité des sciences et de l'industrie

• 16h30 - 17h

- Questionnaires en ligne
- Focus groupes réalisés et à venir
- Feedback / recommandations des conseillers



Focus groupes réalisés phase 1



La première phase de réalisation des focus groupes a eu lieu de janvier à juillet 2021.

Dates	Horaire	Langue naturelle	Recrutement	Matériel de test
07-janv	16h-17h	Français	Adultes tout venant	ST FR Humain
28-janv	14h30 - 15h30	Espagnol	Apprenants Français	ST FR Humain
04-févr	16h-17h	Français	Apprenants Anglais	ST ENG Humain
11-févr	16h-17h	Français	Seniors FR	ST FR Humain
4-mars	16h-17h	Chinois	Apprenants Français	ST FR Humain
25-mars	16h-17h	LSF	SME français courant	LSF
25-mars	18h-19h	LSF	SME français courant	ST FR Humain
15-avril	12h45-13h45	Français	Malentendants	ST FR Humain
15-avril	16h-17h	Français	Apprenants Espagnol	ST ESP Humain
5-juillet	18h-19h	Français	Adultes non SME LSF courant	LSF

Résultats-clés focus groupes - phase 1



Par ordre croissant, les participants ont accordé plus d'importance à :

- **La compréhensibilité (31,5% des avis)**
 - *Les sous-titres seuls/le signeur virtuel seul permettent-ils de comprendre le contenu ?*
- **La fidélité (29,5%)**
 - *Les sous-titres correspondent-ils à l'audio ? Le signeur virtuel exprime-t-il la phrase initialement prévue ?*
- **La visibilité (23,3%)**
 - *Les sous-titres/le signeur virtuel ainsi présentés permettent-ils de saisir le message ?*
- **La lisibilité (15,8%)**
 - *Les sous-titres/le signeur virtuel ainsi présenté sont-ils facilement lisibles ? Confortables ?*

Recommandations des participants aux focus groupes - phase 1



Par ordre croissant, les participants ont émis les recommandations suivantes :

- Réduire le **décalage** entre l'audio et l'apparition des sous-titres (49 commentaires)
- Afficher une **punctuation rigoureuse** afin de transmettre le ton de la voix, l'émotion et l'intention du personnage (24 commentaires)
- Dans les JT, **différencier** les gros titres des sous-titres par une couleur/un graphisme différent. Ou délimiter les zones d'affichage de chacun par des espaces (6 commentaires)
- **Regrouper les sous-titres segmentés** de sorte qu'ils affichent l'ensemble de la phrase en une seule lecture (6 commentaires)
- Porter toute son attention à la **syntaxe** de la **phrase signée** car la **compréhension** en est très fortement impactée (6 commentaires)

Recommandations pour la réalisation des focus groupes - phase 2



Pour les prochains focus groupes, nous suggérons :

- Animer le focus groupes en proposant des **questions semi-directives** afin d'aborder les différentes dimensions ergonomiques (*accessibilité, acceptabilité, utilité, utilisabilité*) avec les participants
- Tester le matériel le plus abouti au moment du focus groupe
- Se servir des personae pour prendre en compte les besoins des différents utilisateurs cibles

Calendrier prévisionnel focus groupes -phase 2



La seconde phase des focus groupes a lieu de septembre à octobre 2021.

Dates	Horaire	Langue naturelle	Recrutement	Matériel de test
23 sept	18h – 19h	Français	Adultes tout venant	ST FR ROSETTA
		Espagnol	Apprenants Français	ST FR ROSETTA
		Français	Apprenants Anglais	ST ENG ROSETTA
16 sept	18h – 19h	Français	Seniors FR	ST FR ROSETTA
		Chinois	Apprenants Français	ST FR ROSETTA
		LSF	SME français courant	LSF
		LSF	SME français courant	ST FR ROSETTA
30 sept	12h-13h	Français	Malentendants	ST FR ROSETTA
		Français	Apprenants Espagnol	ST ESP ROSETTA
		Français	Adultes non SME LSF courant	LSF
		Anglais	Apprenants Français	ST FR ROSETTA
		Français	Apprenants Chinois	ST ZH ROSETTA

6. Discussion



Discussion avec les conseillers et les observateurs

17h - 17h15

- Réflexions sur la suite et la pérennité du projet
- Possibilités d'applications concrètes
- Possibilités de dissémination



Conclusion

17h15 – 17h30

- Conclusion de la journée
- Prochaines étapes
- Réunion #3 : novembre 2021

Merci !



Elise MICHON, coordination ROSETTA elise.michon@systrangroup.com
Hadmut HOLKEN, suivi du comité consultatif holken@holkenconsultants.com