

Sans les mains !

État des lieux des outils de retranscription automatique

1 – Retour d'expérience et points de vigilances



© Mindhunter

Sandrine Ollinger – 11 avril 2024

Plan

- Une aventure MATE-SHS
- La transcription
- Méthodologie proposée

Une aventure MATE-SHS

Les origines

Avril 2019

- Un message anodin

Bonjour,
Je suis à la recherche d'informations sur **Vocalmatic** Est-ce que vous connaissez, utilisez ou avez des infos sur cet outil SVP?
Merci par avance pour vos réponses
Bonne journée

- Une meneuse de projet

Bonjour,

je ne peux malheureusement pas vous répondre sur ce logiciel-ci mais j'avais fait une revue rapide des solutions de transcriptions il y a un an, et ce type de service commençait tout juste à émerger. Pour ma part j'utilisais jusqu'à présent Youtube, mais c'est très mauvais pour le français.

Je me demandais donc si cela intéresserait quelques personnes de faire un benchmark de ces nouveaux outils : on pourrait choisir un entretien commun (libre d'utilisation et dans l'idéal déjà retranscrit), se répartir les plate-formes, et comparer ensuite les résultats obtenus.

Qu'en pensez-vous ?

Cordialement

Elise Tancoigne

Le groupe de travail



The image shows a screenshot of a Jitsi video conference. The interface is dark grey with the Jitsi logo and 'jitsi.org' in the top left corner. Six participants are visible in a 2x3 grid. Each video tile has a name label below it in yellow text. The participants are: Jean-Philippe Corbellini (top left), Daniel Valero (top middle), Gaëlle Deletraz (top right), Sandrine Ollinger (bottom left), Laure Gayraud (bottom middle), and Elise Tancoigne (bottom right). Gaëlle Deletraz's video tile is highlighted with a blue border. Small icons in the top left of each tile indicate video status: a yellow 'x' for muted video, a red 'x' for disabled video, and a blue 'v' for active video.

jitsi.org

Jean-Philippe Corbellini

Daniel Valero

Gaëlle Deletraz

Sandrine Ollinger

Laure Gayraud

Elise Tancoigne

Les résultats

Une méthodologie et sa diffusion

MATE-SHS Tuto MATE-SHS, 4 juin 2020

La transcription automatique : un rêve enfin accessible ?



Elise TANCOIGNE, Jean Philippe CORBELLINI, Gaëlle DELETRAZ,
Laure GAYRAUD, Sandrine OLLINGER, Daniel VALERO
Avec les contributions de Laurie Boyer, Justine Lascar, Giancarlo Luxardo

La transcription automatique : un rêve enfin accessible ?

Analyse et comparaison d'outils pour les SHS.
Nouvelle méthodologie et résultats

Élise TANCOIGNE¹, Jean-Philippe CORBELLINI², Gaëlle DELETRAZ³,
Laure GAYRAUD⁴, Sandrine OLLINGER⁵, Daniel VALERO⁶
Avec les contributions de : Laurie Boyer⁴, Justine Lascar⁴, Marick Fèvre⁷, Giancarlo Luxardo⁸



¹ Université de Genève, elise.tancoigne@unige.ch
² Maison des Sciences de l'Homme Val-de-Loire USR 3501, jean-philippe.corbellini@univ-tours.fr
³ UMR 5319 Passages - CNRS/Université de Pau et des Pays de l'Adour, gaëlle.deletraz@univ-pau.fr
⁴ Centre régional associant le Cereq au Centre Émile Durkheim - IEP de Bordeaux, lgayraud@sciencespobordeaux.fr
⁵ ATILF - UMR 7118 CNRS/Université de Lorraine, Sandrine.Ollinger@atilf.fr
⁶ I.C.A.R. - UMR 5191 CNRS/ENS de Lyon/Université Lyon 2, Daniel.Valero@ens-lyon.fr
⁷ CENS - UMR 6025 CNRS/Université de Nantes - Réseau doctoral EHESP
⁸ UMR 5267 Praxiling - CNRS/Université Paul Valéry Montpellier 3

journals.sagepub.com/home/bms ISSN 0759-1063

BMS

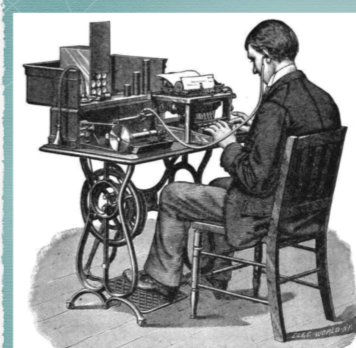
BULLETIN OF SOCIOLOGICAL METHODOLOGY / BULLETIN DE MÉTHODOLOGIE SOCIOLOGIQUE

MICRO-MACRO
 Camille Hamidi and Claire Dupuy Micro-macro: Why focus on levels of data collection and analysis? Introduction to the *BMS* Micro-macro section
 Pierre Wokuri Que peut la comparaison multiniveau dans l'identification de mécanismes causaux? Réflexions à partir d'une enquête sur les projets coopératifs d'énergie renouvelable en Europe

OUTILS ET INSTRUMENTS/TOOLS AND INSTRUMENTS
 Elise Tancoigne, Jean Philippe Corbellini, Gaëlle Deletraz, Laure Gayraud, Sandrine Ollinger et Daniel Valero Un mot pour un autre? Analyse et comparaison de huit plateformes de transcription automatique

IMPLEMENTATION/MISE EN OEUVRE
 Carlos Solar Thinking beyond averages: Quantile regression modelling and military expenditure
 Iris Condamine-Duceux and Marie Jauffret-Roustide Comment se questionnent les pratiques d'injection? Retours méthodologiques de terrain
 Jøhs Hjelbrekke and Vegard Jarness Cultural divisions and time: Mapping diachronic homologies using class-specific MCA (CSA)

ISSUE 155 • JULY 2022



Transcrire automatiquement ses entretiens : rêve ou réalité ?

Sandrine Ollinger (CNRS)
Elise Tancoigne (UNIL)

UNIL, 16 sept. 2021
École doctorale PDEN

Le réseau RUSHS

22 SAISON 3

RUSHS HOUR

Le réseau Mate-SHS

La transcription

Qu'est-ce que c'est ?

- On passe d'un flux sonore à un texte



Mais

Et d'ailleurs ici à l'usine on me prenait pour un Arabe hein

Ben oui tellement tu étais

Et si je me chamaillais avec les gars

Oui mais quel texte ?

Différent selon les usages

Loc 3: Mais...

Loc 2 : Et d'ailleurs euh, ici à l'usine, ils me prenaient pour un Arabe hein

Loc 1: Ben oui, tellement tu étais...

Loc 2 : Si je me chamaillais avec les gars hein

Entretien sociologie

1
00:00:06,951 —> 00:00:06,952
mais

2
00:00:06,952 —> 00:00:07,033
et d'ailleurs ici à l'usine ils me prenaient pour un Arabe

3
00:00:07,033 —> 00:00:07,034
ben oui tellement tu étais

4
00:00:07,051 —> 00:00:07,133
si je me chamaillais avec les gars

Sous-titrage

```
<Turn speaker="spk1" startTime="415.779"
endTime="417.154">
<Sync time="415.779"/>

<Event desc="rire" type="noise"
extent="instantaneous"/>

</Turn>
<Turn speaker="spk3" startTime="417.154"
endTime="417.654">
<Sync time="417.154"/>
mais
</Turn>
<Turn speaker="spk2" startTime="417.654"
endTime="422.192">
<Sync time="417.654"/>
et d'ailleurs euh euh et ici à l'usine ils me ils ils
me prenaient pour un Arabe hein
</Turn>
<Turn speaker="spk1" startTime="422.192"
endTime="423.152">
<Sync time="422.192"/>
ben ouais tellement tu étais
</Turn>
<Turn speaker="spk2" startTime="423.152"
endTime="428.385">
<Sync time="423.152"/>
ah si je me chamaillais avec les les gars hein
<Event desc="pause" type="noise"
extent="instantaneous"/>
</Turn>
```

Analyse du français parlé
(alg_jac_06.trs)

Quelle couverture ?

Selon besoins scientifiques

- Magali Husianyca *Caractérisation de types de discours dans des situations de travail*
 - 74 h 51 d'enregistrement
 - 27 h 22 de transcription
- Julie Glikman et al. *Les Vocaux 0.0.1*
 - 10 h 19 min 08 sec d'enregistrement
 - 10 h 19 min 08 sec de transcription

Une activité chronophage

Temps de transcription pour 1 h de signal

Rioufreyt 2016	Baude et al. 2006	Labex ASLAN
4 h - 6 h	30 h	6 h - 60 h

Comment on s'y prend ?

- Choix des conventions de transcription
- Choix du ou des logiciels
- Choix du matériel



Les erreurs fréquentes

Paul Cappeau (2008)

	Version initiale	Version éditée
Conventions	m'enfin	mais enfin
Oubli	vous l'acceptez vous l'acceptez pas	vous l'acceptez vous ou l'acceptez pas
Ajout	je ne me suis pas	je me suis pas
Transformation	vous avez la permission	avec la permission
Orthographe	on m'a rétablit les américains	on m'a rétabli les Américains

- Poids de la norme
- Manque de connaissances culturelles
- Notion d'Hallucination auditive/Hallucination de l'intelligence artificielle

Qui subsistent toujours un peu

#city et du coup comme j'étais dans le coin bah je suis allé aider mon père effectivement c l'instant ça va et euh et ouais ouais euh concernant #city où habitent euh #name et #name vallée quand on monte euh quand on monte direction le Grand Ballon par la vallée de #cit village qu'on croise quand on vient de de #city après #city il y a #city à l'entrée de la vallée #city c'est c'est déjà un petit peu paumé et et du coup c'est super long pour y aller à vélo tu trop mal et non effectivement alors a- par contre je crois quoi enfin je je moi je comprends si tu t'es déjà fait contrôler en extérieur avec un joint ouais effectivement le #name je vais c'est pas c'est pas c'est pas mm c'est pas rationnel comme euh comme raisonnement juste t #name est très embêté que que d'un coup euh mais bon donc je pense que #name elle a d

Méthodologie proposée

Déroulé complet

1. Sélectionner une ou plusieurs plateformes
2. Sélectionner des échantillons de données
3. Les transcrire manuellement
4. Les transcrire à l'aide des plateformes
5. Évaluer la qualité des résultats
6. Corriger les transcriptions automatiques
7. Évaluer le temps gagné/perdu
8. Estimer le coût de son projet

1. Sélectionner une ou plusieurs plateformes

- Vous pouvez vous appuyer sur la section IV du rapport technique « La transcription automatique : un rêve enfin accessible » (2020)
- Veillez à la sécurité des données (situation géographique des sièges sociaux; hébergement des données; réglementations à l'œuvre; traitement des données nominatives, techniques et transcrites; protocoles de cryptage)
- Comparez les services proposés (tarification; langues supportées; finesse de transcription; métadonnées; formats en entrée/sortie; éditeur de texte; alignement texte/son)

2. Sélectionner des échantillons de données

- Est-ce que toutes les données de mon étude vont être enregistrées avec le même instrument ?
- Dans le même environnement ? bruits de fond ? à distance ? À l'aide d'un téléphone ?
- Combien de locuteurs ? Monologue, entretien face à face, réunion...
- Est-ce que j'aurais plusieurs genres de discours ? Échanges informels, discours planifié ...
- Prévoir un échantillon de 2 minutes par type de situations

3. Les transcrire manuellement

- Selon vos habitudes de transcription
- En prenant soin de noter le temps que cela prend pour chaque échantillon



4. transcrire à l'aide des plateformes

Lancer une transcription

Transcrire

Aligner une transcription

Détection de langues

Document audio ou vidéo

extrait_alg_jac_06.wav

1. Sélectionnez la langue du document (Choisissez "Automatic" seulement si vous ne connaissez pas la langue du document)

Français

Les dernières langues utilisées : Français, Anglais, Français

Cochez cette case s'il s'agit de véritables enregistrements téléphoniques.

ⓘ Attention, les transcriptions issues de traitements d'enregistrements téléphoniques ne contiennent pas de ponctuation.

2. Options d'amélioration de la transcription

Ajouter du vocabulaire ⓘ

Contextualisation ⓘ

Annuler

Lancer la transcription

extrait_alg_jac_06.wav

4 avr. 2024, 16:56:18

∅

Consulter

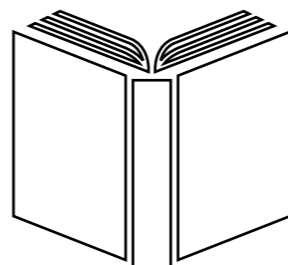
⋮



5. Évaluer qualité des résultats

- Pensez à harmoniser les transcriptions (apostrophes, heures, métadonnées, encodage de caractères)
- Mesure la plus utilisée : Word Error Rate (WER)
- Des alternatives dans la littérature : Information Mutuelle, WKER, H-score ...
- Une proposition de notre part : utilisation d'un outil antiplagiat pour identifier les différences entre transcriptions manuelles et automatiques + analyse des erreurs de transcription

COPYSCAPE

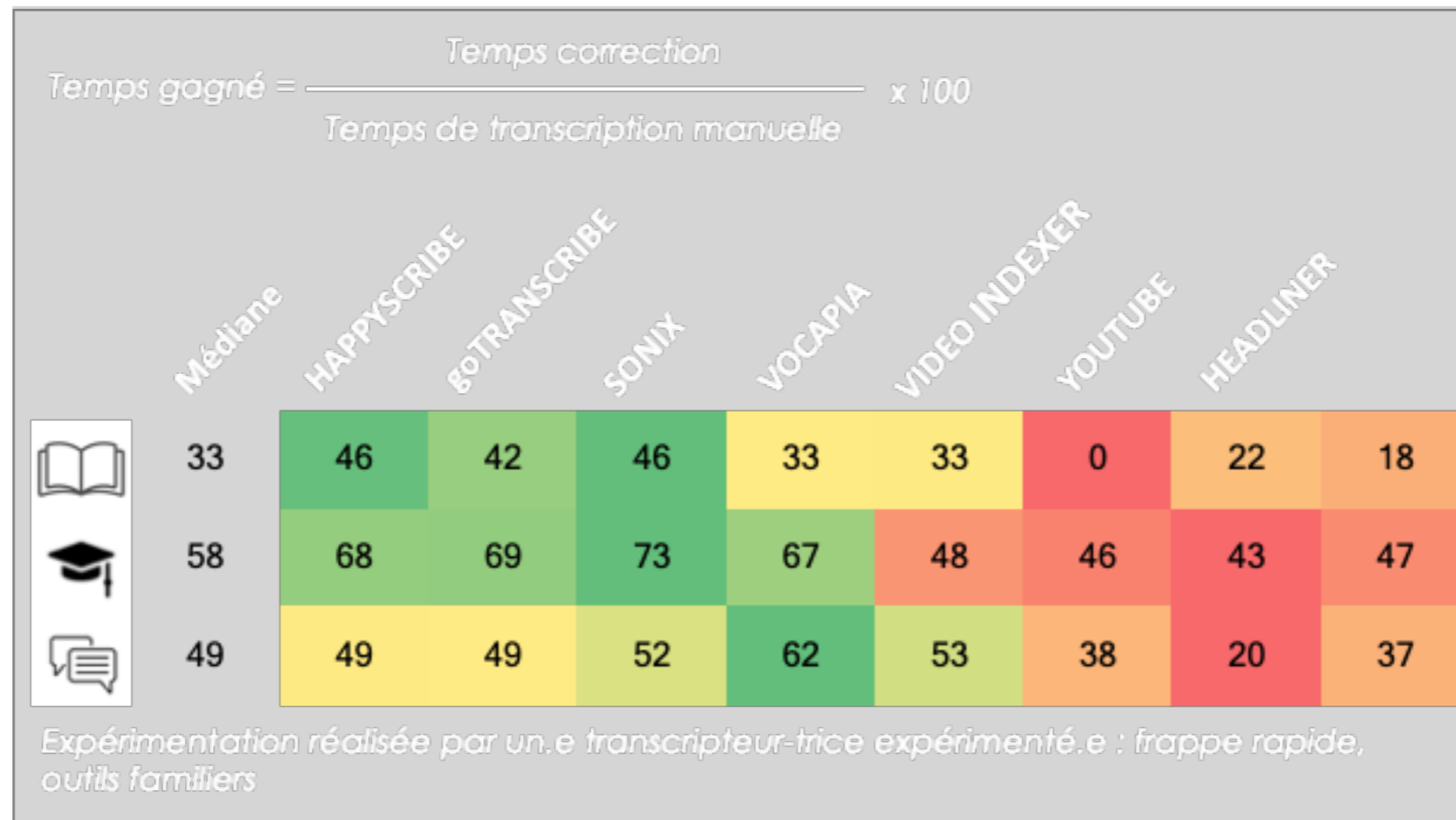


6. Corriger les transcriptions automatiques

- En utilisant les fonctionnalités d'édition disponibles sur les plateformes ou le logiciel que vous pensez utiliser dans le cadre de votre projet
- En prenant soin de noter le temps que cela prend pour chaque échantillon



7. Évaluer le temps gagné/perdu



- Temps de transcription manuelle/correction
- Sans temps de transcription automatique, ni manipulation des fichiers

8. Estimer le coût de son projet

À quoi faut-il penser ?

- Réunions sur le projet
- Rédaction des conventions
- Formation aux conventions du projet
- Formation aux logiciels
- Service de transcription
- Correction des transcriptions
- Préparation des données
- Manipulation des fichiers en entrée + en sortie

8. Estimer le coût de son projet

Exemples LABEX ASLAN

	Phénomènes transcrits/type de convention	Logiciel utilisé	Ratio TT h/ H de signal*	Coût horaire**	Remarques/ Exemple projet Aslan
1-Synopsis	Découpage de l'interaction, séquençage	Word/ Transana	2h	52€	Première phase de plusieurs transcriptions afin de repérer les passages à transcrire
2-Transcription au kilomètre	Tours de paroles	Word	8-12h	208€ - 312€	Premier jet de transcription qui peut éventuellement être repris pour une transcription plus fine. Ou pour un certain type de données du corpus, par ex. projet IMPEC pour la partie 'entretiens'
3-Correction Vocapia	Transcription automatique sur Vocapia Relecture, correction	XML/TXT	6-10h	156€- 260€	Permet d'automatiser une partie mais la phase de correction est nécessaire car peut provoquer de nombreux non-sens, par ex Projet ForPro pour une sous-partie 'entretiens'
7-Transcription ELAN	Grille de transcription Codage gestes, regards, posture et/ ou codage propre au projet	ELAN	45h-60h	1170€- 1560€	On compte 1h de travail pour 1 à 2 minutes de signal, en général une dizaine de tiers sont à coder. Par ex. projet Gedeco

C'est parti !

Merci

Références

- Analyse et traitement informatique de la langue française - UMR 7118 (ATILF) (2020). *TCOF : Traitement de Corpus Oraux en Français* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v2.1, <https://hdl.handle.net/11403/tcof/v2.1>.
- Baude O. et al., 2006. *Corpus oraux, guide des bonnes pratiques 2006*. Presses universitaires d'Orléans; CNRS Éditions, 203 p.
- Cappeau P., 2008. « Ce que nous apprennent les transcriptions », revue *Verbum* T. XXX, n°4. pp. 343-353
- Glikman J., et al., 2024. *LesVocaux* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v0.0.1, <https://hdl.handle.net/11403/lesvocaux/v0.0.1>.
- Husianycia M., 2011. *Caractérisation de types de discours dans des situations de travail*, thèse de doctorat, Université Nancy 2
- Rioufreyt, T., 2018. « La transcription outillée en SHS. Un panorama des logiciels de transcription audio/vidéo ». *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 139(1), 96-133.
- Tancoigne, É. et al., 2020. *La transcription automatique : un rêve enfin accessible ? Analyse et comparaison d'outils pour les SHS. Nouvelle méthodologie et résultats*. [Rapport de recherche] MATE-SHS.
- Tancoigne, É. et al., 2022. « Un mot pour un autre ? Analyse et comparaison de huit plateformes de transcription automatique ». *Bulletin de Méthodologie Sociologique / Bulletin of Sociological Methodology*, 155 (1), pp.45 - 81
- <https://www.youtube.com/watch?v=b7aKj9XdsPg>