



RObot de S**o**us-titrage Et T**o**ute T**r**aduction A**d**aptés

Comité de Conseillers n°3 - Partie 2/2 : Sous-titrage
23 novembre 2021 - Après-midi



Réunion #3

Consortium - Advisory board

Démonstrateurs finaux, ouvertures et perspectives

france•tv
access

Hôtel Courtyard Paris Porte de Versailles + Zoom Meeting

23 Novembre 2021

Bienvenue après-midi / Tour de table avec les nouveaux participants



Le comité consultatif

ATAA

Unanimes

BUCODES SurdiFrance

FICAM

FNSF

ROSETTA
accès multilingue

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

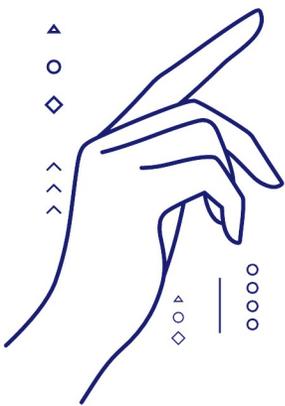
Objectifs de la réunion / Ordre du jour 2



Après-midi
consacrée
à la génération automatique
des sous-titres

Après-midi consacré à la génération automatique des sous-titres		
14:30	SYSTRAN	Bienvenue Tour de table Rappel des principaux objectifs de ROSETTA – partie ST
14:45	LISN – france.tv access	Génération automatique de sous-titres français multi-domaines : divertissement, apprentissage, information <ul style="list-style-type: none">• Enjeux• Progrès et limites au cours du projet• Démonstration des systèmes en action• Analyse des résultats de l'évaluation métier en regard de ceux des métriques automatiques• Perspectives de recherche
15:15	SYSTRAN – france.tv access	Génération automatique de sous-titres multilingues (anglais, espagnol, italien, portugais) <ul style="list-style-type: none">• Vue d'ensemble des réalisations et évaluations• Démonstrations• Conclusions et Perspectives• Réflexions sur la place de l'humain dans la boucle
15:45	Tous	<ul style="list-style-type: none">• Discussion avec les conseillers et observateurs
16:00	Pause Café & Photo de groupe	
16:15	LUTIN	Résultats-clés issus des évaluations concernant la génération automatique des sous-titres français et multilingues <ul style="list-style-type: none">• Présentation des échantillons• Progression des retours d'utilisateurs au cours du projet• Atouts et limites des protocoles utilisés
16:45	Tous	Perspectives/ouverture pour de futurs projets <ul style="list-style-type: none">• Comment comparer de façon juste différentes méthodes de production des sous-titres (qualité, quantité, productivité, coût, effort cognitif, créativité...) ?• Quels processus de concertation dans l'évolution des métiers et des outils du sous-titrage ?• Conclusion de l'après-midi
17:15	Systran, HC, FTV	Conclusions de la journée et prochaines étapes
17:30	Fin de la réunion	

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

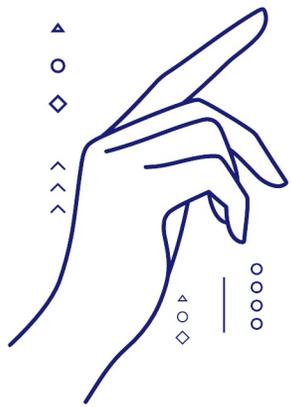
→ Un pas vers une automatisation de la traduction de contenus journalistiques du français vers la langue des signes française (LSF) via l'animation d'un avatar 3D

O3

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



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



ROSETTA
accès multilingue

François YVON
François BUET
Eric FLORENCE

• 14h45 - 15h15

- Enjeux
- Progrès et limites au cours du projet
- Démonstration des systèmes en action
- Analyse des résultats de l'évaluation métier en regard de ceux des métriques automatiques
- Perspectives de recherche

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



↪ *Contraintes spatiales et temporelles*

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



Direct



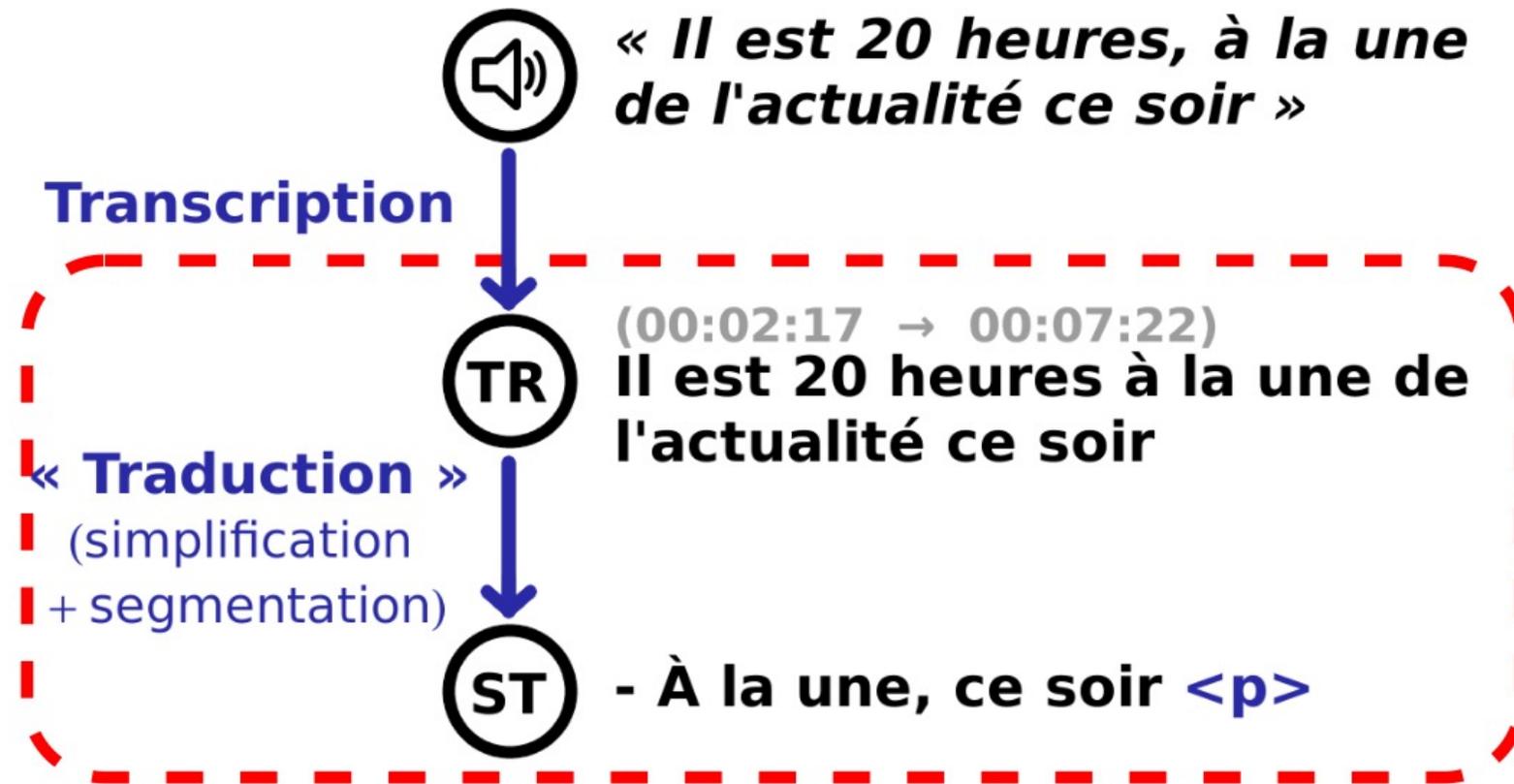
Stock



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

↳ *Adaptation multi-genre*

Le sous-titrage comme une traduction monolingue



Récolte et traitement des données france.tv access



Transcription
automatique

TR

Alignement
automatique

ST

```
<span tts:textAlign="center" tts:color="white" style="background-color: yellow;">es collègues ont tout repeint.</span>
</p>
<p xml:id="caption83" ttm:role="caption" begin="00:04:27:15" end="00:04:32:13">
<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>
```

```
me="265.51" dur="0.24" conf="0.982"> plafonds </Word>
me="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>



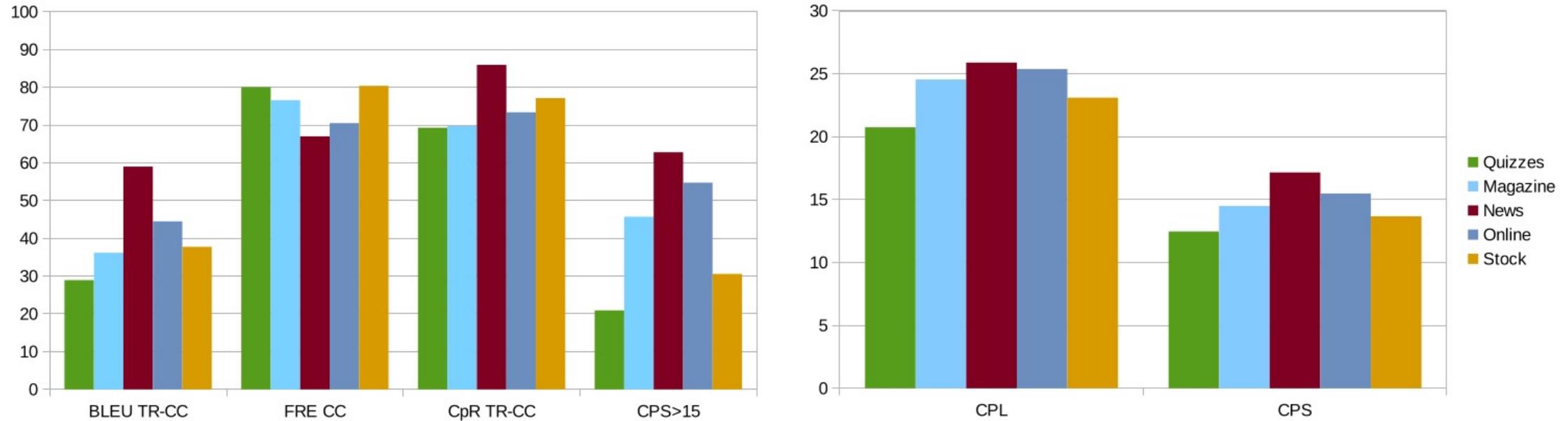
Volume de données alignées (novembre 2021) :

- 2 900 heures de vidéo,
- 780 000 paires de phrases.

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

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

Variabilité des sous-titres selon le genre dans le corpus



- les sous-titres en direct sont plus proches de la transcription (ex. journal),
- les sous-titres en stock sont moins denses (ex. jeu),
- des écarts pour la lisibilité et la compression également.

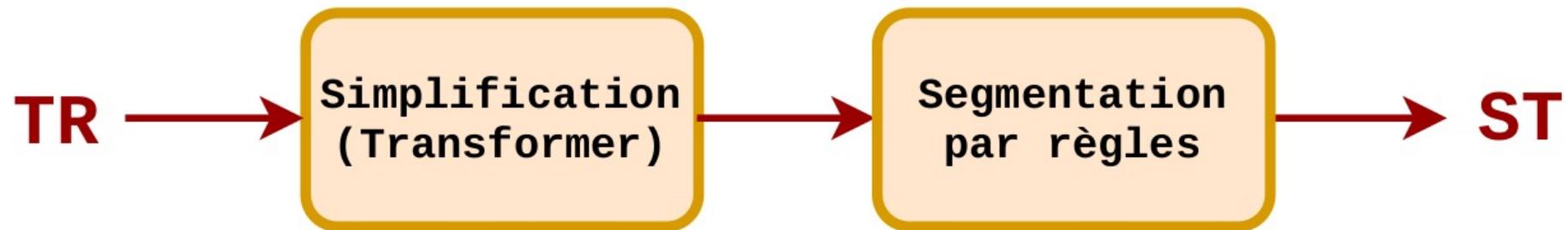
Traduction vers ST : systèmes de départ avec module de segmentation



TR + règles (module de segmentation séparé) :



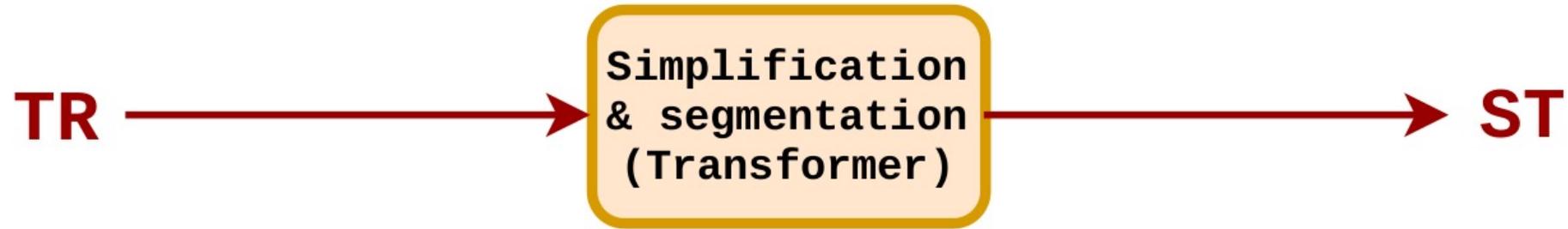
Transformer (IA de traduction) + règles :



Traduction vers ST : systèmes bout-en-bout intégrant les frontières



Transformer + frontières (intégrées parmi les mots) :



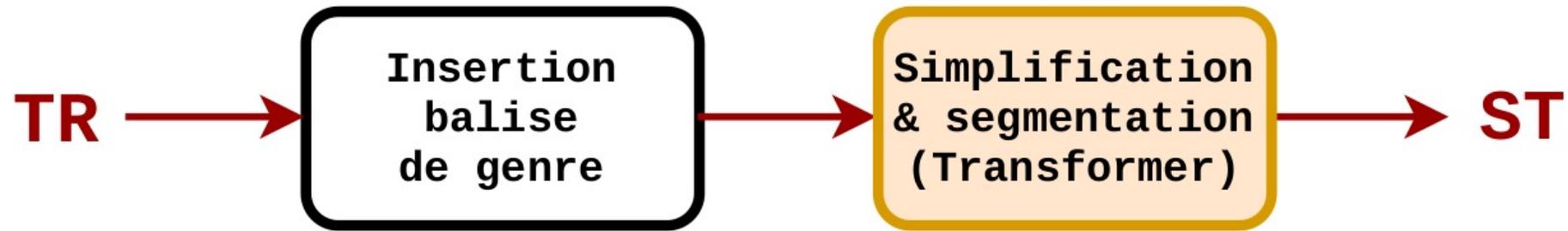
Transformer + frontières + contrôle de longueur :



Traduction vers ST : systèmes adaptés au genre



Transformer + frontières + balises de genre (BG) :



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
CPS = 17,2

CpR = 0,70
CPS = 14,5

CpR = 0,69
CPS = 12,5

CpR = 0,86
CPS = 12,2

Exemples d'apprentissage selon les modèles



Systèmes avec module de segmentation

TR Donc nous on monte les mâts oui mais bon, on va laisser finir.
CC Nous, on monte les mâts. Ah oui. Stop. On va les laisser finir.

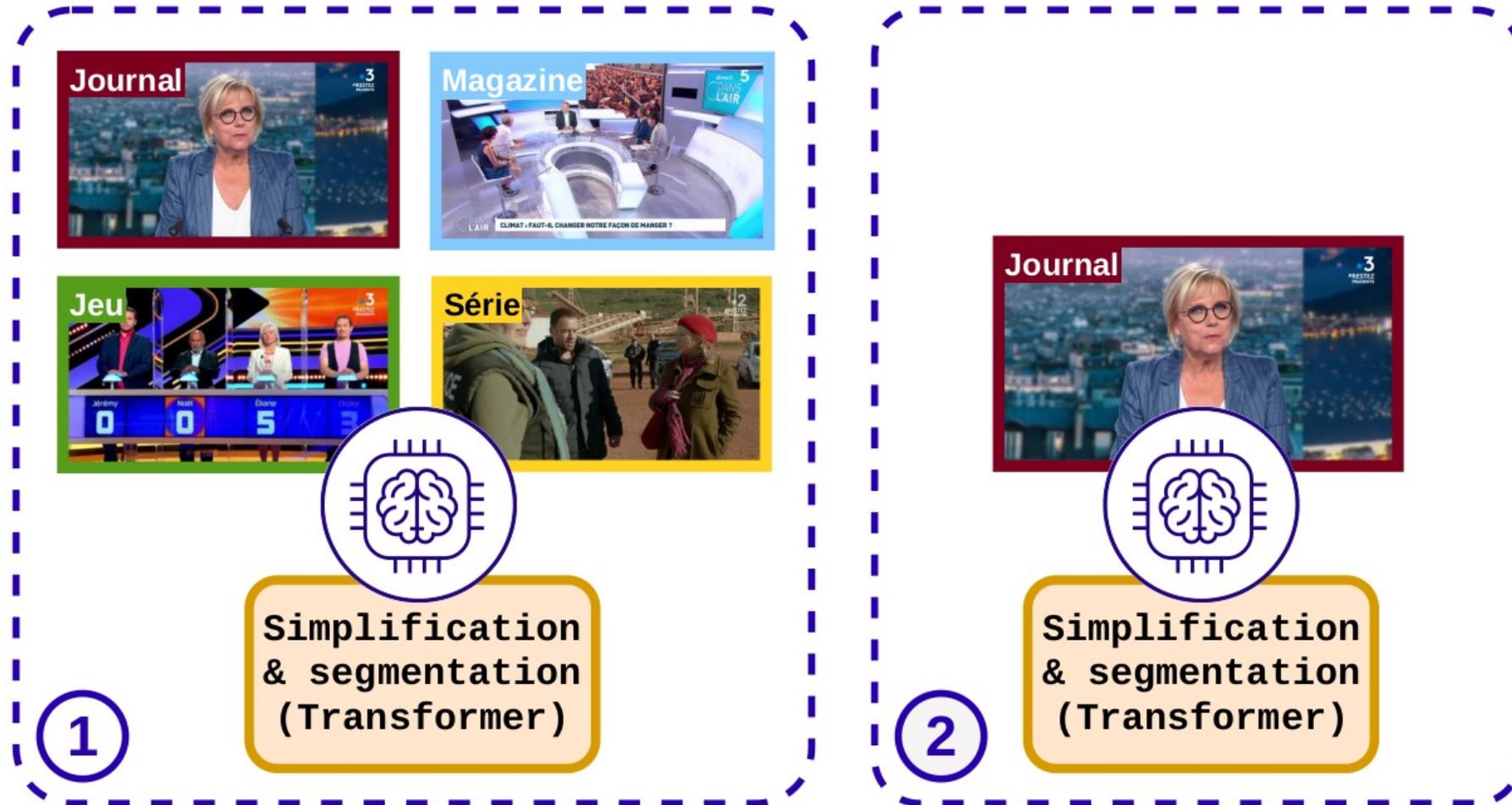
Systèmes bout-en-bout

TR Donc nous on monte les mâts oui mais bon, on va laisser finir.
CC Nous, on monte les mâts. **<p>** Ah oui. Stop. **
** On va les laisser finir. **<p>**

Systèmes bout-en-bout avec balises de genre

TR **<stock><mag>** Donc nous on monte les mâts oui mais bon, on va laisser finir.
CC Nous, on monte les mâts. **<p>** Ah oui. Stop. **
** On va les laisser finir. **<p>**

Traduction vers ST : systèmes affinés par domaine



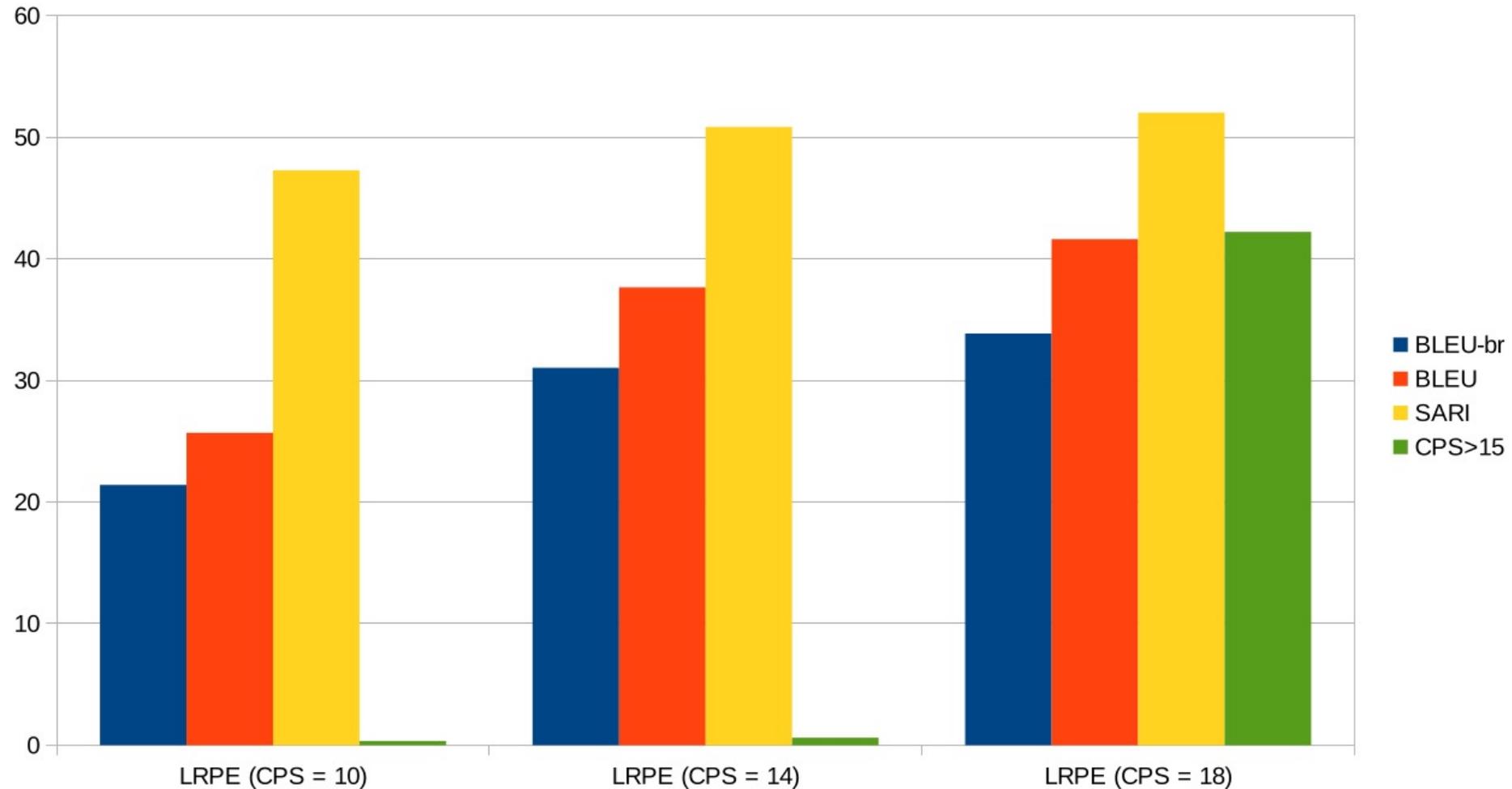
Performances : progression pour les systèmes adaptés au genre



Systemes	↑BLEU-br	↑BLEU	↑SARI	↓TER-br	↓CPL>36	↓CPS>15
<i>Systemes indifferents au genre (février 2021, V2)</i>						
+ front.	33,3	43,8	52,9	0,38	6,1	63,8
+ front. + RT	33,5	44,0	53,0	0,38	2,9	64,1
+ front. + LRPE (CPS = 14,5)	30,3	39,5	51,2	0,31	5,3	1,2
<i>Systemes adaptés au genre (février 2021, V2)</i>						
+ front. + BG	34,1	44,5	53,2	0,34	7,0	58,0
+ front. + RT + BG	34,3	44,6	53,4	0,37	3,3	62,1
+ front. + LRPE (CPS adapté)	31,4	40,9	51,5	0,31	5,3	10,8
+ front. + LRPE (CPS) + BG	32,5	42,2	52,2	0,31	5,6	13,9

- Amélioration globale avec balise de genre.
- Le contrôle de longueur adapté au genre reste moins bon en qualité.
- Résultats intermédiaires en combinant balise et contrôle de longueur adapté.

Contrôle de longueur et relâchement des contraintes sur les sous-titres



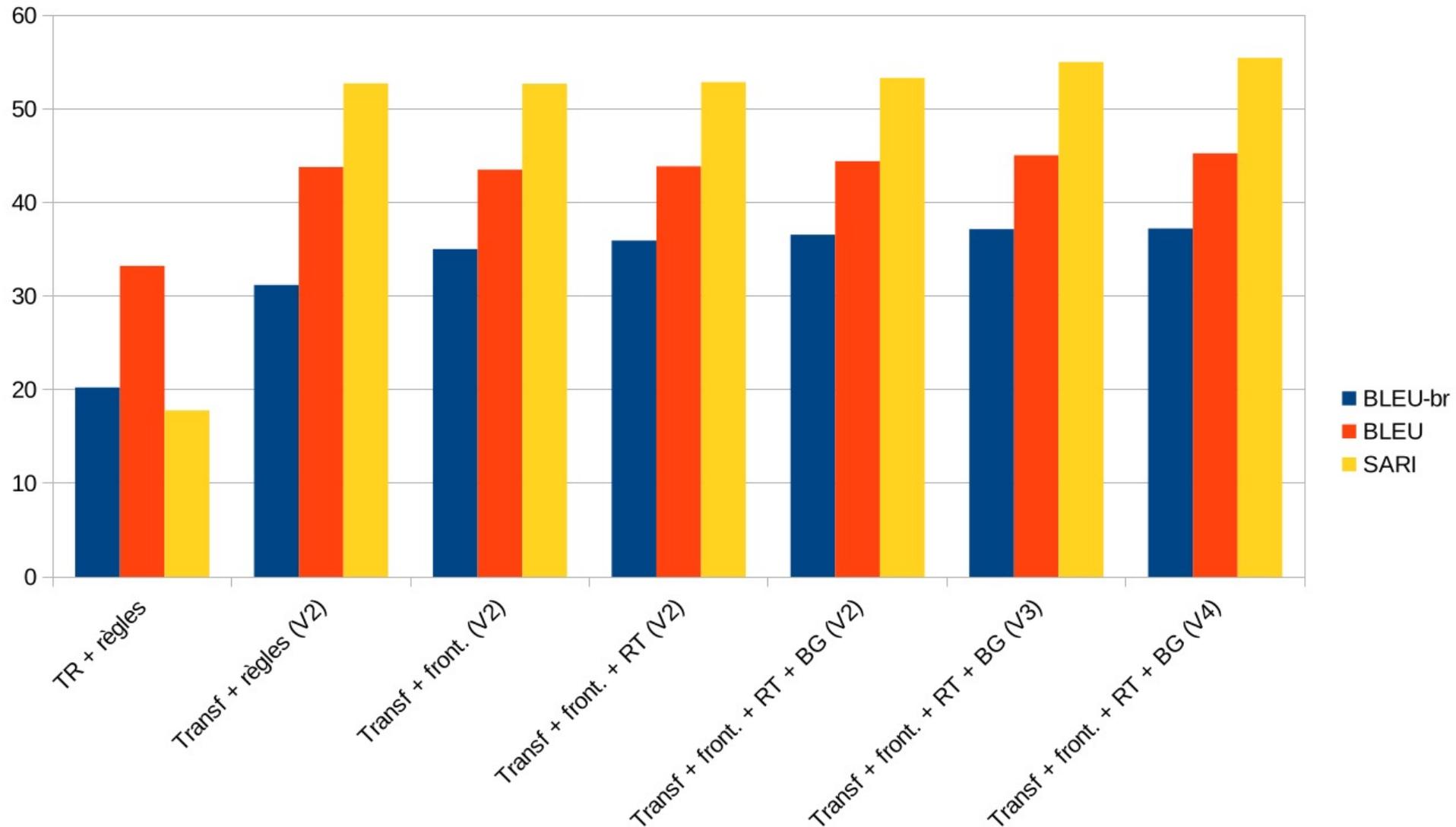
Performances : progression entre les versions du corpus



Systemes	↑BLEU-br	↑BLEU	↑SARI	↓TER-br	↓CPL>36	↓CPS>15
<i>février 2021 (V2)</i>						
+ front. + RT + BG	36,6	44,4	53,3	0,31	2,7	53,3
<i>juillet 2021 (V3)</i>						
+ front. + RT + BG	37,2	45,0	55,0	0,31	2,4	53,8
<i>novembre 2021 (V4)</i>						
+ front. + RT + BG	37,2	45,3	55,5	0,31	1,7	54,3

- +200 000 paires de phrases dans le corpus entre février et juillet 2021.
- +150 000 paires de phrases dans le corpus entre juillet et novembre 2021.

Performances : progression générale



Performances : variabilité selon les genres et adaptation



Systemes	↑BLEU-br	↑BLEU	↑SARI	↓TER-br	↓CPL>36	↓CPS>15
<i>Des chiffres et des lettres (jeu)</i>						
+ front. + RT	19,7	22,1	45,3	0,35	1,0	24,6
+ front. + RT + BG	20,8	22,2	46,7	0,36	1,4	28,5
+ front. + affin.	20,6	22,2	46,1	0,33	0,7	28,7
<i>Le magazine de la santé (magazine)</i>						
+ front. + RT	36,6	43,0	54,5	0,45	2,1	71,6
+ front. + RT + BG	40,4	46,6	57,5	0,37	2,2	58,8
+ front. + affin.	39,3	45,7	57,1	0,39	1,7	63,4
<i>Journal 13h00 (journal)</i>						
+ front. + RT	55,3	69,5	66,2	0,19	2,1	71,7
+ front. + RT + BG	56,8	69,1	66,6	0,19	2,6	73,1
+ front. + affin.	57,3	69,2	66,5	0,18	2,2	72,6

Performances : impact de la qualité de la reconnaissance de parole



Systèmes	BLEU-br	BLEU	SARI	TER-br	CPL>36	CPS>15
<i>Systèmes partant de la transcription automatique</i>						
+ règles	30,0	43,3	52,2	0,39	0,0	63,5
+ front.	33,3	43,8	52,9	0,38	6,1	63,8
+ front. + RT + BG	34,3	44,6	53,4	0,37	3,3	62,1
+ front. + LRPE (CPS = 14,5)	30,3	39,5	51,2	0,31	5,3	1,2
<i>Systèmes partant de la transcription manuelle</i>						
+ règles	32,3	45,8	50,2	0,37	0,0	48,3
+ front.	35,1	46,0	50,5	0,36	6,1	53,7
+ front. + RT + BG	35,3	45,6	49,8	0,37	3,7	50,4
+ front. + LRPE (CPS = 14,5)	31,6	41,1	49,3	0,32	5,3	1,3

Performances : point de vue de l'évaluation métier



Systemes	↑orthographe	↑contresens	↑moyenne	↓erreurs/ST
<i>février 2021 (V2)</i>				
+ front.	2,18	2,63	2,84	0,500
+ front. + RT + BG	2,44	2,70	2,89	0,475
<i>juillet 2021 (V3)</i>				
+ front.	2,56	2,64	2,84	0,460
+ front. + BG	2,46	2,90	2,92	0,461
+ front. + RT + BG	2,44	2,65	2,80	0,486

- 6 notes sur 5 : orthographe, contresens, lisibilité, mots manquants, ponctuation, césure.
- Compte d'erreurs par catégorie.
- 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>

Exemples de sous-titres produits



LRPE (CPS = 10)

Vous en tant que metteur en scène,
 je ne sais pas.

LRPE (CPS = 14)

Vous en tant que metteur en scène,
 je ne sais pas. <p> Je me méfie.

LRPE (CPS = 18)

Vous en parlez tant que metteur
 en scène, je ne sais pas. <p> Je me méfie beaucoup d'élus.
<p>

BG = jeu, stock

Présentez avec les mains. <p> En tant que metteur en scène,
 on n'est pas... <p> Je me méfie beaucoup d'élus. <p>

BG = magazine, direct

Présenter avec les mains, le truc... <p> En tant que metteur en scène,
 on n'est pas... <p>
Je me méfie beaucoup d'élus. <p>

BG = journal, stock

Présenter avec les mains, le truc,
 en tant que metteur en scène, <p> je ne sais pas. <p> Je me méfie beaucoup d'élus. <p>

Conclusion et perspectives



Segmentation des sous-titres :

amélioration par rapport aux systèmes de départ en intégrant les frontières dans les données.

Adaptation au genre :

gain pour la simplification et la segmentation en utilisant des balises en entrée.

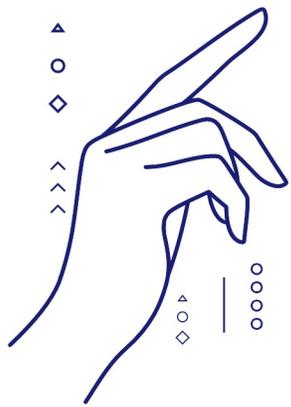
Contrôle de longueur :

compromis entre le respect des normes et la qualité des phrases.

En cours :

- Comparaison des stratégies selon une évaluation métier réalisée par france.tv access.
- Réflexion sur les métriques d'évaluation automatique (avec FBK).
- Réflexion sur les architectures pour l'adaptation au genre, ou le passage en langue étrangère (avec SYSTRAN).

4. Génération automatique de sous-titres multilingues : anglais, espagnol, italien, portugais



ROSETTA
accès multilingue

Elise BERTIN-LEMÉE
Eric FLORENCE

- Vue d'ensemble des réalisations et évaluations
- Démonstrations
- Conclusions et perspectives
- Réflexions sur la place de l'humain dans la boucle

• 15h15 – 15h45



SYSTRAN
beyond language

france•tvaccess

Réalisations

V0 – 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**

V1 - 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**

V2 – 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 CONTINUE DES MODELES



Evaluations

V0 – Preuve de concept interne
Modèles de traduction de ST
Août 2020

3 DOMAINES / USAGES

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

V1 - Modèles de génération de ST
à partir de la reconnaissance vocale
Mai 2021

3 TYPES D'EVALUATION

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

V2 – Modèles de traduction de ST
basée sur le contexte
Juillet 2021



INTEGRATION DANS LE DEMONSTRATEUR

Evaluation interne

1. MOOCs / tedtalks – apprendre

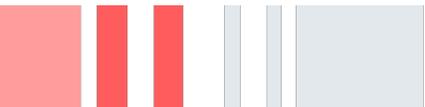


Language	Source	Translation		Segmentation TER-break	Reading ease			
		BLEU	BLEU-nobreak		Char Per Line	Char Per Second	Char Per Sec - CSA	
fr		source			100%	89%	59%	
en		ref			100%	98%	79%	
	sous-titres FR	v0.mono	19.37	17.7	25	95%	97%	68%
		v0.multi	18.25	16.29	27	96%	97%	65%
		v0.nosource	14.32	12.86	31	96%	95%	53%
	transcript	v1	13.32	12.5	31	97%	96%	57%
		v1.noftv	14.13	13.4	31	96%	93%	57%
		v2.no_context.transcript	15.55	13.96	32	96%	96%	64%
		v2.context.transcript	15.49	13.86	34	95%	97%	58%
	sous-titres FR	v2.no_context.fr	19.09	17.13	26	97%	96%	58%
		v2.context.fr	18.39	16.76	27	96%	96%	51%

2. Séries – se divertir



Language	Source	Translation		Segmentation TER-break	Reading ease			
		BLEU	BLEU-nobreak		Char Per Line	Char Per Second	Char Per Sec - CSA	
fr		source			100%	97%	48%	
		Opensubtitles			100%	78%	51%	
en		ref			100%	95%	50%	
	sous-titres FR	v0.mono	23.75	22.93	27	97%	94%	54%
		v0.multi	22.09	21.4	29	98%	95%	54%
		v0.nosource	16.08	15.63	36	97%	71%	52%
	transcript	v1.noftv	14.07	13.63	41	98%	66%	50%
		v2.no_context.transcript	16.2	15.28	36	97%	94%	49%
		v2.context.transcript	14.84	14.2	38	93%	58%	40%
	sous-titres FR	v2.no_context.fr	20.6	19.65	28	98%	94%	49%
		v2.context.fr	20.84	19.78	28	95%	61%	40%



Démonstrations

Sous-titre original
(france.tv access)

V1 - Modèles de génération de ST
à partir de la reconnaissance vocale
Mai 2021

V0 – Preuve de concept interne
Modèles de traduction de ST
Août 2020

V2 – Modèles de traduction de ST
basée sur le contexte
Juillet 2021



Améliorations

Référence à un sous-titre précédent
"ils y arrivent"



Évidemment, c'est des bébés
dont je vous parle.



Of course, these are babies
I'm talking about.



Obviously, it was babies
I'm talking to you about.



Obviously, these are babies
I'm talking about.

Améliorations

Expressions
"relève de l'exploit"



Mais par quel bout commencer pour
apprendre ma langue et savoir ce



But what part of the beginning
to learn my language and know what it tells me?



But how much does it begin to learn
my language and know what it tells me?



But how much do I start
to learn my language

Améliorations

Expressions
"dégage le ciel"



•2

de protection
quisitionne pour les soignants
dans des pharmacies aujourd'hui.



•2

de protection
quisitionne for treaters in pharmaces today.



•2

de protection
quisitionne to the drug care workers today.



•2

de protection
quisitionne The coronavirus clearing
the skies of China.

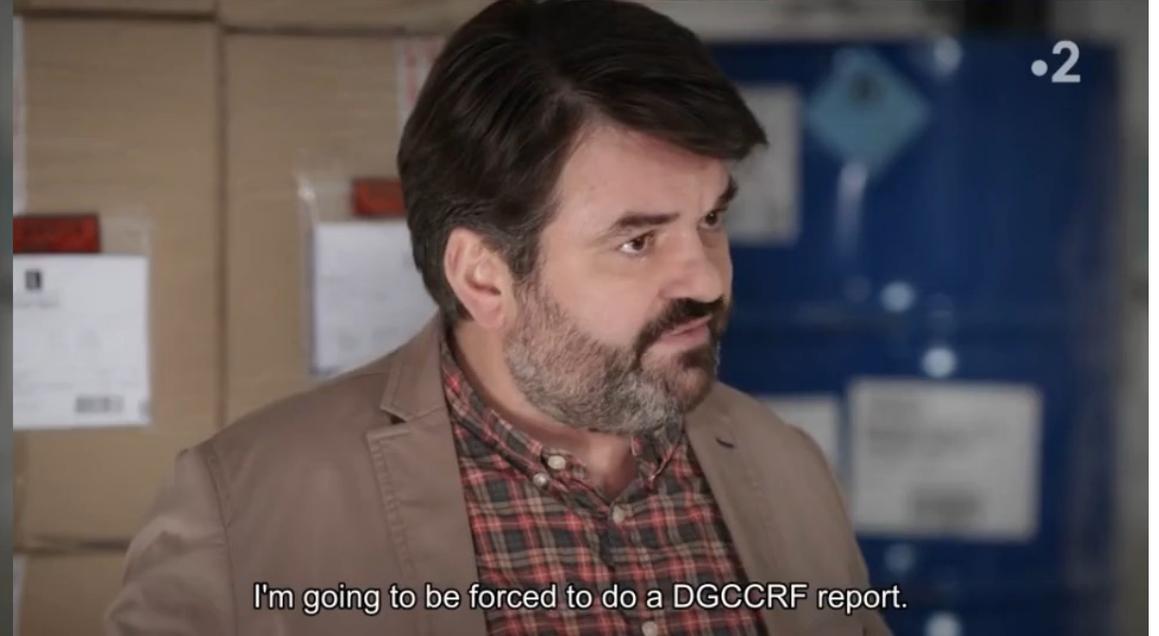
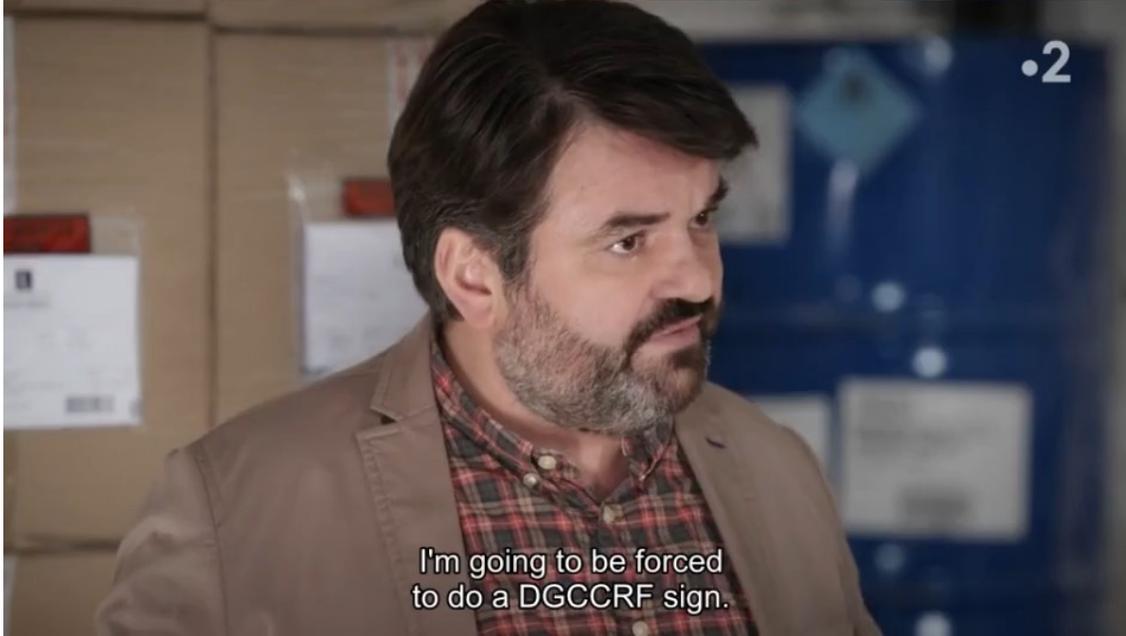


Améliorations

Cohérence du dialogue
"l'aider, la laisser comme ça"









Parce qu'autrement,
ils étaient beaucoup trop éloignés



Because otherwise, they were too far away



Because otherwise, well,
they were too far away



Because otherwise,
they're way too far away

Défis

Références culturelles
"Dinah"



Défis

VO > VF > VO
"gros niais"

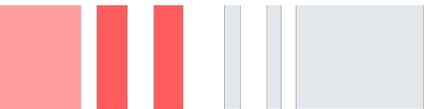


Source	Reference	V1	V2
l'âge de Pierre	(EN) the Stone Age (ES) La Edad de Piedra	(EN) Pierre's age (ES) la edad de Pierre	(EN) the Stone Age
Ca me démonte.	(EN) I hate that. (ES) Lo detesto.	(EN) It's disassembling me. (ES) Me desmonta.	(EN) I'm just walking apart.
mais 40 000 dollars, c'est rondelet	(EN) but \$40,000 is a lot for a wedding (ES) ...pero 40.000 dólares es mucho dinero	(EN) But 40,000 dollars is round (ES) Pero 40.000 dólares son rígidos	(EN) but 40,000 dollars is roller coaster



Evaluation métier

- Evaluation humaine de france.tv access par échantillonnage statistique : picking de 5 minutes aléatoires sur une sélection d'émissions variées
- 1691 extraits en français
85 en anglais avec le système V1
116 en anglais avec le système V2
60 en espagnol
- Critères :
 - Orthographe
 - Contresens
 - Éléments manquants (mots, phrases)
 - Ponctuation (qui change le sens) Césure
 - Lisibilité



Evaluation métier

EN v1 –mai 2021

Usage	Domaine	#	moyenne	total_Orthographe	total_Contresens	total_mots_manquants	total_Phrases_Manquantes	total_Ponctuation	cesure	lisibilite
Apprendre	Documentaire	1	1,98	8,00	50,00	3,00	0,00	5,00	1,50	2,50
	Magazine	26	2,35	1,54	76,92	3,69	1,58	4,85	1,81	2,00
S'informer	Informations	19	2,68	3,16	72,05	4,37	0,68	2,95	2,75	2,38
Se divertir	Fiction	20	2,25	0,85	78,65	16,65	11,25	7,25	1,44	1,33
	Jeu	19	2,22	1,11	97,74	16,16	7,21	4,68	1,34	1,46
TOTAL		85	2,37	1,72	80,58	9,67	4,89	4,95	1,82	1,81

EN v2 –juil 2021

Usage	Domaine	#	moyenne	total_Orthographe	total_Contresens	total_mots_manquants	total_Phrases_Manquantes	total_Ponctuation	cesure	lisibilite
Apprendre	Documentaire	10	3,57	3,30	18,60	2,60	2,20	0,50	3,56	3,81
	Magazine	76	3,14	4,45	32,51	5,39	2,64	0,83	3,33	3,54
S'informer	Informations	14	2,76	4,14	37,43	3,21	3,79	1,36	3,09	2,24
Se divertir	Fiction	14	3,20	1,93	33,64	3,86	3,79	0,93	3,12	3,08
	Jeu	2	2,96	7,00	21,00	1,00	4,50	2,50	3,50	4,25
TOTAL		116	3,13	4,05	31,84	4,63	2,91	0,91	3,30	3,36

ES v1 –mai 2021

Usage	Domaine	#	moyenne	total_Orthographe	total_Contresens	total_mots_manquants	total_Phrases_Manquantes	total_Ponctuation	cesure	lisibilite
Apprendre	Magazine	25	2,25	2,88	66,80	11,92	2,60	10,24	2,06	1,45
S'informer	Informations	15	2,19	4,13	64,07	9,93	1,27	7,13	1,71	1,37
Se divertir	Fiction	11	2,26	3,82	62,27	27,27	14,82	2,91	2,37	1,84
	Jeu	9	2,34	2,22	72,44	11,67	8,89	4,11	1,88	1,86
TOTAL		60	2,25	3,27	66,13	14,20	5,45	7,20	2,01	1,56

Conclusions du module



- Résultats automatiques ne permettant pas une diffusion directe à l'antenne
 - moyenne ayant évolué de 2,25/5 pour v1 à 3/5 pour v2, alors qu'un minimum de 4/5 est attendu dans par l'équipe de france.tv access pour être estimé utilisable
 - l'utilité de ces sorties automatiques reste à évaluer dans d'autres contextes et chaînes de production
- Erreurs variables en fonction du type d'émission
 - reconnaissance vocale trop bruitée dans le cas de séries, jeux, dessins animés qui sont les principaux domaines ciblés pour des outils automatiques
 - lorsque la reconnaissance vocale est de bonne qualité, les erreurs les plus fréquentes sont les contresens : des erreurs qu'on remarque... donc facilement détectables en post-édition ?
- Progrès, non seulement sur la qualité de la traduction, mais aussi l'intégration
 - amélioration de la synchronisation, des couleurs, locuteurs
 - grand impact sur l'expérience utilisateur

Pour la suite des recherches



- Penser la traduction au niveau de l'émission (voire de la saison, la série)
- Séparer le son (musiques, jingles, bruits de fond) de la parole, si cela reste impossible d'accéder à la version avant mixage
- Détecter des informations dans l'image (positionnement, couleurs, timecodes, plans)

Pour de futurs projets



- Trouver des façons d'avoir accès à plus de données multilingues de qualité professionnelle
- Accompagner les résultats automatiques d'une interface pour une interaction avec l'humain facilitée
- Etablir une juste comparaison avec d'autres outils automatiques du marché et avec les méthodes existantes

Perspectives d'exploitation à SYSTRAN



- poursuivre la recherche pour atteindre la qualité attendue dans le cas d'un sous-titrage destiné à la télévision
 - considérer pour les types de modèles développés dans le cadre de ROSETTA d'autres cibles, pour lesquelles :
 - l'automatique est nécessaire du fait des volumes
 - la qualité peut être suffisante pour le cas d'usage
 - et/ou la post-édition peut être intégrée
- (ex. campagnes de marketing, vidéos de communication interne, modules de formation, tutoriels, webinars, etc.)

Un exemple de passage progressif vers la Traduction Automatique



- Qui : entreprise sportive avec forte image de marque
- Quoi : contenus très créatifs de marketing
- Avant (2019) : agence de traduction / localisation peu emballée par la MT
- Pendant (2020) : gestion du changement
 - 3 demi-journées de formation à la post-édition, accompagnement par des experts du métier externes à SYSTRAN
 - premiers modèles développés par SYSTRAN avec les données existantes
 - réunions régulières de suivi
- Maintenant (2021) :
 - changement de méthodes pour les traducteurs, méthodes qui au départ prennent plus de temps, mais font gagner en productivité
 - ré-entraînement de modèles automatiques en tenant compte des retours

SYSTRAN Market place

END-USERS



DATA PROVIDERS



TRAINERS





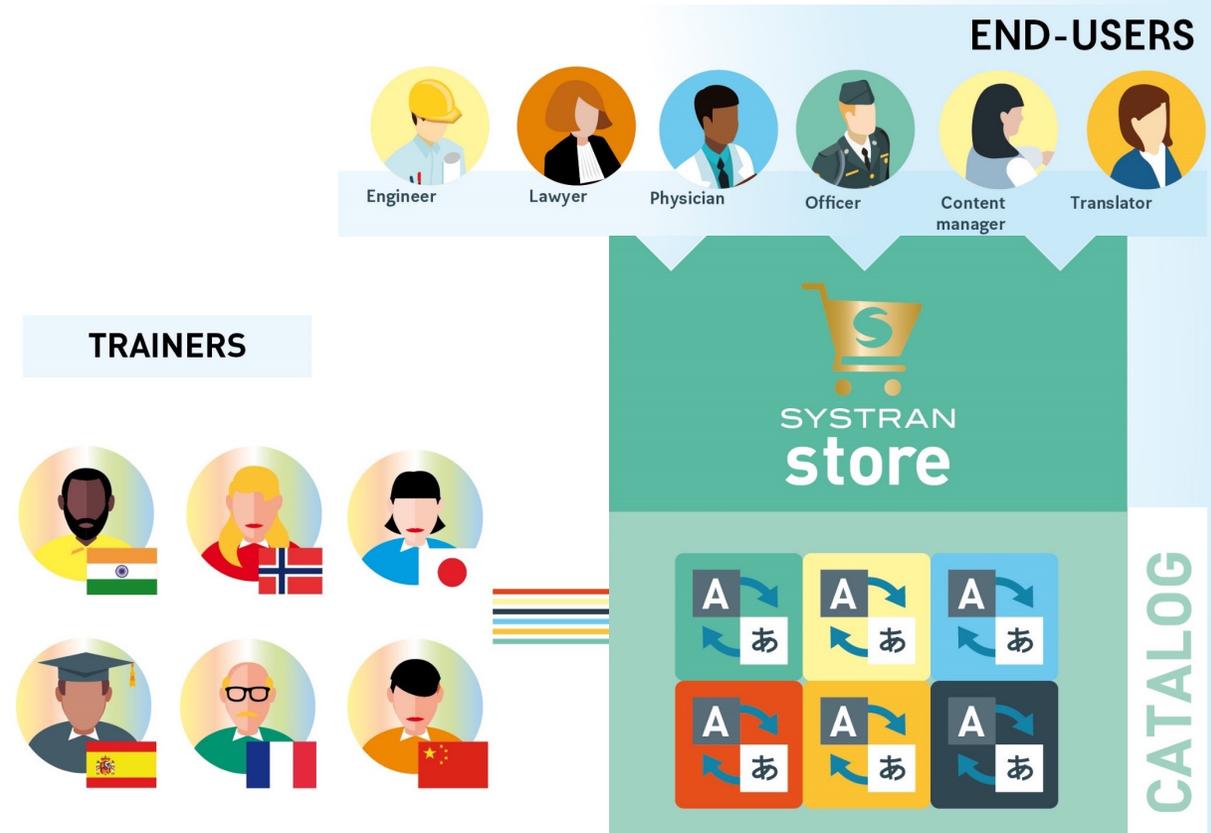
The SYSTRAN Marketplace

The SYSTRAN Marketplace is an online platform where:

- Language experts (Trainers) can build, share, give world wide visibility and sell their models to MT Industry Users.



- MT industry users can freely test and evaluate available models and buy access to these models





Responsible Platform

- Training data is not exposed to end-users or third-party trainers
- Translated data belongs to user and is not used
- Certified datacenter infrastructure through OVHCloud





Responsible Platform

intellectual
property protection



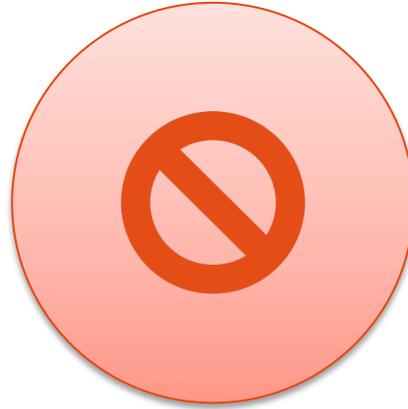
- Trainer keeps full IP on the generated model
- SYSTRAN modelStudio checks that no restricted-license data is used for the training of a model
- No dilution of IP when a model is used for specialization



Trainer Obligation



**TRAINER COMMITS NOT TO USE
DATA WITH RESTRICTED
CONTENT WHEN TRAINING
MODEL**



**TRAINER COMMITS NOT TO
INTENTIONALLY BIAS A MODEL
TRANSLATION OR TO TEACH A
MODEL TO PRODUCE
DEFAMATORY TRANSLATION**



**TRAINER COMMITS FOR
REASONABLE EFFORTS TO
MAINTAIN TRANSLATION
MODELS BASED ON USER
FEEDBACK**

Community-based

→ Responsible Portal

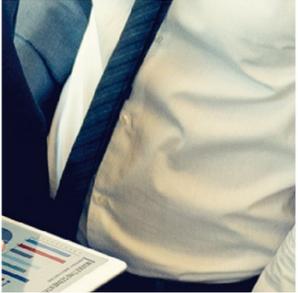


institutions

corporates



data providers



LSPs

freelance translators

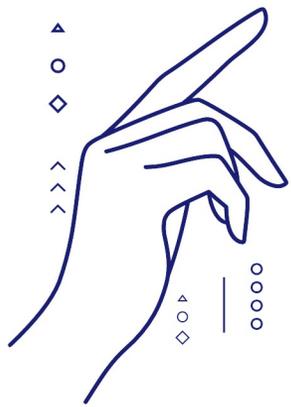
academics



Open Source contributors



SYSTRAN
Market place
The professional Choice



ROSETTA
accès multilingue

Tous

• 15h45 - 16h



Discussion avec les conseillers
et observateurs



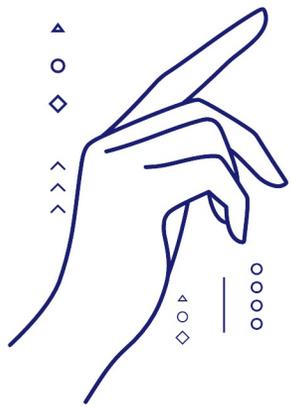


16h - 16h15

Pause café & Photo de groupe



5. Résultats-clés issus des évaluations concernant la génération automatique des sous-titres français et multilingues



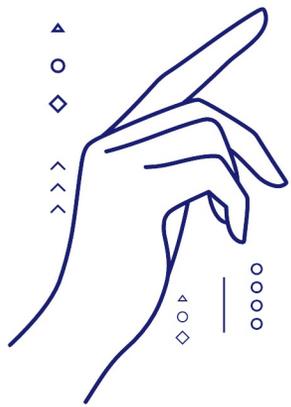
ROSETTA
accès multilingue

Charles Tijus
Taïna Victor
Léa Lachaud

- 16h15 – 16h45

- Présentation des échantillons
- Progression des retours d'utilisateurs au cours du projet
- Atouts et limites des protocoles utilisés

Discussion : Perspectives/ouverture pour de futurs projets



ROSETTA
accès multilingue

Tous

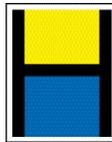
• 16h45 - 17h15

- Comment comparer de façon juste différentes méthodes de production des sous-titres (qualité, quantité, productivité, coût, effort cognitif, créativité...) ?
- Quels processus de concertation dans l'évolution des métiers et des outils du sous-titrage ?
- Conclusion de l'après-midi



Conclusion et prochaines étapes

17h15 – 17h30



Holken Consultants
& Partners

france•tvaccess

Merci !



Elise BERTIN-LEMÉE, coordination ROSETTA elise.bertinlemee@systrangroup.com
Eric FLORENCE, l'hôte de la reunion Advisory Board #3 florenceeric@francetvaccess.fr
Hadmut HOLKEN, suivi du comité consultatif holken@holkenconsultants.com