

La transcription outillée en SHS. Un panorama des logiciels de transcription audio/vidéo

Bulletin de Méthodologie Sociologique

2018, Vol. 139 96–133

© The Author(s) 2018

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0759106318762455

journals.sagepub.com/home/bms**Thibaut Rioufreyt***Université catholique de Louvain, Belgium***Abstract**

This article proposes an overview of the main software available to the transcriber to facilitate its work and ensure the quality of transcripts. The goal is twofold. On the one hand, a summary of the existing tools (24 softwares are treated) and the main technical characteristics of each of them are available. There is indeed an overabundant offer in the field, to the point that it is difficult to navigate for the average user. On the other hand, this article does not simply propose a list of tools but nine simple criteria allowing to choose the software most adapted to the needs of its research, its corpus and its competences: role of the problematic and the research subject, specificity of the field, platform on which the tool works, availability of the software, mode of representation of the data, articulation of the supports, format of the data, tools of reading and treatment of the sound, functions of analysis proposed by these software.

Résumé

Cet article propose d'esquisser un panorama des principaux logiciels à la disposition du transcripateur permettant de lui faciliter le travail et de garantir la qualité des transcriptions. L'objectif visé est double. Il s'agit d'une part de mettre à disposition une synthèse des outils existants (24 logiciels sont traités) et des caractéristiques techniques principales de chacun d'entre eux. Il existe en effet une offre pléthorique en la matière, au point qu'il est difficile de s'y retrouver pour l'utilisateur lambda. D'autre part, cet article ne propose pas simplement une liste d'outils mais neuf critères simples permettant de choisir le logiciel le plus adapté aux besoins de sa recherche, à son corpus et à ses compétences : rôle de la problématique et de l'objet de recherche, spécificité du terrain,

Corresponding Author:

Thibaut Rioufreyt, Université catholique de Louvain (UCL)/ISPOLE, SPLE, Place Montesquieu, Louvain-la-Neuve, 1348, Belgium

Email: thibaut.rioufreyt@uclouvain.be

plate-forme sur laquelle fonctionne l'outil, disponibilité du logiciel, mode de représentation des données, articulation des supports, format des données, outils de lecture et de traitement du son, fonctions d'analyse proposées par ces logiciels.

Keywords

transcription, software, research data, qualitative methods, methodology

Mots clés

transcription, logiciels, données de la recherche, méthodes qualitatives, méthodologie

La transcription des entretiens est une activité centrale de l'enquête qualitative¹. Incontournable mais chronophage, elle est souvent préalable à l'analyse mais ses modalités de mise en œuvre conditionnent celle-ci. Les publications et manuels sur la manière de transcrire sont désormais nombreux. Cependant, nombre d'entre eux négligent encore largement l'usage de logiciels. Et lorsque c'est le cas, ils tendent à se restreindre à quelques-uns des outils existants². Dans cette perspective, cet article propose d'esquisser un panorama des différents logiciels à la disposition du transcripteur permettant de lui faciliter le travail et de garantir la qualité des transcriptions. L'objectif visé est double. Il s'agit d'une part de mettre à disposition une synthèse des outils existants et des caractéristiques techniques principales de chacun d'entre eux. Il existe en effet une offre pléthorique en la matière, au point qu'il est difficile de s'y retrouver pour l'utilisateur lambda. D'autre part, cet article ne propose pas simplement une liste d'outils mais une série de critères simples permettant de choisir le logiciel le plus adapté aux besoins de sa recherche, à son corpus et à ses compétences. Entrer dans le processus de transcription, c'est d'abord faire des choix : Quels phénomènes transcrire ? Comment les transcrire ? Comment matérialiser la production orale ? Quelle forme visuelle choisir ? Le propos ici n'est pas de donner les « bonnes réponses » ; il n'y en a pas, sauf à faire preuve d'un impérialisme méthodologique ou d'une cécité disciplinaire rares. En revanche, il y a de bonnes questions à se poser. Plus exactement, il y a quelques critères à avoir en tête (seront ici abordés 9 critères généraux déclinés en sous-critères) pour pouvoir choisir de manière pertinente : le rôle de la problématique et de l'objet de recherche (1), la spécificité du terrain (2), la plateforme sur laquelle fonctionne l'outil (3), la disponibilité du logiciel (4), le mode de représentation des données (5), l'articulation des supports (6), le format des données (7), les outils de lecture et de traitement du son (8) et les fonctions d'analyse proposées par ces logiciels (9).

Avant d'entrer dans le vif du sujet, quelques précisions sont nécessaires. Premièrement, le propos sera centré sur les logiciels de transcription *stricto sensu* (de l'audio ou de la vidéo vers l'écrit ou en tout cas d'alignement audio ou vidéo-textuel). Ont ainsi été retenus les logiciels de transcription manuelle généraliste ainsi que les logiciels multimodaux qui, tout en permettant une transcription fine des sources audio et/ou vidéo, proposent également des fonctionnalités plus avancées d'annotation et d'exploration des corpus multimodaux (c'est-à-dire comprenant plusieurs modes de communication comme l'écrit, la parole ou l'image)³. Des solutions techniques pour la transcription seront donc ici mises de

côté et méritent à n'en pas douter un traitement spécifique, qu'il s'agisse de l'usage averti du couplage d'un logiciel de traitement de texte et d'un logiciel de lecture audio/vidéo⁴ ou encore des logiciels de reconnaissance vocale⁵. D'autres outils à la disposition du transcritteur comme les pédaliers de transcription ou les dispositions de clavier alternatives ne seront abordés que de manière marginale. Par ailleurs, certains CAQDAS (*Computer-assisted qualitative data analysis software*) comme par exemple MAXQDA permettent également de réaliser des transcriptions. Outre le fait que les publications méthodologiques sur ces outils sont plus nombreuses et facilement accessibles, ces derniers relèvent d'une démarche différente (le codage interprétatif des matériaux et non simplement leur transcription en vue de l'analyse) qui aurait exigé d'amples développements, développements réalisés au détriment de la présentation des outils plus directement dédiés à la seule transcription. Enfin, le choix a été fait ici de ne pas traiter des outils de *re-transcription* (d'un texte manuscrit ou tapuscrit papier à sa version numérique ou d'un fichier texte nativement numérique à une version enrichie)⁶ qui s'inscrivent dans une approche également différente ; cela aurait en particulier impliqué de revenir sur les opérations de traitement numérique des données textuelles (comme la numérisation, l'océrisation⁷, la conversion sous des formats pérennes ou encore l'encodage sous un standard de métadonnées dont le plus connu est la TEI⁸).

Deuxièmement, même en se focalisant sur les logiciels de transcription, il serait présomptueux de prétendre traiter ici de tous les logiciels, et davantage encore de toutes leurs fonctionnalités. Ma contribution ne prétend donc pas à l'exhaustivité. Elle portera sur 24 logiciels, sélectionnés à la fois parce que ce sont les plus utilisés et qu'ils sont développés par une entreprise ou une communauté qui garantit *a minima* leur relative pérennité⁹. Je ne suis pas un spécialiste de tous les logiciels traités. Il convient en outre de préciser *d'où je parle*, depuis quelle position disciplinaire, méthodologique et technique. Je suis en l'occurrence un politiste, qualitatif de formation et formé à la science politique dans une optique fortement marquée par la sociologie. Je suis par ailleurs un utilisateur de PC, rétif vis-à-vis des Mac (tout en leur reconnaissant certaines vertus), utilisant Linux et dans une moindre mesure Windows. Enfin, j'ai transcrit pour l'essentiel des entretiens menés dans le cadre d'enquêtes individuelles et collectives. C'est pour cette raison que je connais mieux certains logiciels de transcription que d'autres et surtout que je viens d'une tradition de transcription *située*. Par conséquent, je connais et utilise régulièrement seulement une petite partie d'entre eux. J'en ai téléchargé d'autres afin de pouvoir les tester en vue de la rédaction de cet article. Enfin, certains autres ont été ajoutés afin de compléter le panorama, et notamment par souci de prendre en compte les utilisateurs de systèmes d'exploitation comme Mac.

Troisièmement, il s'agit d'un *work in progress* et en aucune façon d'un discours définitif sur ces outils. D'une part, refusant de payer certains logiciels propriétaires (Express Scribe, Transcribe, Eurescribe, InqScribe, Transcription Helper) ou d'indiquer mes données personnelles pour pouvoir accéder au manuel d'utilisation, je n'ai pu vérifier l'intitulé du format natif¹⁰ ou la véracité des certaines informations sur les formats et l'interopérabilité¹¹. D'autre part, en matière informatique les logiciels, les formats ou encore les langages évoluent rapidement : ce qui est observé dans le cadre de cet article est donc susceptible d'évoluer ultérieurement¹².

Quatrièmement, enfin, ce papier a bénéficié des conseils et suggestions de plusieurs collègues chercheurs, enseignants-chercheurs et ingénieurs¹³. Il reprend également les apports d'un certain nombre de publications comparant certains de ces outils (Rohlfing et al., 2006 ; Parisse et Morgenstern, 2010 ; Fiorèse, 2016) ou travaux de repérages publiés sur le web¹⁴. À ce titre, c'est déjà un travail collectif. Si je garde le monopole du blâme (toutes les éventuelles erreurs ou approximations commises ici sont miennes), je ne saurais prétendre à celui de la louange. Cette contribution se veut simplement une première base d'informations et de discussion permettant aux collègues spécialistes de tel ou tel outil ou méthode de pouvoir amender, compléter, contribuer à ce panorama.

I. Le rôle de la problématique et de l'objet de la recherche

Transcrire, ce n'est pas simplement écouter et mettre sur le papier ce qu'on entend (Blanche-Benveniste et Jeanjean, 1987 ; Green, Franquiz et Dixon, 1997). En ce sens, le mythe de la transcription comme reproduction fidèle et du transcripteur comme copiste est bel et bien révolu (Dister et Simon, 2008 : 76). Tout corpus est une construction, au sens où il est toujours déjà le produit de l'ancrage méthodologique et théorique du chercheur. Il n'y a pas de transcription neutre puisqu'elle dépend de ce qui est recherché dans le corpus, des problématiques du chercheur, de ses choix méthodologiques, etc. (Ochs, 1979). En ce sens, le premier principe de la transcription est qu'elle n'est pas qu'un préliminaire à l'analyse mais qu'elle en fait véritablement partie. Ce principe est encore plus pertinent lorsque l'on utilise certains logiciels qui ont des fonctionnalités d'annotation et d'exploration intégrées. Par conséquent, il existe un grand nombre de formes de transcription en fonction des buts scientifiques visés, mais aussi en fonction de la discipline, de la méthodologie ou de l'épistémologie sous-jacentes (Mondada, 2008a).

Dans un premier temps, il s'agit pour le transcripteur de se poser la question de la « granularité de la transcription » (Mondada, 2008b : 151)¹⁵, c'est-à-dire de faire des choix sur le nombre, la précision et les types de phénomènes à transcrire. Cela revient à décider jusqu'où aller dans les détails, ce d'autant plus que l'on peut toujours recourir, grâce à certaines fonctionnalités comme le minutage interactif ou l'alignement (voir *supra*), aux données sources. Ce choix dépend en réalité de ce qui est recherché dans les matériaux empiriques. Les transcriptions peuvent se faire à de multiples niveaux du langage : phonétique, phonologique, lexicale, orthographique, syntaxique, sémantique, actes de langages, pragmatique, mimo-gesto-posturale (gestes, pointages, *hexis*, etc.), interactions avec autrui, avec la scène, etc. Ces différents niveaux correspondent schématiquement à six sous-contextes, présentés ici du plus micro au plus large :

- le cotexte ou contexte immédiat, *i.e.* l'environnement sémiotique voisin du signe au sein de la même unité, le plus souvent la phrase ou le paragraphe ;
- le contexte linguistique, *i.e.* le système virtuel de la langue, comme l'anglais, le français, le latin de telle époque ;
- le contexte situationnel, *i.e.* les circonstances dans lesquelles les locuteurs sont en relation, particulièrement étudié par la pragmatique ;
- l'intertexte, *i.e.* l'ensemble des énoncés effectifs auquel l'énoncé se réfère de manière implicite ou explicite, consciente ou non, qui le domine et à partir duquel il fait sens ;

- le discours, *i.e.* un ensemble d'énoncés spécifiques et socialement différenciés, qui renvoient par ailleurs à des degrés de généralité plus ou moins grands : formation, champ, genre et registre discursifs ;
- l'interdiscours, *i.e.* les relations entre unités discursives¹⁶ qui finissent par correspondre au « contexte large », soit tout « l'arrière-plan socioculturel des co-énonciateurs » (Joly et O'Kelly, 1990 : 79)

L'articulation de ces six contextes permet d'esquisser les contours du contexte sémiotique dans lequel prennent place les énoncés des enquêtés. Selon le chercheur, la méthode, la théorie ou la discipline, on privilégiera plutôt un aspect sans pour autant sacrifier les autres. Un peu comme la distribution des rôles au théâtre, on détermine qui va jouer le premier rôle et qui seront les personnages secondaires.

Tableau 1. Types d'approche

LOGICIEL	TYPE D'APPROCHE
ANVIL	Phonétique ; Linguistique interactionnelle
Casualtranscriber	Transcription de contenu
CLAN	Acquisition du langage chez les enfants ; Analyse conversationnelle ; Etude de récit ; Linguistique interactionnelle
EasyTranscript	Transcription de contenu
ELAN	Interactions multi-locuteurs ; Transcription du langage des signes ; Documentation des langues en voie de disparition
EureScribe	Transcription de contenu
EXMARALDA	Analyse du discours ; Sociolinguistique ; Linguistique interactionnelle ; Dialectologie
Express Scribe	Transcription de contenu
F4&F5 Transkript	Transcription de contenu
FOLKER	Analyse conversationnelle ; Linguistique interactionnelle
HyperTranscribe	Transcription de contenu
InqScribe	Transcription de contenu
Listen N Write	Transcription de contenu
otranscribe	Transcription de contenu
PHON	Phonologie (Développement phonologique (incluant le babillage), acquisition d'une seconde langue, troubles phonologiques, phonologie des emprunts); Phonétique ; Sociolinguistique
PmTrans	Transcription de contenu
PRAAT	Phonétique
Sonal	Transcription de contenu ; Codage audio-textuel ; Annotation (proche d'un CAQDAS)
TRANSANA	Linguistique interactionnelle ; Analyse conversationnelle
Transcribe	Transcription de contenu
Transcriber AG	Transcription de contenu ; Etude des récits ; Analyse des interactions ; Annotation de corpus oraux
Transcription Helper	Transcription de contenu
Transcriber Pro	Transcription de contenu
Transcriva	Transcription de contenu

La linguistique dans sa version logico-grammaticale (Rastier, 1998) ou la philologie tendent ainsi à étudier préférentiellement, voire parfois exclusivement, le cotexte et la langue. La phonologie s'intéresse à l'agencement des phonèmes d'une langue pour former des mots, la phonétique aux sons eux-mêmes (les unités phonétiques, les « phones »), à leur production, leur variation. Dans cette perspective, certains logiciels très spécialisés comme PRAAT et PHON sont parfaits pour l'utilisateur qui désire faire de la phonétique ou de la phonologie.

La sociologie, la science politique ou l'histoire privilégient, quant à elles, plutôt le contenu des énoncés des enquêtes en le rapportant au contexte discursif et inter-discursif propres à une époque et/ou à un groupe social donnés. Il s'agit d'étudier les informations manifestes contenues dans les propos de l'enquêté pour par exemple en savoir plus sur un événement, le fonctionnement d'un groupe, etc. Il peut s'agir également d'étudier plutôt le sens, donc y compris le contenu latent, comme les schèmes de représentation et de perception, les catégories que mobilise l'enquêté. Certes, le contexte d'énonciation intéresse aussi le *social scientist* ; celui-ci va donc intégrer du *paratexte* dans la transcription sous forme de passages balisés entre parenthèses ou entre crochets du type [Il rit], [long silence], etc. De même, la manière dont c'est dit est intéressante car cela révèle du sens. L'usage intensif du « y » marquant un registre langagier populaire, ou un accent campagnard, sont de précieux indices qui signent les origines sociales du locuteur. Il reste que le contexte d'énonciation comme le discours ne constituent pas l'objet de la recherche ; ce n'est pas la fin mais le moyen de la recherche. Ils sont donc indiqués de manière plus sommaire que par exemple en linguistique. Dans cette approche, les logiciels de transcription de type généraliste proposant simplement un minutage interactif facilitant la navigation et l'exploration du corpus sont tout à fait adaptés.

Entre ces deux pôles se trouvent des formes intermédiaires de contextualisation : situation d'énonciation pour la pragmatique, de la situation à l'interdiscours pour l'analyse de discours, la sémantique interprétative et la socio-linguistique. Dans le cas de l'analyse conversationnelle (AC) ou de la linguistique interactionnelle, on pose la question des contenus mais aussi de la façon dont l'interaction se co-construit, et des processus interactifs (Sacks, Schegloff, et Jefferson, 1974 ; Vincent, 2001). Loin d'être réservée à la linguistique ou à l'AC proprement dite, l'analyse des interactions peut intéresser le sociologue qui souhaiterait étudier l'effet enquêteur sur les réponses de l'interviewé. Tous les phénomènes peuvent être pertinents, il est donc judicieux et nécessaire d'adopter une transcription fine du verbal incluant par exemple les tronctions, les pauses, les hésitations, etc. L'analyse des interactions, a fortiori lorsque l'on adopte une approche multimodale, est importante. ELAN (EUDICO Linguistique Annotator) est sans doute le logiciel le plus adapté lorsqu'il s'agit d'aligner plusieurs lignes de transcription (plusieurs interlocuteurs ou/et plusieurs événements) de manière temporelle précise. CLAN (Computerized Language Analysis) est particulièrement indiqué dans le traitement de récit (comme Transcriber AG) mais aussi à la description des interruptions, reprises, erreurs de performance ou variantes de prononciation grâce à son riche système de codage. EXMARaLDA (Extensible Markup Language for Discourse Annotation) ou FOLK (FOLK Editor) présentent, quant à eux, une configuration intermédiaire entre ces deux pôles. Dans le cas de l'analyse du discours, ce qui est recherché dans la transcription, ce n'est plus seulement le contenu mais le lexique, la

syntaxe, la grammaire, etc. Cela a des effets très directs sur le mode de transcription. L'orthographe standard sera ainsi souvent privilégiée (voir *supra*), ce qui offre davantage de lisibilité qu'une restitution stricte et permet les analyses automatiques (lexico-métriques ou textométriques) basées sur des conventions orthographiques standards.

2. La spécificité du terrain

La spécificité du terrain est un autre critère déterminant dans le choix du logiciel. Ce critère renvoie à la fois au type de locuteurs et aux conditions d'enregistrement des données. En premier lieu, certains locuteurs ayant des caractéristiques vocales spécifiques rendent parfois l'écoute, et donc la transcription, plus malaisée : accents locaux plus ou moins prononcés, variation dans la modulation de la voix (voix chevrotante de certaines personnes âgées, énonciation dans des situations affectives critiques : pleurs, angoisse, sidération, etc.), problèmes de diction (déformation palatine, bégaiement, zozotement), etc. Certains enquêtés/enquêteurs parlent parfois très rapidement, voire trop vite. D'autres ont tendance à mâcher leurs mots. Dans ce cas, les logiciels proposant des outils de traitement du son (fréquence, débit, etc.) peuvent parfois faciliter le travail du transcripateur (voir *infra*). Une fonctionnalité offerte par la plupart des logiciels présentés consiste par exemple à ralentir la vitesse de lecture de la bande et ainsi de découper artificiellement l'enchaînement des phonèmes de manière à comprendre certains mots ou bouts de phrases inaudibles. D'autres interviewés encore ont un débit et une articulation normaux mais la fréquence de leur voix fait qu'on distingue mal certains passages. Dans ce cas, certains logiciels comme PRAAT offrent au transcripateur la possibilité de modifier la tonalité des voix et parvenir à distinguer et reconnaître des mots jusque-là inaudibles. De telles fonctionnalités s'avèrent en outre très pratiques dans certains cas d'enregistrements sonores de mauvaise qualité qui produisent parfois des voix modifiées (très aiguës ou très basses, ralenties) à cause de l'usure de la bande magnétique. Enfin, de manière plus ciblée, certain.e.s enquêté.e.s peuvent présenter des spécificités langagières : bébés, jeunes enfants, adultes, étrangers, personnes âgées, handicapés physiques et/ou mentaux, etc. Pour le traitement du langage de l'enfant, le logiciel aujourd'hui le plus adapté est CLAN (Parisse et Morgenstern, 2010). CLAN, conçu pour traiter les corpus servant à étudier l'acquisition du langage permet de calculer la LME (Longueur Moyenne d'Énoncé), mesure proposée par Brown (Brown, 1973) et systématiquement utilisée dans la littérature sur l'acquisition comme permettant d'évaluer le niveau approximatif de développement de langage d'un enfant.

Deuxièmement, les conditions de recueil des sources orales modifient également la manière de transcrire et les codes utilisés. La majorité des entretiens et observations en SHS est réalisée sur le terrain (sur les lieux où évoluent les enquêtés) ou en tout cas en dehors du laboratoire (lieu de rendez-vous convenu entre enquêtrices/eurs et enquêté.e.s). Dès lors, le lieu où est réalisé l'enregistrement des données orales ne dépend que partiellement de la volonté de l'enquêteur. Le fait qu'il ait lieu dans un lieu plutôt calme ou au contraire dans un endroit bruyant a des effets sur le mode de transcription et peut jouer dans le choix des outils. Dans ce dernier cas, un traitement du signal audio (comme l'élimination du souffle ou du bruit de fond sur la bande) serait sans doute

bienvenu, que ce soit par des outils spécialisés dans cette tâche avant la transcription – comme par exemple Audacity (<https://audacity.fr.softonic.com/>) – ou par des logiciels de transcription proposant de telles fonctionnalités (voir *infra*). Bien que plus marginalement, les entretiens peuvent également être effectués en laboratoire et/ou dans le cadre d'un protocole expérimental (captation vidéo, mise en situation, jeux de rôles, etc.). Utilisé souvent en linguistique, en sciences cognitives ou en information-communication pour analyser les interactions en situation (entre humains et/ou humains-machines), ce type de dispositifs implique dès la préparation du projet de réfléchir au mode de transcription. Le choix de logiciels multimodaux proposant l'alignement audio-véo-textuel comme CLAN, ELAN ou EXMARaLDA est le plus adapté et est, de fait, le plus récurrent.

3. La plate-forme

Le troisième paramètre à regarder avant d'adopter tel ou tel logiciel, c'est évidemment la plate-forme, c'est-à-dire à la fois : 1) le **système** (PC/Mac) ; 2) la **configuration minimale requise** du système d'exploitation (OS) ; 3) la **localisation du programme** (disque dur/navigateur web/application smartphone) sur lequel il fonctionne.

Le premier paramètre technique à envisager ici est le **système** (PC et/ou Mac) sur lequel fonctionne le logiciel. 2 des logiciels référencés ne fonctionnent que sous Mac, 5 exclusivement sous PC et les 17 autres sur les deux systèmes. Le deuxième paramètre technique concerne le **système d'exploitation** (par la suite « OS » dans l'article)¹⁷ qu'exige le logiciel. L'une des règles d'or lorsque l'on cherche un logiciel pour sa recherche – pour la transcription comme pour d'autres finalités – est de ne jamais oublier que les informations disponibles sur le site de chaque logiciel ne sont pas toujours fiables ou exactes. Qu'il s'agisse de logiciels payants ou gratuits, chaque promoteur de logiciel a tendance à dire qu'il fait un peu tout et qu'il marche sur n'importe quel système ou OS. En réalité, les choses sont souvent un peu plus compliquées. Si l'installation est possible, certaines fonctionnalités ne sont ainsi disponibles que sous un système ou OS. De même, certains logiciels sont promus comme fonctionnant sous Linux. Sauf que Linux connaît plein de versions différentes : Ubuntu ; Xubuntu, etc. Par ailleurs, il existe parfois une confusion entre Linux et Unix (Mac OS fonctionne sous une base Unix). Le dernier paramètre technique à prendre en considération concernant la plateforme est la **localisation du programme**. Trois des outils recensés ici sont des applications en ligne (Easy-Transcript, oTranscribe ou Transcribe). Attention : lors de l'utilisation d'un logiciel en ligne, il faut veiller à le faire depuis une connexion stable et également vérifier que son navigateur habituel est compatible avec celui-ci. À titre d'exemple Transcribe ne fonctionne sans problèmes qu'avec Google Chrome.

Par ailleurs, les outils en ligne nécessitent certaines précautions en termes de sauvegarde. oTranscribe sauvegarde automatiquement dans l'espace de stockage du navigateur chaque seconde. Toutefois, il perd parfois le travail effectué pour diverses raisons, notamment une modification du cache sur le navigateur web de l'utilisateur ou bien une mémoire locale pleine. Étant un serveur en ligne, il n'y a pas de fonction « Enregistrer ». La seule possibilité de sauvegarder le travail est donc de l'exporter. D'où la nécessité absolue d'exporter régulièrement et de ne pas laisser une transcription

Tableau 2. La plate-forme

PLATE-FORME			
LOGICIEL	Ordinateur	Configuration minimale OS	Localisation
ANVIL	PC ; Mac	Windows ; Macintosh ; Solaris ; Linux	Disque dur
Casualtranscriber	Mac	Mac OS X 10.11	Disque dur
CLAN	PC ; Mac	Windows ; Macintosh ; Unix	Disque dur
EasyTranscript	PC ; Mac	Windows (Windows XP SP 2) ; Mac (OS X 10.7) ; Linux (Ubuntu 12.04 LTS, Fedora 1)	Disque dur ; application smartphone
ELAN	PC ; Mac	Windows ; Macintosh ; Linux	Disque dur
EureScribe	PC	Windows Vista ; Windows XP ; Windows 2000	Disque dur
EXMARALDA	PC ; Mac	Windows ; Macintosh ; Linux	Disque dur
Express Scribe	PC ; Mac	Windows 95-Vista et + ; Linux GTK Native ; Mac OS X 10.1+	Disque dur
F4&F5 Transkript	PC ; Mac	F4Transkript : Windows (Vista, 7, 8, 10+) et Linux (Ubuntu 14.04, 15.10). F5Transkript pour Mac (OS X 10.7)	Disque dur
FOLKER	PC ; Mac	Windows ; Mac ; Linux	Disque dur
HyperTranscribe	PC ; Mac	Windows (XP, Vista, 7, 8.x, 10) ; Apple OS X (de 10.5.8 à 10.12)	Disque dur
InqScribe	PC ; Mac	Windows ; Macintosh	Disque dur
Listen N Write	PC	Windows XP/Vista/7/8/10	Disque dur (une version portable sur clé USB existe)
otranscribe	PC ; Mac	—	Service web (Chrome 31+ and Safari 7+) + application Google
PHON	PC ; Mac	Windows ; Mac ; Linux	Disque dur
PmTrans	PC ; Mac	Windows ; Mac ; Linux	Portable (sans installation)
PRAAT	PC ; Mac	Windows ; Macintosh ; Linux ; Unix ; SGI ; Solaris ; HP/UX	Disque dur
Sonal	PC ; Mac	Windows ; Macintosh (mais instable sous Mac)	Disque dur
TRANSANA	PC ; Mac	Windows ; Macintosh	Disque dur
Transcribe	PC	Windows NT/2000/XP	En ligne (ne fonctionne sans problèmes qu'avec Google Chrome)
Transcriber AG	PC ; Mac	Windows 2000 ; Mac ; Linux	Disque dur
Transcription Helper	PC	Windows 2000/XP/2003/Vista/ Server 2008/7	Disque dur
Transcriber Pro	PC	Windows 10 ; Windows 10 Mobile (x86, x64, ARM)	Disque dur ; application smartphone
Transcriva	Mac	Mac OS-X 10.5	Disque dur/en ligne

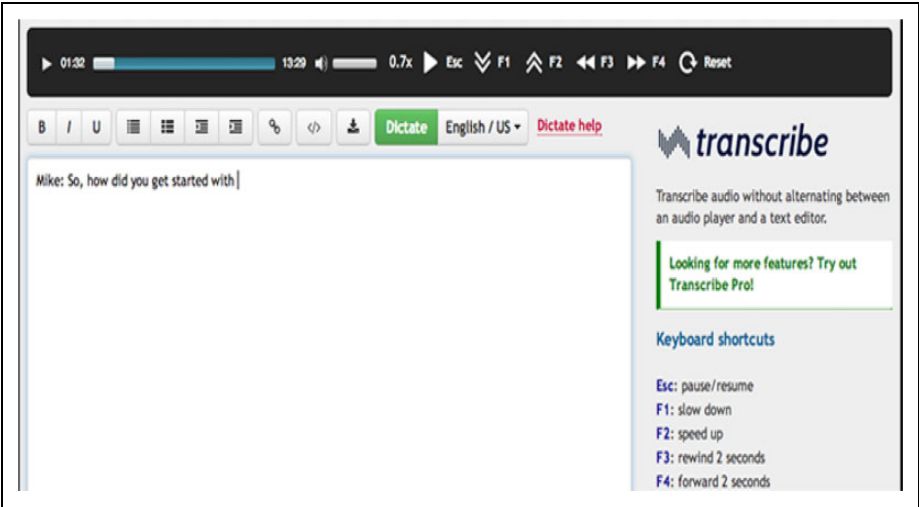


Figure 1. Le logiciel Transcribe (Source: <https://transcribe.wreally.com/>)

non exportée sur un ordinateur pendant des heures. Enfin, tout comme le choix de l’OS ou de son navigateur web, la localisation du programme n’est pas simplement une question technique ; elle est tout autant politique et déontologique. À titre d’exemple, on sait que Google sur lequel fonctionne certains des logiciels présentés comme Transcribe duplique l’ensemble des données des utilisateurs dans ses *data centers*. De même, oTranscribe garantit le respect du caractère **privé des données** : les fichiers audio et les transcriptions ne quittent jamais l’ordinateur. Toutefois, il est lié à une application Google sur smartphone. Ce point est central car les données dans les entretiens ne concernent pas seulement les chercheurs mais le droit des enquêtés à voir leurs données personnelles ne pas être *crawlées* par Google.

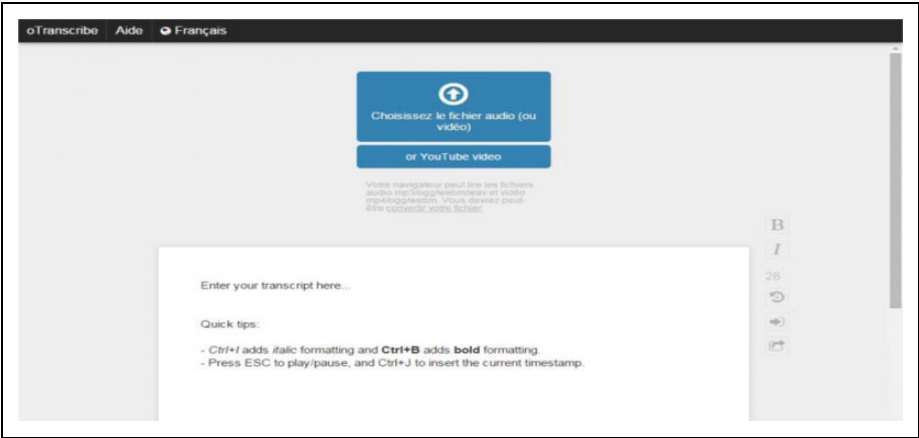


Figure 2. Le logiciel oTranscribe (Source: <https://il.wp.com/outilstice.com/wp-content/uploads/2014/01/oTranscribe.jpg>)

Les autres logiciels sont installés uniquement sur le disque dur de l'ordinateur et n'ont pas besoin de connexion Internet. Transcriba propose, quant à lui, les deux options : sur le disque dur ou sur le serveur web. Par ailleurs, trois d'entre eux proposent également des applications smartphone (EasyTranscript, oTranscribe, Transcriber Pro). Enfin, deux des logiciels de transcription, parmi les plus simples, proposent également une version portable sur clé USB sans installation : Listen N Write et PM Trans (The Poor's Man Transcriber). Toutefois, il convient de ne pas enregistrer ses transcriptions sur la clé USB car ce n'est pas un support fait pour la sauvegarde mais uniquement pour le transfert (contrairement à un disque dur externe par exemple).



Figure 3. Le logiciel Listen N Write (Source : http://telecharger.cnet.com/Listen-N-Write/3000-2170_4-75416316.html)

Hormis les précautions en matière de sauvegarde et de protection des données, il n'y a pas de « bonne » solution mais plutôt des usages. Le choix de tel ou tel outil dépend avant tout des habitudes de travail du transcripteur. Certains ont l'habitude de travailler en ligne ou de recourir beaucoup à leur smartphone. D'autres transcrivent dans les transports ou chez eux, etc. Il revient donc à chaque transcripteur (s'il travaille seul) de choisir l'outil qui correspond à ses usages.

4. La disponibilité

Le quatrième paramètre est la **disponibilité** du logiciel, notion sous laquelle j'englobe trois dimensions :

- l'**accès**, c'est-à-dire la réponse à la question « où puis-je le trouver ? » : le plus souvent téléchargeable directement en ligne sur le site du développeur, plus rarement, il faut le demander au concepteur ;
- la **licence**, c'est-à-dire savoir si le logiciel est propriétaire ou bien *open source*, c'est-à-dire dont le code est ouvert et accessible, ce qui lui garantit une certaine pérennité et la possibilité de l'adapter à ses propres besoins si l'on a des compétences en termes de codage ou que l'on travaille avec un ITA qui les possède ;
- le **coût**, c'est-à-dire savoir si le logiciel est gratuit et/ou payant.

Tableau 3. La disponibilité

LOGICIEL	DISPONIBILITÉ		
	Accès	Licence	Coût
ANVIL	http://www.anvil-software.de/	Open Source	Gratuit
Casualtranscriber	https://sites.google.com/site/casualconc/utility-programs/casualtranscriber	Propriétaire	Gratuit
CLAN	http://alpha.talkbank.org/clan/	Propriétaire	Gratuit
EasyTranscript	http://www.e-werkzeug.eu/index.php/en/	Open source (GNU GPLv3 licence)	Gratuit
ELAN	https://tla.mpi.nl/tools/tla-tools/elan/	Propriétaire	Gratuit
EureScribe	http://www.eurescribe.com/	Propriétaire	Payant
EXMARALDA	http://exmaralda.org/ Téléchargement avec un mot de passe fourni par le développeur	Open Source	Gratuit
Express Scribe	http://www.nch.com.au/scribe/fr/	Propriétaire	Version gratuite/ Version payante (Express Scribe=Pro)
F4&F5 Transkript	https://www.audiotranskription.de/english/f4.htm	Propriétaire	Payant
FOLKER	http://agd.ids-mannheim.de/folker_en.shtml	Propriétaire	Gratuit
HyperTranscribe	http://www.researchware.com/products/hypertranscribe.html	Propriétaire	Version d'essai gratuite/Version payante
InqScribe	https://www.inqscribe.com/	Propriétaire	Version gratuite/ Version payante
Listen N Write	http://elefantsoftware.weebly.com/listen-n-write.html	Propriétaire	Gratuit
otranscribe	http://otranscribe.com/	Open source (sous licence MIT)	Gratuit
PHON	https://www.phon.ca/phontrac	Open Source	Gratuit
PmTrans	https://github.com/juanerasmoe/pmTrans	Open source (GNU GPLv3 licence)	Gratuit
PRAAT	www.praat.org	Propriétaire	Gratuit
Sonal	http://www.sonal-info.com/	Propriétaire	Gratuit
TRANSANA	http://www.transana.org/	Propriétaire	Payant
Transcribe	https://transcribe.wreally.com/	Propriétaire	Payant
Transcriber AG	http://transag.sourceforge.net/	Open source (GNU GPLv3 licence)	Gratuit
Transcription Helper	http://www.easy-to-use-software.com/	Propriétaire	Payant
Transcriber Pro	http://www.hasfocus.com/transcribe-pro-windows-10-transcription-software/	Propriétaire	Payant
Transcriva	http://www.bartastechnologies.com/products/transcriva/	Propriétaire	Version d'essai gratuite/Version payante

Il existe bien souvent une confusion entre la licence et le coût, qui vient de la polysémie du terme « libre » (et plus encore de « *free* » en anglais). Nombre de sites référencent certains des logiciels comme étant « *free* », désignant tantôt qu'ils sont gratuits, tantôt que ce sont des *freeware* (gratuits mais propriétaires), tantôt enfin qu'ils sont *open-source*. De même, les *shareware* étant gratuits temporairement, le temps de l'essayer, contrairement aux *payware*, pour lesquels il faut payer d'emblée, ils sont confondus avec des *freeware*.

Tableau 4. Licence/Coût des logiciels

	Propriétaire	Open-source
Payant	Payware ou Shareware/partagiciel/contribuciel	Logiciel libre/Free Software
Gratuit	Logiciel gratuit/gratuiciel/Freeware	Logiciel libre/Free Software

Concernant le coût, il existe des logiciels qui ont une version gratuite et une version professionnelle payante (comme Express Scribe et InqScribe) ou encore des logiciels qui ont des versions gratuites pour un temps donné, c'est-à-dire des *shareware* comme Transcriva ou HyperTranscribe. Dans ce cas, il convient de vérifier quelles options (fonctionnalités, formats, compatibilité avec les logiciels extérieurs, etc.) existent dans la version gratuite et/ou temporaire et celles qui au contraire sont réservées à la version payante. Là encore, il convient de se méfier des effets d'annonce et de regarder dans le détail. Les versions d'essai sont pratiques pour tester un outil avant d'éventuellement l'acheter. En revanche, les utiliser pour une recherche est fortement déconseillé. Des imprévus dans la tâche de transcription qui retardent le planning prévu peuvent toujours surgir alors que les périodes d'essai sont souvent courtes (en général 30 jours). Ensuite, toutes les fonctionnalités ne sont pas offertes dans les versions d'essai. Enfin et surtout il existe des logiciels gratuits qui évitent d'avoir à recourir à cette solution. Dans l'ensemble je suggère, lorsque c'est possible, de recourir à des logiciels *open source*, ceux-ci garantissant bien mieux que leurs homologues propriétaires, à la fois la pérennité¹⁸ et la protection de vos données¹⁹.

5. Représentation graphique et visuelle des données

Il n'existe pas de représentation « parfaite » des données transcrites, mais toute une série de modèles qui sont plus ou moins adaptés aux différents usages que l'on peut en faire. Dans cette perspective, on peut distinguer la *représentation graphique* et la *représentation visuelle* des données orales.

En premier lieu, le transcripteur doit faire un choix quant au type de représentation graphique utilisée pour transcrire l'oral. Trois principales possibilités s'offrent en l'occurrence à lui : l'*orthographe adaptée*, l'*orthographe standard* et la *transcription phonétique/phonémique*. La première a l'avantage d'être assez proche de l'énonciation orale. En revanche, elle peut s'avérer difficile à lire notamment lorsqu'il y a beaucoup d'élisions. Par ailleurs, elle peut donner parfois une image péjorative du locuteur (Blanche-Benveniste, 1997). L'orthographe standard, quant à elle, est facile à lire puisqu'elle correspond à nos habitudes de lecture ; elle ne produit de ce fait pas d'effet

péjoratif. En revanche, elle lisse l'oral et donne une « version optimiste » (Blanche-Benveniste, 1997) de la production orale des locuteurs. Elle offre davantage de lisibilité et permet les analyses automatiques basées sur des conventions orthographiques standards (Thibault et Vincent, 1988). En revanche, toute une série d'éléments sont perdus avec l'orthographe standard comme la tonalité, les pauses, les chevauchements, etc. qui doivent être identifiés par des marques conventionnelles spéciales (comme les conventions de notations développées en AC).

Enfin, le troisième type de transcription, dite étroite ou *transcription phonétique*, reproduit les détails des interactions conversationnelles, comme l'accentuation et l'articulation de certains mots, la prononciation plus bruyante d'autres mots, la superposition des voix, etc²⁰. Avec cette transcription, le résultat est très fidèle à la production orale sans en donner une image péjorative puisque le code graphique est différent de nos habitudes de lecture. En revanche, ce type de transcription est chronophage et peut s'avérer difficile à lire pour les néophytes. L'alphabet phonétique peut être utilisé ponctuellement lorsque notamment un terme n'existe pas dans la langue standard ou que la présence d'un accent pose un problème d'intercompréhension. Le choix quant à l'orthographe va lui aussi découler des objectifs de la recherche et des formes présentes dans le corpus.

Qu'il utilise ou non un logiciel de transcription, le transcripteur devra choisir un type de représentation graphique. En revanche, en fonction de ce choix, certains outils seront

Tableau 5. La représentation visuelle des données

LOGICIEL	PRÉSENTATION
ANVIL	Partition
Casualtranscriber	Texte
CLAN	Texte
EasyTranscript	Texte
ELAN	Partition
EureScribe	Texte
EXMARALDA	Partition
Express Scribe	Texte
F4&F5 Transkript	Texte
FOLKER	Partition
HyperTranscribe	Texte
InqScribe	Texte
Listen N Write	Texte
otranscribe	Texte
PHON	Texte
PmTrans	Texte
PRAAT	Partition
Sonal	Texte
TRANSANA	Texte
Transcribe	Texte
Transcriber AG	Texte
Transcription Helper	Texte
Transcriber Pro	Texte
Transcriva	Texte

plus adaptés que d'autres. Il est clair que la transcription phonétique ou phonémique ne pourra se faire que sous un logiciel capable de gérer ce type de présentation et le degré de granularité et de multimodalité qu'il exige, comme ANVIL, CLAN, ELAN EXMARaLDA, FOLKER, PHON ou PRAAT. La transcription orthographique standardisée pourra également se faire soit sous ces logiciels afin de bénéficier des conventions langagières standardisées intégrés qu'ils proposent, soit sous des logiciels de transcription plus généralistes moyennant le recours à des conventions de notations en complément. Enfin, les logiciels de transcription généraliste sont en principe suffisants pour la transcription orthographique adaptée, au plus près de la singularité de l'enquête et de l'oralité.

Deuxièmement, une fois la représentation graphique choisie, le transcripteur devra choisir ensuite la *représentation visuelle des données*. Il en existe deux grands types disponibles, qui renvoient à ce qu'on appelle dans la terminologie des logiciels multimodaux la *présentation texte* et la *présentation partition*.

Le *type de présentation texte* est une présentation verticale des transcriptions comme un texte classique. L'axe vertical, de haut en bas, représente l'axe du temps ou la succession temporelle. Dans les logiciels multimodaux, l'espace d'écriture de chaque ligne est parfois utilisé comme moyen complémentaire de codage ou d'alignement temporel. Chaque ligne principale est en général suivie d'un ou plusieurs titres qui contiennent des informations dépendantes de la ligne principale. Un exemple classique de ce type de présentation est celui du logiciel CLAN, mais c'est aussi le cas de Transcriber AG ou de PHON.

```

*JULIEN: je {/} j'appuie sur vert ?
%mor: pro:subilic v:lappuyer prep:art/sur adj:vert ?
%pho: za apui syk vek
%add: à ABDEL.
%act: JULIEN appuie sur un bouton de la machine à café.
*ABDEL: oui yas+y.
%mor: adv:ynloui v:mdllex/vas+y.

```

Figure 4. Exemple de présentation de type texte (logiciel CLAN)²⁹

Les logiciels recourant à ce type de présentation, largement majoritaires (19 sur 24), proposent des outils de mise en page plus ou moins avancés. Dans oTranscribe, il n'y a presque aucune option de mise en page du texte, en dehors de l'italique et du gras. Toutefois, cet aspect est quelque peu secondaire puisqu'une fois écrit, le texte peut être exporté dans un logiciel de traitement de texte.

Comme le notent Parisse et Morgenstern, la présentation texte a ses forces et ses faiblesses :

« Le type de présentation texte est très bien adapté à l'analyse textuelle, en particulier les récits, mais aussi les successions d'interactions lorsque l'analyste s'intéresse à des suites d'échanges qui sont plus facilement visualisés ainsi. Le type de présentation texte permet également d'afficher un grand nombre de lignes dépendantes en même temps. Il présente par contre des limites importantes dans la représentation des dialogues et en particulier des chevauchements entre locuteurs. Il est également moins bien adapté à une situation d'interaction impliquant plus de deux ou trois personnes. » (Parisse et Morgenstern, 2010 : 10)

Si les interlocutions croisées et les interlocuteurs ne sont pas trop nombreux, la présentation texte fonctionne très bien. En revanche, dès que l'entrelacement des énoncés est fort comme dans une conversation à plusieurs, il devient nécessaire d'adapter ce type de présentation en recourant au mode de transcription utilisé par les chercheurs travaillant avec la méthode des *focus groups* (Duchesne et Haegel, 2004) (voir la figure 5).

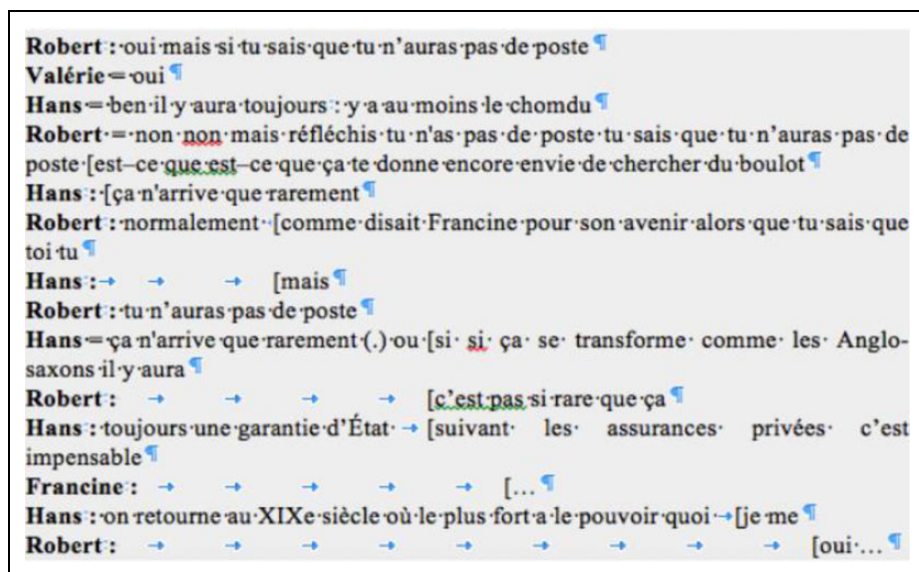


Figure 5. Présentation texte pour les *focus groups*

Si cette adaptation de la partition texte s'avère insuffisante dans un certain nombre de cas, le transcripteur peut recourir au second grand type de présentation : la *présentation partition*. Celle-ci est une présentation horizontale des transcriptions qui ressemble à une partition de musique. Le déroulement du temps est représenté par l'horizontal de gauche à droite. Selon les logiciels, cet axe horizontal est proportionnel au déroulement temporel ou à la taille du texte transcrit. Dans ce format, à chaque interlocuteur correspond une ligne de la partition. Des lignes supplémentaires permettent de coder les lignes dépendantes de chaque interlocuteur. Ce format est largement utilisé dans les logiciels servant à coder les interactions de langage (par exemple ELAN ou EXMARaLDA), mais c'est aussi la présentation utilisée dans un logiciel comme PRAAT (qui, bien que spécialisé dans le traitement du son, permet la transcription et l'alignement temporel).

Comme le texte, la présentation partition présente des avantages et des inconvénients :

« Les avantages sont une précision extrême dans les chevauchements d'énoncés et dans les alignements temporels en général, la possibilité de visualiser facilement les interactions de nombreux interlocuteurs et la liberté de maniement des annotations dépendantes par rapport aux annotations principales. Les longs énoncés ou les tours

de paroles et les annotations secondaires, si elles sont en grand nombre, peuvent cependant être assez peu lisibles. » (Parisse et Morgenstern, 2010 : 10)

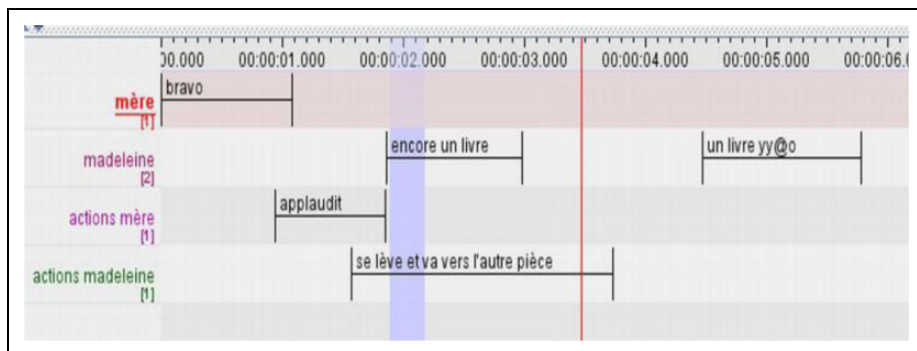


Figure 6. Exemple de présentation de type partition (logiciel ELAN)³⁰

6. L'articulation des supports (audio, texte, vidéo)

Le sixième paramètre renvoie aux possibilités d'articulation des différents supports (audio, texte, vidéo). Toute transcription consiste en une *textualisation* de l'oral : « One could consider a text as a *trace* of an interaction. But interactions do not leave traces by themselves except in the memory of those who took part in them or witnessed them. This action of creating such a trace can be called *entextualization*, a term which is possible a calque of the German term *Vertextung* » (Haberland, 2012 : 3). La textualisation peut être définie comme « the process by which circutable texts are produced by extracting discourse from its original context and reifying it as a bounded object » (Sung-Yul Park et Bucholtz, 2009 : 485). Le transcripteur produit un nouvel objet – le texte – dont le but est qu'il soit détaché (au moins partiellement) du contexte d'énonciation initial pour pouvoir être transportable, traitable et faciliter ainsi l'analyse. Mais le prix à payer de cette opération est le risque d'oublier que le texte n'est pas l'événement empirique lui-même qu'il s'agit d'étudier (un entretien ou bien une réunion militante en situation d'observation, etc.) mais une trace partielle. Dans cette perspective, les logiciels de transcription peuvent contribuer à garder les apports de la textualisation tout en gardant un lien avec les sources audio ou vidéo.

Parmi les logiciels présentés, 10 d'entre eux comme oTranscribe ou Transcribe rendent possible le minutage interactif. Ils permettent également d'insérer dans la transcription un *timestamp* (marqueur temps), c'est-à-dire le minutage correspondant à la bande audio. Il s'agit concrètement d'un lien qui permet, lorsque l'on clique dessus, de retrouver la partie audio/vidéo correspondante. Cette fonctionnalité peut non seulement faciliter la navigation dans le corpus pendant la transcription mais s'avérer très utile par la suite, notamment pour la restitution auprès des enquêtés ou la communication auprès des collègues. Toutefois, les outils de synchronisation audiotextuelle restent très limités par rapport aux logiciels multimodaux.

Tableau 6. L'articulation des supports

LOGICIEL	ARTICULATION SUPPORTS
ANVIL	Multimodalité
Casualtranscriber	Minutage interactif
CLAN	Multimodalité
EasyTranscript	Minutage interactif
ELAN	Multimodalité
EureScribe	Minutage interactif
EXMARALDA	Multimodalité
Express Scribe	Minutage interactif
F4&F5 Transkript	Minutage interactif
FOLKER	Multimodalité
HyperTranscribe	Minutage interactif
InqScribe	Minutage interactif
Listen N Write	Minutage interactif
otranscribe	Minutage interactif
PHON	Multimodalité
PmTrans	Minutage interactif minimal
PRAAT	Multimodalité
Sonal	Multimodalité
TRANSANA	Multimodalité
Transcribe	Minutage interactif
Transcriber AG	Multimodalité
Transcription Helper	Minutage interactif
Transcriber Pro	Minutage interactif
Transcriva	Minutage interactif

De tels logiciels permettent la transcription de dialogues oraux, avec *alignement* temporel du texte à la vidéo ou au son. CLAN, ELAN, EXMARaLDA, PHON, PRAAT, Transcriber AG ou Sonal permettent ainsi non seulement de transcrire au sens strict, c'est-à-dire de passer du son au texte mais d'articuler de manière beaucoup plus fine le texte, le son et, parfois, l'image animée (vidéo). L'analyse multimodale consiste à mettre en relation des informations langagières produites dans différentes modalités, chacune d'elles contribuant à l'élaboration et à la perception du message communiqué. Ainsi, l'on peut distinguer la modalité verbale qui comporte plusieurs niveaux – phonèmes, choix du lexique, organisation syntaxique, organisation discursive – de la modalité orale – prosodie, qualité de voix – et enfin de la modalité visuelle – gestualité et expressions faciales. Cette multimodalité les rend adaptés à l'analyse du langage oral et des interactions.

7. Les formats des données

Le septième critère de sélection concerne les formats des données. En premier lieu, les logiciels de transcription reconnaissent un certain nombre de *formats d'import audio et/ou vidéo* (voir tableau 7). Tous les logiciels ne sont pas capables d'importer tous les supports. Ainsi, sur les 24 logiciels recensés, seul PHON n'a pas de format de fichier

audio importable (il traite le son *via* les fichiers vidéos) et 4 d'entre eux (FOLKER, PM Trans, PRAAT et Transcriber AG) ne peuvent importer de fichiers vidéos, étant spécialisés dans la transcription vocale. Quelques remarques sont par ailleurs nécessaires. Tout d'abord, si certains logiciels payants comme Express Scribe mettent en avant une liste impressionnante de formats audio et vidéos pris en charge, l'importance de ce paramètre doit être nuancé. Il est toujours possible de convertir pour la transcription ces fichiers audio en *.wav ou *.mp3 qui sont les formats audio les plus courants et donc reconnus par tous ces logiciels, quitte à faire une copie en format pérenne (*.wav ou *.flac). Pour cela, le transcripteur dispose de logiciels de conversion des formats audio gratuits comme Any Audio Converter, FreeMake Audio Converter, Format Factory ou Fre:ac. Il existe des outils analogues pour la conversion de formats vidéos comme Format Factory, Freemake Video Converter, HandBrake, MediaCoder, VidCoder, etc.

Ensuite, dans le cas des logiciels payants, il n'est pas rare que la liste des formats importables soit beaucoup plus réduite dans la version gratuite que dans la version professionnelle, voire même, comme dans le cas d'Express Scribe, qu'il ne soit possible d'importer de la vidéo qu'avec la version payante. Par ailleurs, il convient de se méfier des informations fournies sur les sites des logiciels qui, pour mettre en avant leur produit, tendent parfois à occulter certaines choses concernant les formats d'import. À titre d'exemple, sur oTranscribe, certaines vidéos YouTube ne supportent pas la fonction « ralentir la vitesse de lecture ». Enfin, ce n'est pas parce qu'un logiciel est capable de lire un format qu'il peut l'utiliser pleinement. Ainsi, Sonal supporte les formats compressés de type *.mp3, *.wma, *.ogg, *.amr, *.m4a, mais il travaille préférentiellement avec des documents *.wav (PCM) encodés au format 16 bits mono dans la mesure il ne peut représenter visuellement la bande son que sous *.wav.

Troisièmement, seuls 7 logiciels référencés proposent l'import de fichiers texte (dans le tableur « Format import » → « Texte »). Cela s'explique par le fait que les 17 autres ont été pensés pour être la première interface pour passer de l'audio et/ou la vidéo au texte. Toutefois, ce paramètre a moins d'importance car il est toujours possible d'utiliser la fonction copier-coller depuis un fichier texte dans un logiciel. Quatrièmement, enfin, certains logiciels sont capables d'importer non seulement des données enregistrées mais également des données déjà structurées et alignées. Il s'agit logiquement des 9 logiciels multimodaux (ANVIL, CLAN, ELAN, EXMARaLDA, FOLKER, PHON, PRAAT, Transana et Transcriber AG) : d'une part, leur fonction même est de produire, d'annoter et d'analyser les corpus multimodaux et d'autre part, ils sont souvent utilisés de manière complémentaire par la communauté des spécialistes de linguistique et d'AC.

Les logiciels de transcription présentent également un certain nombre de **formats d'export** dans lequel il est possible d'enregistrer les données. Tout d'abord, sur les 24 logiciels, 23 d'entre eux offrent des possibilités d'édition. Seul ANVIL ne permet pas l'export des transcriptions dans des fichiers de traitement de texte. L'export de texte peut en revanche prendre des formes très variées :

- des fichiers de traitement de texte (*.rtf, *.txt, *.doc ; *.docx ; *.odt) ou assimilés (*.pdf, *.html5) ;
- des fichiers e-book (*.epub) ;

Tableau 7. Les formats natifs et d'import

FORMAT IMPORT					
LOGICIEL	FORMAT NATIF	Audio	Texte	Vidéo	Formats corpus alignés
ANVIL	*.anvil	*.wav	NON	*.avi ; *.mov	Praat (*.Textgrid, PitchTier)
Casualtranscriber	?	?	*.rtf ; *.txt ; *.doc ; *.csv	Vidéos You Tube	NON
CLAN	chat (*.cha) ; ca (*.ca)	*.wav ; *.mp3 ; *.aif ; *.aiff	NON	*.mpeg ; *.mpg ; *.dat ; *.mov	*.eaf>*.cha ; *.Text-Grid>*.cha ; exmaralda (*.xml)>*.cha ; *.tag>*.cha ; Shoebox/Toolbox (*.sf et *.txt)>*.cha
Easy Transcript	?	*.wav ; *.mp3 ; *.ogg ; *.wma ; *.m4a ; *.flac ; *.aiff	NON	*.avi ; *.mp4 ; *.mov ; *.3gp ; *.ogm ; *.mkv ; *.wmv ; *.m2ts ; *.mts	NON
ELAN	*.eaf	*.wav	NON	*.mpeg1-2 ; *.mov	Shoebox/Toolbox (*.sf et *.txt)>*.eaf ; *.cha>.eaf ; *.tag>*.eaf
EureScribe	?	*.wav ; *.wma ; *.mp3	*.srt ; *.txt ; *.rtf	*.avi ; *.mpeg ; dvd videos	NON
EXMARALDA	*.xml	*.wav	NON	*.mpg	TASX ; HIAT-DOS ; *.TextGrid

(continued)

Tableau 7. (continued)

FORMAT IMPORT				
LOGICIEL	FORMAT NATIF	Audio	Texte	Vidéo
Express Scribe	?	Pour la version gratuite : *.wav ; *.mp3 ; *.wma ; *.dct. Pour la version payante sous Windows : *.aac ; *.act ; *.aifc (aiff compressé) ; *.aiff ; *.amr ; *.ape ; *.au ; *.caf (format Apple Core Audio) ; *.dct (Express Dictate, MSRS, etc.) ; *.dsp TrueSpeech ; *.dss ; *.dvf (enregistreur Sony) ; *.dvs (enregistreur Royal) ; *.flac ; *.ogg ; *.m4a ; *.mp2 ; *.mp3 ; *.mpc ; *.msv (enregistreur Sony) ; *.qcp ; *.ra/rm (RealAudio) ; *.shn ; *.spx ; *.voc ; *.vox ; *.wav (dont codecs adpcm, aLaw, gsm 6.10, pcm et uLaw) ; *.wma ; *.mp3 ; *.wav ; *.ogg ; *.wma ; *.wav	NON	Pour la version Pro uniquement : *.3g2 ; *.3gp ; *.asf ; *.avi ; *.DivX ; *.dv ; *.flv ; *.m4v ; *.mkv ; *.mov ; *.mp4 ; *.mpeg ; *.wmv
F4&F5 Transkript FOLKER	*.rtf *.flk		*.rtf ; *.txt NON	*.mpg ; *.avi NON
HyperTranscribe	*.htd	*.3gp ; *.aif ou *.aiff ; *.mp3 ; *.mov ; *.wav ; *.u	NON	*.dv ; *.avi ; *.mpg ; *.mpeg ; *.mov
InqScribe	?	*.wav ; *.mp3 ; *.aiff ; *.wma ; *.aac	*.txt ; Tab-delimited text ; *.xml	H.263 ; H.264 ; QuickTime movies (*.mov ; *.qt) ; *.mpeg 1 ; *.mpeg 2 ; *.mpeg 4 ; *.avi ; *.wmv
Listen N Write	?	*.wav ; *.mp3 ; *.ogg ; *.wma ; *.avi	NON	*.mpg ; *.wmv ; *.ogv ; *.flv ; *.vob ; *.ts
otranscribe	*.otr	*.wav ; *.mp3 ; *.ogg ; *.webm	NON	*.mp4 ; *.ogg ; *.webm + liens vidéos Youtube

(continued)

Tableau 7. (continued)

FORMAT IMPORT					
LOGICIEL	FORMAT NATIF	Audio	Texte	Vidéo	Formats corpus alignés
PHON	*.TextGrid ; TalkbankXml	NON	NON	OUI	*.TextGrid ; TalkbankXml
PmTrans	?	*.wav ; *.mp3	NON	NON	NON
PRAAT	*.TextGrid	*.wav ; *.mp3 ; *.aif ; *.aifc ; *.au ; *.nist	NON	NON	NON
Sonal	Fichier corpus (*.crp) ; fichiers transcription (*.rtt) ; fichiers de lecture (*.lct) ; fichiers d'image (*.bmp)	*.wav ; *.mp3 ; *.ogg ; *.wma ; *.amr ; *.m4a. (mais la bande son ne peut être représentée visuellement dans le logiciel que sous *.wav)	*.doc ; *.docx ; *.txt	*.avi ; *.mpg	NON
TRANSANA	*.xml (pour base de données) ; *.rtf (pour transcriptions)	*.wav ; *.mp3 ; *.wmv ; *.mov	*.rtf	*.mpg ; *.mpeg -2 ; *.avi ; *.mov	*.rtf (pour transcriptions) ; *.txt (pour statistiques)
Transcribe	?	*.wav ; *.mp3 ; *.mp4 ; *.m4a ; *.amr ; *.wma	NON	*.mp4 ; vidéo Youtube	NON
Transcriber AG	*.tag	*.wav ; *.mp3 ; *.au ; *.aiff/*.*.aif ; *.sph directement importables et *.wma ; *.flac ; *.mp2 ; Vorbis (*.ogg) ; *.ra via le convertisseur Ffmpeg	*.txt, *.html ou *.pdf ; *.xml ; *.sgml ; *.utf ; *.stm ; *.typ ; *.ctm ; *.phn ; *.wrd ; *.lola ; *.lbl ; *.mdtm ; *.lab	NON	*.trs, with DTD trans-13.dtd or trans-14.dtd ; *.stm ; *.ctm ; *.cha ; *.mdtm ; *.sgml
Transcription Helper	?	*.wav ; *.mp3 ; *.wma	NON	*.avi ; *.mpeg ; *.mov ; *.wmv	NON
Transcriber Pro	?	*.mp3 ; *.wav ; *.wma ; *.m4a	NON	Vidéos You Tube *.wmv ; *.wmv ; *.avi ;	NON
Transcriba	?	*.mp3 ; *.wav ; ?	NON	*.divX ; *.flv ; *.mkv ; *.gvi ; *.vp6 ; *.yfw	NON

Tableau 8. Les formats d'export

LOGICIEL	FORMAT EXPORT	
	Édition	Corpus aligné
ANVIL	NON	Fichier texte ; *.xml
Casualtranscriber	*.txt ; *.csv	NON
CLAN	Impression du texte seul sans <i>time code</i> audio ou vidéo	chat>ca ; chat>elan ; chat>praat ; chat>xmar ; chat>praat
EasyTranscript	*.doc ; *.docx ; *.odt ; *.txt ; *.html ; *.xml ; *.epub	NON sauf les timesmarks compatibles avec F4&F5 Transkript
ELAN	Création d'un fichier excel avec les time codes et le texte	*.eaf>shoebox ; *.eaf>*.cha ; *.eaf>transcript text ; *.eaf>-quicktime text
EureScribe	*.rtf pour le texte ; *.srt pour les sous-titres vidéos	NON
EXMARALDA	*.html ; *.pdf ; *.rtf	*.TextGrid ; *.eaf ; *.xml (TASX Annotator) ; *.xml (Interlinear Text XML)
Express Scribe	Word (*.doc et *.docx). Fonctionne avec d'autres logiciels de traitement de texte mais les raccourcis clavier ne fonctionnent pas toujours avec autre chose que Word	NON
F4&F5 Transkript	*.rtf ou *.txt (pour les transcriptions) ; *.csv (pour les matrices) ; *.xml (pour les bases de données) ; *.srt (pour les sous-titres vidéos)	Oui (format en partie compatible avec ELAN et Transana via *.xml)
FOLKER	*.doc ; *.docx	*.xml
HyperTranscribe	*.txt ; *.rtf ; *.sbv (sous-titres vidéos YouTube et format de captures vidéo)	NON
InqScribe	*.txt ; Tab-delimited text ; *.xml ; *.html ; Final Cut Pro XML ; Spruce STL ; Subrip ; WebVTT	Oui via *.xml
Listen N Write	*.rtf ; *.txt	NON
otranscribe	markdown (*.md), texte brut (*.txt) ou oTranscribe format (*.otr) + Google Drive	NON
PHON	*.pdf ; MS Excel (*.xls) ; MS Word (*.doc ; *.docx) ; Texte Libre Office (*.odt) ; *.html ; *.csv	Clan (*.cha ; *.ca) ; Praat (*.TextGrid)
PmTrans	*.txt	NON
PRAAT	*.pdf	NON
Sonal	Pour les transcriptions : *.doc ; *.docx ; *.txt ; *.rtf ; *.html5 ; Pour l'audio : *.wav ; *.mp3 ; Pour les graphiques : *.bmp ; Pour la base de données : *.xls ; *.bdf (Trideux) ; Pour les vidéos : *.srt (à venir)	NON

(continued)

Tableau 8. (continued)

LOGICIEL	FORMAT EXPORT	
	Édition	Corpus aligné
TRANSANA	*.pdf ; *.rtf	*.xml
Transcribe	*.doc ; *.docx	NON
Transcriber AG	*.html ; *.txt	*.stm ; *.cha ; user defined format
Transcription Helper	*.doc ; *.docx ; Autres *.odt ; *.txt ?	NON
Transcriber Pro	*.txt ; *.html	NON
Transcriva	*.rtf ; *.doc ; *.docx ; *.txt	NON

- des fichiers en format texte plus structuré comme le *.xml. Cette possibilité d'export est très intéressante en ce qu'elle permet l'analyse des tables produites sous des logiciels statistiques comme SPSS ou Statistica (c'est le cas pour ANVIL) ;
- des fichiers tableurs (*.csv, *.xls) avec juste le texte, les *timecodes* et le texte (dans le cas d'ELAN) ou encore pour exporter la base de données qualitatives constituées sous Sonal.

Ensuite, certains logiciels proposent l'export de fichiers audio ou vidéo dans la mesure où ils offrent des fonctionnalités permettant de les modifier : des formats pour les sous-titres vidéos (principalement en format *.srt ou *.sbv) avec InqScribe, Sonal ou HyperTranscribe ou encore des formats audio lorsque les logiciels proposent un traitement du signal comme c'est le cas de PRAAT. Enfin, comme pour l'import, certains logiciels offrent des possibilités d'export des données non plus séparément, par type de support (texte, audio, vidéo), mais sous forme de corpus multimodal de données alignées. C'est là encore le cas des 9 logiciels multimodaux, mais également de certains logiciels de transcription généraliste. À titre d'exemple, F4 & F5 Transkript peut ainsi exporter en format *.xml des données compatibles avec ELAN et Transana.

La question des formats dépasse de très loin la simple dimension technique tant elle est lourde d'enjeux. Le premier d'entre eux est la sécurité et la protection des données personnelles contenues dans les fichiers. oTranscribe propose ainsi l'export de ses données transcrites dans Google Docs ou Google Drive alors même que les conditions générales d'utilisation que l'utilisateur doit accepter pour pouvoir utiliser les services Google spécifie que les données pourront être dupliquées et exploitées par l'entreprise californienne. Le second enjeu concerne l'**interopérabilité** des outils. Les logiciels de transcription ont une panoplie plus ou moins grande de fonctions qui leur offrent une certaine adaptabilité et n'oblige pas à systématiquement changer de logiciel lorsque l'on change de mode de travail. Cette souplesse ne devrait toutefois pas empêcher l'utilisateur averti de changer de logiciel lorsque cela peut lui apporter des solutions techniques efficaces. On peut profiter des capacités de conversion entre logiciels pour passer d'un logiciel à l'autre et utiliser intelligemment le logiciel le plus approprié en fonction

de ses besoins du moment. Dans cette perspective, l'interopérabilité des formats constitue un critère important pour faciliter la migration ou l'usage parallèle de différents logiciels. Cela ne concerne bien évidemment pas que le passage d'un logiciel de transcription à un autre mais aussi et surtout d'un logiciel de transcription à un logiciel d'analyse comme un CAQDAS, un logiciel d'analyse lexicométrique ou textométrique ou encore un logiciel de statistiques. C'est pourquoi je suggère fortement d'opter pour des formats standards ouverts et non propriétaires.

Si le choix des formats est déterminant, il existe néanmoins, comme mentionné plus haut, des outils logiciels permettant de convertir les formats : logiciels de conversion de l'audio et/ou de la vidéo, logiciels de traitement de texte permettant la conversion de formats texte, etc. Toutefois, ces outils permettent simplement de convertir un format en un autre. Ils sont donc utiles en amont de tout travail de transcription. Or, le transcripteur peut tout à fait commencer le travail sur un logiciel et vouloir en changer ou bien vouloir en utiliser un autre, complémentaire, en parallèle. Dans cette perspective, il existe également, pour ce qui concerne les formats d'entrée, des outils plus intéressants permettant de passer d'un format à l'autre en conservant toutes les informations supplémentaires ajoutées au matériau lui-même (alignement, annotations, *tags*, mise en page, etc.)²¹. Ces outils sont toutefois pensés pour les logiciels multimodaux car la sauvegarde des informations ajoutées à la transcription proprement dite leur est la plus essentielle.

Un dernier point central relatif aux formats est la question de leur **pérennité**. Je ne prendrai que quelques exemples. Il n'est pas garanti qu'un entretien enregistré en format *.wma sera toujours lisible d'ici 5 à 10 ans. C'est également le cas des formats de fichiers compressés parmi lesquels le *.mp3 apparaît fortement déconseillé en termes de longévité. Enfin, les logiciels propriétaires de traitement de texte comme Microsoft Word n'assurent en aucun cas la lisibilité d'un fichier ancien par la nouvelle version. Fort heureusement, on assiste progressivement à une prise de conscience collective de ce problème massif pour les SHS. Dans ce cadre, le transcripteur pourra se reporter au site du CINES (Centre informatique national de l'enseignement supérieur) qui propose une interface en ligne (baptisée Facile : <https://facile.cines.fr/>) sur laquelle on peut trouver une liste des formats conseillés et vérifier la pérennité de ses fichiers. Dans cette perspective, si l'idéal consiste à importer et exporter les données dans les logiciels de transcription dans des formats déjà pérennes, il est tout à fait possible de distinguer le format des fichiers pour la transcription et leur format à des fins de conservation.

8. Outils de lecture et de traitement du son

Les outils spécifiques qu'ils proposent pour faciliter le travail de transcription proprement dit constitue un autre critère important de sélection des logiciels : lecteur audio/vidéo, éditeur de texte, traitement du signal sonore, codage des locuteurs, développeur de texte, reconnaissance vocale, etc. En corollaire, savoir si ces fonctionnalités sont implémentées dans le programme ou, à défaut, si le logiciel est compatible avec les outils d'aide à la transcription qui existent doit être pris en considération.

En premier lieu, le fait de proposer une interface visuelle unique apparaît comme l'un des principaux intérêts de ces logiciels : les bascules entre le lecteur audio et le traitement de texte disparaissent, ce qui représente un gain de temps et un confort de travail non

négligeables. On peut ainsi distinguer les *systèmes intégrés* (ce qui est le cas de l'immense majorité des logiciels présentés ici) des logiciels qui ont besoin d'un lecteur audio ou d'un éditeur de texte supplémentaire, à l'instar d'Express Scribe dont c'est sans doute le principal défaut.

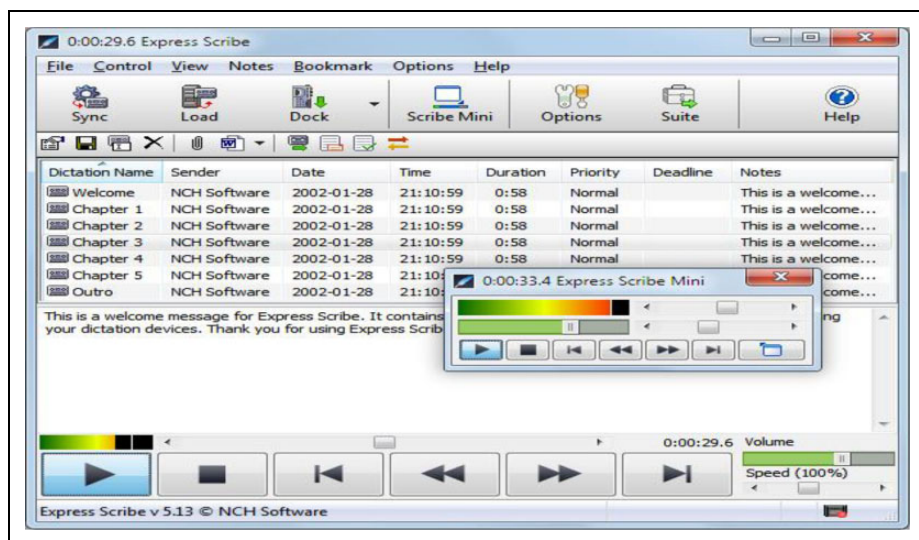


Figure 7. Le logiciel Express Scribe (Source : <http://www.nch.com.au/scribe/fr/screenshots.html>)

Intégration ne signifie néanmoins pas que le logiciel « contient » tous ces outils ; mais il communique en revanche avec d'autres logiciels, tout en permettant au transcripteur de rester sur la même interface. Les logiciels présentés ici n'ont ainsi pas leur propre lecteur audio et/ou vidéo. Certains requièrent l'installation d'un certain type de *player* audio comme QuickTime. D'autres comme ELAN utilisent le lecteur du fichier installé sur l'ordinateur et/ou le navigateur, comme par exemple Windows Media Player, QuickTime ou JMF (Java Media Framework) pour l'audio. Par conséquent, une large variété de formats audio et vidéo est supportée par ELAN. Le transcripteur doit donc veiller à ce que le lecteur audio/vidéo compatible avec le logiciel de transcription soit installé sur son ordinateur.

Le second ensemble de fonctionnalités concerne le *paramétrage de la vitesse de lecture audio*. Un grand nombre de ces logiciels, même les plus simples comme oTranscribe, permettent de ralentir la vitesse de lecture audio pour l'adapter à ses besoins et à sa vitesse de frappe ou à la qualité de l'enregistrement audio traité. Certains logiciels comme F4&F5 Transkript permettent à l'inverse d'accélérer la vitesse de lecture pendant que le transcripteur relit ce qu'il a tapé, ce qui offre un gain de temps. Le programme rembobine automatiquement un peu chaque fois que le transcripteur fait une pause. Cela lui permet d'entendre à nouveau les deux ou trois derniers mots avant de continuer et contribue à la fluidité du travail. De même, sous Sonal, le mode dictée permet de retranscrire des portions de bande successives de quelques secondes (représentées par



Figure 8. Le mode dictée dans Sonal (Source : <http://www.sonal-info.com/sites/all/exec/Mode%20d'emploi%20Sonal.pdf> (p. 96))

un rectangle blanc sur la bande). Le lecteur progresse jusqu'à la fin de la portion courante, puis revient au début et s'arrête, attendant le signal du transcripateur. Cela lui laisse le temps de transcrire tranquillement avant de passer à la suite.

Toutefois, le changement de vitesse entraîne une modification de la tonalité du fichier audio. Une musique ralentie aura par exemple une sonorité plus grave. Aussi certains logiciels plus sophistiqués comme Express Scribe, F4&F5Transkript ou InqScribe garantissent une vitesse de lecture variable mais à tonalité constante.

La troisième fonctionnalité intéressant directement le travail de transcription est **l'identification de tours de parole avec fonction de codage automatique des locuteurs** (comme F4&F5 Transkript, Transcriber AG, Sonal ou Transcriva). Ils identifient en outre les énonciateurs dans différentes couleurs, ce qui permet d'identifier les pauses et les changements d'énonciateurs ainsi que le moment du temps de parole.

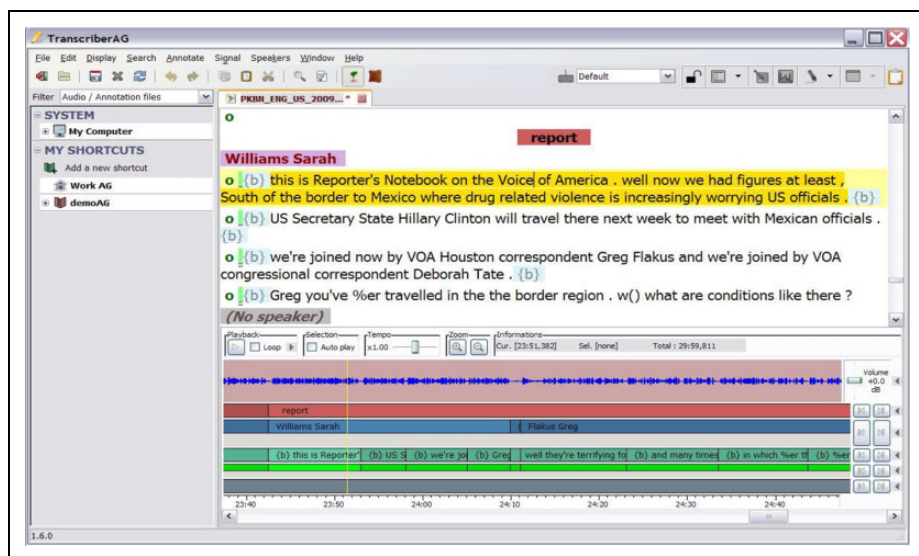


Figure 9. Le codage des locuteurs dans Transcriber AG. (Source: http://transag.sourceforge.net/images/screenshot_1.jpg)

Certains des logiciels de transcription offrent également des fonctionnalités de *traitement du son*. À titre d'exemple, Express Scribe comprend trois outils de traitement spécial de l'audio pour aider à clarifier certains enregistrements audio s'ils sont de mauvaise qualité : « réduction du bruit d'arrière-plan », « amélioration supplémentaire du volume » et « Filtre passe-haut ». De même, lorsque l'audio est enregistré en stéréo, il est possible de réduire le volume de chaque canal pour augmenter la clarté du signal à transcrire. De telles fonctionnalités sont offertes par n'importe quel logiciel de traitement audio comme Audacity. La force d'EXpress Scribe vient donc moins de l'originalité des fonctionnalités qu'il propose que du fait qu'il réunit sous un seul programme des outils auxquels on peut accéder gratuitement mais séparément. PRAAT, d'abord conçu pour

traiter et afficher du signal de parole offre des fonctionnalités beaucoup plus avancées d'automatisation et de modification des traitements sur ce signal. Il requiert cependant des compétences techniques poussées.

Cinquièmement, certains de ces logiciels ont une fonction *développeur de texte*. Il s'agit d'un programme de saisie semi-automatique qui apprend les phrases utilisées le plus souvent et les suggère pour permettre de sélectionner et d'insérer le texte « à la volée ». Le transcripteur peut ainsi sélectionner et stocker les segments de texte les plus importants dont il a régulièrement besoin : phrases récurrentes, phrases longues et difficiles propres à un langage spécialisé (médical ou juridique par exemple), caractères transcriptionnels spéciaux²², noms des personnes impliquées, etc. Il peut par la suite à l'aide de raccourcis clavier insérer ces segments textuels et gagner ainsi du temps. Certains logiciels proposent cette fonctionnalité intégrée (comme F4&F5 Transkript ou Easy Transcript). D'autres sont compatibles avec des développeurs de texte (comme Express Scribe version payante avec FastFox).

Sixièmement, sept des logiciels présentés ici sont compatibles avec des *pédaliers USB professionnels*, qui permettent à l'utilisateur de contrôler la lecture des enregistrements avec les pieds. La plupart des logiciels proposent toute une série d'options de lecture permettant de jouer un rôle analogue, à l'instar de la fonction « SoundWalker » de CLAN qui reproduit l'utilisation d'une pédale de contrôle du défilement du son. Mais il est vrai que déléguer une partie des commandes des mains aux pieds (en général, au minimum trois : le retour rapide, la lecture/pause et l'avance rapide) apporte en soi un confort (en limitant les douleurs aux doigts et poignets, prévenant ainsi d'éventuels troubles musculo-squelettiques) et un gain de temps. Plus besoin d'arrêter sa frappe pour lancer, couper ou reculer sur la bande audio. Pendant que le transcripteur finit de taper la fin de sa phrase, son pied repositionne la bande au bon endroit et il peut valider sans faire de pause dans sa transcription. Attention à bien veiller à ce que le pédalier que l'on a ou que l'on souhaite acheter soit compatible avec le logiciel !

Enfin, parmi les logiciels présentés ici, seul Express Scribe est compatible avec un logiciel de *reconnaissance vocale* (Dragon Naturally Speaking) permettant de convertir (semi-)automatiquement la parole en texte. Toutefois, là encore, cet argument doit être nuancé car il suffit au transcripteur, s'il le souhaite, d'utiliser d'abord un logiciel de reconnaissance vocale pour opérer une transcription semi-automatique, puis d'importer sous son logiciel de transcription le texte (soit en important le fichier texte, soit par un copier-coller) et l'audio pour opérer un travail complémentaire avec la transcription manuelle.

Les logiciels de transcription proposent d'autres fonctionnalités intéressantes mais qui, pour des raisons de place, ne peuvent être traitées ici. Je me contenterai donc d'en citer quelques-unes d'entre elles comme le vérificateur/correcteur orthographique, l'évaluation quantitative du travail de transcription²³ ou encore l'édition de corpus textuel²⁴.

9. Les fonctions d'analyse

La transcription fait partie prenante de l'analyse, nous l'avons vu. Toutefois, c'est d'autant plus vrai que certains logiciels de transcription comprennent également des outils d'analyse du corpus embarqués : codage interprétatif, annotation, recherche de mots, calcul de fréquence, analyse lexicométrique, analyse interactionnelle, analyse

Tableau 9. L'intégration des outils

INTÉGRATION OUTILS					
LOGICIEL	Player	Éditeur texte	Pédaliers	Reconnaissance vocale	Développeur texte
ANVIL	OUI : Quicktime	OUI	NON	NON	NON
Casualtranscriber	OUI : QuickTime 7	OUI	?	NON	NON
CLAN	OUI : Quicktime)	OUI	NON	NON	NON
EasyTranscript	OUI (VLC Player 2)	OUI	OUI	NON	NON
ELAN	OUI : Windows Media Player ; Quick- Time ; JMF (Java Media Framework)	OUI	NON	NON	NON
EureScribe	OUI	OUI	NON	NON	NON
EXMARALDA	OUI	OUI	NON	NON	NON
Express Scribe	OUI	NON	OUI	OUI pour la version payante : Dragon Naturally Speaking. Pour la version gratuite, seul est proposé lors du téléchargement un logiciel de dictée (Express Dictate)	OUI pour la version payante : Compatible avec FastFox
F4&F5 Transcript	Oui : Windows Media Player 9 (pour Windows) ; VLC Player (pour Ubuntu) ; Direct X9.0c ou + (pour Mac)	OUI	OUI : Scythe USB 1,2,3 ; Olympus RS-2 ; Olympus RS-3 I ; Philips LFH 2310 ; Philips LFH 2330 ; Grundig 540	NON	OUI
FOLKER	OUI : Direct Show (pour Windows uniquement) ; QTJ (Quicktime for Java) ; JMF (Java Media Framework)	OUI	NON	NON	NON
HyperTranscribe	OUI : Quicktime	OUI	NON	NON	NON

(continued)

Tableau 9. (continued)

INTÉGRATION OUTILS					
LOGICIEL	Player	Éditeur texte	Pédaliers	Reconnaissance vocale	Développeur texte
InqScribe	OUI : Windows Media Player 11 et QuickTime 7	OUI	OUI	NON	NON
Listen N Write otranscribe	OUI OUI : lecteur audio dans le navigateur web	OUI OUI	NON NON	NON NON	NON NON
PHON	OUI	OUI	NON	NON	NON
PmTrans	OUI	OUI	NON	NON	NON
PRAAT	OUI	OUI	NON	NON	NON
Sonal	OUI	OUI	NON	NON	NON
TRANSANA	OUI : Windows Media Player ; Quicktime	OUI	NON	NON	NON
Transcribe	OUI	OUI	OUI (The Infinity USB Foot Pedal ; Olympus RS27 ; Olympus RS26)	NON	NON
Transcriber AG	OUI	OUI	NON	NON	NON
Transcription Helper	OUI	OUI	OUI	NON	NON
Transcriber Pro	OUI	OUI	NON	NON	NON
Transcriba	OUI : Quicktime	OUI	OUI : Pedalpax Pedis 3-Button Programmable USB Foot Switch ; AudioTranscrip- tion Foot Switch ; vPedal	NON	NON

morphosyntaxique, etc. La présentation détaillée de ces fonctions mériterait à elle seule un article spécifique. Aussi je me contenterai ici d'en énumérer simplement quelques-unes des fonctions principales.

Le premier type de fonctionnalités, proposé par la plupart des logiciels de transcription sous des formes plus ou moins avancées, relève du **codage et de l'annotation**. Dans le passage de l'oral à l'écrit, on perd nécessairement une partie des informations et des indices signifiants. On peut néanmoins limiter les pertes, en ne se contentant pas de transcrire de manière brute les données audio ou vidéo mais en les structurant et en les contextualisant. Il peut s'agir de fonctions simples comme le *codage des énonciateurs* : la plupart d'entre eux proposent ainsi un système d'identification de tours de parole avec fonction d'étiquetage de locuteurs (voir *infra*). Il est également possible dans certains cas de contextualiser quelque peu les conditions de l'entretien. À titre d'exemple, dans Easy Transcript, le bouton « information » permet d'ajouter des faits importants, tels que le lieu ou l'heure de l'entretien dans le fichier de transcription. Une autre fonctionnalité intéressante est la possibilité de marquer certains passages dans un fichier audio ou vidéo en y insérant des *signets* pour y faire référence ultérieurement ou pour y accéder rapidement, comme le propose par exemple Express Scribe. Par ailleurs, une partie des logiciels permet de coder les éléments contextuels contribuant à restituer le sens des énoncés transcrits comme par exemple « [rires] », « [le téléphone sonne] », « [une quinte de toux le prend] » ou « [il s'allume une cigarette] » et qui forment ce qu'on appelle le *paratexte*²⁵. Ils permettent ainsi de coder au fur et à mesure les indices propres à l'oralité, c'est-à-dire les indices propres à la voix des énonciateurs (changements de rythme, intonation, etc.) mais aussi tous les indices sonores sur le contexte dans lequel s'est déroulé l'entretien :

[...] viennent s'ajouter les indications propres à l'oralité que sont les silences, les soupirs et les respirations, les rires, les onomatopées, les sifflements, les claquements des lèvres, les changements de tonalité, les bafouillements, les hésitations, les émotions qui percent sous un ton brusquement assourdi ou au contraire strident, les bruits parasites (crayon qui tape nerveusement, pied qui frappe le sol, plat de la main sur la table), indications qui suggèrent déjà une première interprétation : ce silence correspond-il à une hésitation face à la question, à un temps de réflexion, à une réprobation, à une méfiance, ou à une incompréhension ? (Descamps, 2001)²⁶

L'annotation et le codage sont évidemment possibles avec les logiciels multimodaux comme par exemple Transcriber AG. PHON permet même de lier directement les annotations (par exemple la transcription phonétique) à des parties de la ligne principale (par exemple les mots puis les phonèmes). Mais elle l'est également avec certains logiciels de transcriptions généralistes qui s'inspirent sur ce point des CAQDAS. C'est le cas de F4&F5 Transkript qui propose une fonction « mémo », dont le format permet l'import par Atlas.ti, NVivo ou MaxQDA. Sonal, quant à lui, permet d'insérer des commentaires dans le corps du texte de la transcription mais qui pourront, grâce à un système de balises, être ou non affichés au moment de la lecture, et qui seront exclus de toutes les analyses (lexicométrie, dynamique des échanges, etc.). Mais sa principale originalité est de permettre par ailleurs le codage thématique directement de la bande

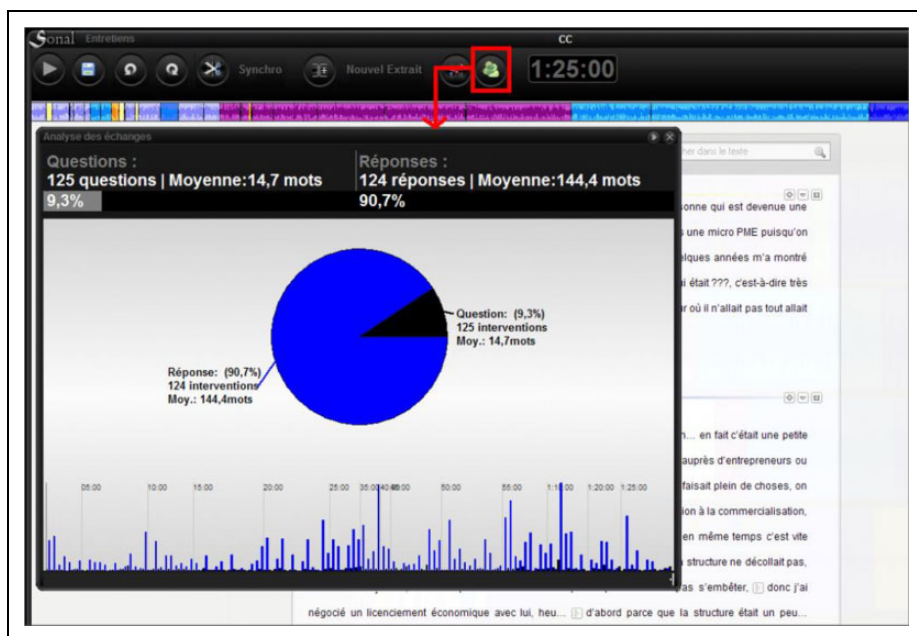


Figure 10. L'analyse conversationnelle quantitative dans Sonal (Source: <http://sonal.hypotheses.org/category/methodes-danalyse/analyse-conversationnelle>)

audio. Sonal offre également d'autres fonctionnalités originales telles que le système de « pondération qualitative » qui permet d'augmenter la taille du texte des passages les plus intéressants dans la fenêtre de lecture.

Le deuxième type de fonctionnalités offertes par certains de ces logiciels consiste à faciliter la **navigation et l'exploration documentaire**. En dehors du codage proprement dit, on peut ainsi citer la fonction « recherche de mots » qui, pour être basique, n'en constitue pas moins un outil intéressant. D'autres fonctionnalités plus avancées d'analyse statistique sont également proposées par quelques-uns des logiciels. Sonal par exemple offre deux options : une analyse lexicométrique et une analyse chronométrique. La première permet, par le décompte des mots, d'étudier systématiquement le vocabulaire contenu dans une série d'extraits définis par leur thématique et/ou par les caractéristiques des locuteurs. La seconde, elle, permet de mesurer le temps passé par l'enquêté à parler de telle thématique. Sonal propose en outre une fonctionnalité d'analyse conversationnelle quantitative limitée mais intéressante en permettant d'opérer des dénombrements sur les prises de parole des différents locuteurs dans les entretiens.

Le troisième type de fonctionnalités offertes cette fois uniquement par les logiciels multimodaux spécialisés en linguistique interactionnelle ou en analyse conversationnelle est l'**analyse morphosyntaxique**. Ils offrent ainsi une structure de codes standardisés permettant de caractériser les écarts à la norme syntaxique. Il existe ainsi de nombreuses conventions parmi lesquelles on peut citer CHILDES (<https://chilides.talkbank.org/>), IPA (<http://www.internationalphoneticalphabet.org/ipa-charts/ipa-symbols-chart-complete/>)

ou encore VALIBEL (<https://uclouvain.be/fr/instituts-recherche/ilc/valibel/outils-tools.html>). CLAN, ELAN ou PHON proposent des fonctionnalités d'analyse du corpus très poussées. Notons cependant qu'ils utilisent de fait des langages de programmation qui ne sont que rarement connus des chercheurs/euses en SHS. Le transcripteur se trouve ainsi devant un choix à faire, la richesse des fonctionnalités d'analyse étant difficilement compatible avec une exigence d'ergonomie *easy-to-use*. Se dégage ainsi un critère central : souhaite-t-on utiliser une interface « WYSIWYG »²⁷ ou bien un logiciel aux fonctionnalités d'analyse très poussées ? Si certains logiciels comme Transcriber AG ou Sonal proposent un entre-deux intéressant, il n'en reste pas moins que le transcripteur doit aussi sélectionner le logiciel dont il a besoin en fonction des compétences qui sont les siennes ou alors choisir de se former à des langages de programmation et/ou des logiques informatiques encore peu enseignés dans les cursus de SHS.

Conclusion

Au terme de cette présentation, on constate qu'il n'existe pas d'outil miracle pour la simple et bonne raison que le choix de tel ou tel logiciel dépend de l'ancrage théorique, disciplinaire et méthodologique du chercheur, de ses habitudes, de ses compétences techniques ou encore de la spécificité du terrain. Dans cette perspective, cet article pourra je l'espère aider les chercheurs/euses à mieux se repérer dans l'offre logicielle existante et sélectionner le dispositif adapté. Il se veut ainsi une contribution au tournant numérique des sciences sociales en facilitant l'appropriation des outils par les collègues. Entre le manuel de méthodes généraliste et le tutoriel d'utilisation de tel ou tel outil, il manque encore une ressource de type intermédiaire. Cet article constitue une telle ressource, en cherchant à articuler de manière plus resserrée méthodologie et technique. Par ailleurs, ce travail est d'abord conçu comme une base qui pourra être améliorée et actualisée par d'autres membres de la communauté universitaire (ou d'autres personnes). Alors que les logiques d'évaluation individuelle et la pression à l'« excellence » amènent parfois à omettre ou minimiser les lacunes dans nos publications, il s'est agi ici au contraire de publiciser aussi les informations manquantes sur les outils référencés tant par probité qu'afin de permettre à d'autres de compléter ce travail. Ce n'est donc pas un discours définitif mais plutôt une invitation à la contribution collective. Dans la même veine, ce travail ne prétend en aucune façon à l'exhaustivité²⁸. On pourrait ainsi mentionner un autre critère de sélection : les apports de ces différents logiciels de transcription pour le travail collectif. Aucun outil n'offre en l'état la possibilité d'un véritable travail collaboratif synchronisé. Cependant, plusieurs des critères énoncés comme la plateforme et plus encore les formats sont autant d'enjeux pour permettre la transcription à plusieurs ou l'export des données dans des logiciels d'aide à l'analyse comme par exemple les CAQDAS. Tout.e collègue souhaitant étendre et compléter ce panorama en les y intégrant serait évidemment bienvenu.e. La rubrique « Outils et instruments » nouvellement créée dans la nouvelle formule du BMS pourrait ainsi contribuer à mettre en relation des collègues intéressé.e.s par ces questions et à la cumulativité des compétences et des savoirs.

Notes

1. Le terme *transcription* sera privilégié à celui, plus courant en français, de retranscription. Le « re- » semble en effet abusif car il est question ici du passage de l'oral à l'écrit donc bien d'une *trans*-scription. Conséquemment, le terme *retranscription* sera employé dans son sens strict : celui de seconde transcription d'un document déjà transcrit au moins partiellement, comme c'est parfois le cas dans les analyses secondaires ou dans l'édition numérique de données textuelles.
2. De même, certains développeurs de logiciels (comme par exemple F4&F5 Transkript (<https://www.audiotranskription.de/english/downloads-en.html#book>) proposent des manuels de transcription généralistes disponibles en ligne mais qui ne mentionnent pour l'essentiel que les outils développés par l'entreprise et/ou compatibles avec ceux-ci.
3. Une modalité est une forme concrète particulière d'un mode de communication. Par exemple, le bruit, la musique, la parole sont des modalités du mode sonore. En informatique, un système est dit multimodal lorsqu'il est capable d'intégrer plusieurs modes de communication (son, audio, texte). Cependant, on désigne également par ce nom tout système capable d'intégrer plusieurs modalités de communication (même s'il n'intègre qu'un seul mode).
4. Pratique certainement la plus répandue, le couplage d'un logiciel de traitement de texte (comme Word ou Libre Office) et d'un logiciel de lecture audio et/ou vidéo reste néanmoins assez peu pensé en tant que tel afin d'être optimisé.
5. Certains logiciels de transcription que nous aborderons sont néanmoins couplés avec un logiciel de reconnaissance vocale, à l'instar d'*Express Scribe* avec *Dragon Naturally Speaking*.
6. Comme par exemple le *plug-in* Scripto (<http://scripto.org/>) que propose la plateforme Omeka, (<https://omeka.org/>), le wiki de transcription du projet Bentham (<http://blogs.ucl.ac.uk/transcribe-bentham/>) ou encore le logiciel Transkribus (<https://transkribus.eu/Transkribus/>), tous trois très intéressants pour la transcription collaborative de textes manuscrits.
7. L'océrisation désigne la transformation automatisée d'un fichier contenant l'image d'un document en un fichier texte grâce à un logiciel de reconnaissance optique de caractère (en anglais OCR pour *Optical Character Recognition*).
8. https://fr.wikipedia.org/wiki/Text_Encoding_Initiative
9. À titre d'exemple, je n'ai pas référencé *Transcriber*, qui fut pourtant l'un des premiers logiciels de transcription *open-source* performants. Bien que toujours utilisable, il a en effet été « remplacé » par *Transcriber AG*.
10. Le format natif est le format utilisé en interne par un programme. On le visualise par l'extension dans l'intitulé du fichier (comme par exemple *.docx pour Microsoft Word ou *.tag pour Transcriber AG).
11. Les informations manquantes sont signalées dans les tableaux par un point d'interrogation.
12. Les pages web indiquées dans ce papier ont été vérifiées pour la dernière fois le 15 janvier 2018.
13. Je tiens en particulier à remercier Séverine Gedzelman, Justine Lascar et Vincent Ventresque pour leur aide et m'avoir permis de découvrir certains outils que je ne connaissais pas.
14. À ce titre, voir les pages web des sites *Social Science Software* (<http://www.sosciso.de/fr/software/datenumwandlung/transcription/?wcteid=27>), *Text Analysis Info* (<http://textanalysis.info/pages/text-analysis-software—classified/transcribing-software.php>), *Speech analysis*

- and transcription tools* (http://liceu.uab.es/~joaquim/phonetics/fon_anal_acus/herram_anal_acus.html) ou encore, celles très utiles du laboratoire ICAR réservés aux logiciels de transcription et/ou d'annotation de corpus oraux (http://icar.univ-lyon2.fr/projets/corinte/documents/CORINTE_Tableau_ComparatifPA.pdf).
15. La notion de *granularité*, très utilisée dans le traitement et la gestion des données, définit la taille du plus petit élément, de la plus grande finesse d'un système. Quand on arrive au niveau de granularité d'un système, on ne peut plus découper l'information.
 16. Pour une présentation synthétique de ces notions, voir Maingueneau, 1996.
 17. O.S. : *Operating System*, en français « système d'exploitation », est un « ensemble de programmes qui dirige l'utilisation des ressources d'un ordinateur par des logiciels applicatifs » (Stuart, 2008 : 2). Parmi les plus connus, on peut citer Windows, Apple ou Linux.
 18. La licence GPL garantit ainsi que, quelles que soient les options prises au cours du développement du projet, une version stable et efficiente du logiciel, entièrement libre, est toujours disponible. Cela signifie que la dernière version du logiciel pourra toujours lire vos données enregistrées sous une version précédente, ce qui n'est pas toujours le cas des logiciels propriétaires. Le code source étant ouvert, même s'il n'est plus développé par son concepteur original, il peut être repris et amélioré par d'autres, ce qui, là encore, lui assure une pérennité plus grande.
 19. Pour un argumentaire plus détaillé en faveur de l'*open source* concernant la protection des données personnelles, voir Stallman, 2009.
 20. Si de tels détails ont moins d'importance, par exemple si l'analyste s'intéresse plus à la structure globale de la conversation, alors il est possible d'utiliser une variante de ce type de transcription, connue sous le nom de *transcription large* ou *transcription phonémique*.
 21. On peut ainsi citer *The Transformer* (<http://www.oliverehmer.de/transformer/>), développé par Oliver Ehmer et *Transcriber.js* (<http://modyco.inist.fr/transcriberjs/teiconvertbeta/>), développé par Christophe Parisse.
 22. À titre d'exemple, certaines des abréviations de transcription standard sont déjà installées et proposées sous F4&F5 Transkript.
 23. Sonal permet de calculer la vitesse de frappe en mots par minute ou encore le nombre d'heures de travail nécessaire par heure de bande à transcrire. Cela peut s'avérer utile pour évaluer en continu l'avancée du travail et le temps nécessaire pour l'achever.
 24. EXMARaLDA, conçu pour traiter des données textuelles de volume important, dispose ainsi de logiciels annexes pour traiter et éditer des corpus entiers.
 25. Dans *Seuils*, Gérard Genette désigne par le terme « paratexte » ce qui entoure et prolonge le texte (Genette, 1987). G. Genette distingue deux sortes de paratexte regroupant des discours et des pratiques hétéroclites émanant de l'auteur (paratexte auctorial) ou de l'éditeur (paratexte éditorial). Il s'agit du paratexte situé à l'intérieur du livre – le **péritexte** – (le titre, les sous-titres, les intertitres, les nom de l'auteur et de l'éditeur, la date d'édition, la préface, les notes, les illustrations, la table des matières, la postface, la quatrième de couverture...) et celui situé à l'extérieur du livre – l'**épitexte** – (entretiens et interviews donnés par l'auteur avant, après ou pendant la publication de l'œuvre, sa correspondance, ses journaux intimes...). Dans cette perspective, le paratexte dont il est question ici englobe le péritexte (l'« intérieur » du texte renvoyant à ce qui a été effectivement énoncé lors de l'entretien) et l'épitéxte (l'« extérieur » du texte ayant trait à la situation d'énonciation).

26. <http://books.openedition.org/igpde/592>. La fin de cette citation permet de souligner une fois encore combien la transcription ne précède pas l'analyse ; elle y participe. Inscrire un commentaire paratextuel sur la signification d'un rire ou d'un silence, c'est déjà interpréter.
27. « WYSIWYG » est l'acronyme de la locution anglaise « *what you see is what you get* ». Le terme désigne en informatique une interface « intuitive » : l'utilisateur voit directement à l'écran à quoi ressemblera le résultat final.
28. Certains outils ont ainsi été identifiés au cours du travail préparatoire en amont de la rédaction de ce papier mais n'ont pu, faute de temps et car ils sont moins centraux, y trouver place. C'est notamment le cas de : Happy Scribe (<https://www.happyscribe.co/>), MacSpeech Scribe (<http://www.application-systems.eu/scribe/>), SoundScriber, Transkriptshun (<http://www-personal.umich.edu/~ebreck/code/sscriber/>), Voice Walker (<http://voicewalker.software.informer.com/2.0/>).
29. Cet exemple est tiré (et modifié) de Parisse et Morgenstern (2010: 8). Dans CLAN, les « ● » représentent des indicateurs d'alignement temporel texte/vidéo.
30. Parisse et Morgenstern, 2010: 10. Pour une description des logiciels au format partition, voir Rohlfing et al., 2006.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Références

- Blanche-Benveniste C (1997) *Approches de la langue parlée en français*. Paris/Gap : Ophrys.
- Blanche-Benveniste C and Jeanjean C (1987) *Le français parlé, Transcription et édition*. Paris : Institut national de la langue française/Didier Érudition.
- Brown R (1973) *A First Language: The Early Stages*. Cambridge, MA : Harvard University Press.
- Descamps F (2001) Chapitre premier : de l'extraction des informations à la transcription intégrale. In : Descamps F (ed.) *L'historien, l'archiviste et le magnétophone. De la constitution de la source orale à son exploitation*. Paris : Comité pour l'histoire économique et financière de la France, 433-49.
- Dister A and Simon A-C (2008) La transcription synchronisée des corpus oraux. Un aller-retour entre théorie, méthodologie et traitement informatisé. *Arena Romanistica* 1(1) : 54-79.
- Duchesne S and Haegel F (2004) *L'enquête et ses méthodes : les entretiens collectifs*. Paris : Nathan.
- Fiorèse S (2016) « Méthodologie de la constitution et de l'exploitation de corpus oraux pour une analyse multimodale ». *Carnet des jeunes chercheurs du Crem (Centre de recherche sur les médiations)* (blog). 26 novembre 2016.
- Genette G (1987) *Seuils*. Paris : Éditions du Seuil.
- Green J, Franquiz M and Dixon C (1997) The myth of the objective transcript : transcribing as a situated act. *TESOL Quarterly* 31(1) : 172-176.

- Haberland H (2012) The politics of transcription : issues of multilingual and other non-standard data. In : Keiko I and Brandt A (eds) *Challenges and New Directions in the Micro-Analysis of Social Interaction*. Osaka : Division of International Affairs/Kansai University, 3-8.
- Joly A and O'Kelly D (1990) *Grammaire systématique de l'anglais : structures fondamentales*. Paris : Nathan.
- Maingueneau D (1996) *Les termes clés de l'analyse du discours*. Paris : Éditions du Seuil.
- Mondada L (2008a) Contributions de la linguistique interactionnelle. In : Durand J, Benoît Habert B and Laks B (eds) *Discours, pragmatique et interaction*. Paris : EDP Sciences.
- Mondada L (2008b) Documenter l'articulation des ressources multimodales dans le temps : la transcription d'enregistrements vidéos d'interactions. In : Bilger M (ed.) *Données orales. Les enjeux de la transcription*. Cahiers de l'Université de Perpignan 37. Perpignan : Presses universitaires de Perpignan, 127-156.
- Ochs E (1979) Chapter 3 : transcription as theory. In : Ochs E and Schieffelin B (eds) *Developmental Pragmatics*. New York/San Francisco/London : Academic Press, 43-72.
- Parisse C and Morgenstern A (2010) Transcrire et analyser les corpus d'interactions adulte-enfant. In : Veneziano E, Salazar Orvig A and Bernicot J (eds) *Acquisition du langage et interaction*. Paris : L'Harmattan, 201-222.
- Rastier F (1998) Le problème épistémologique du contexte et le statut de l'interprétation dans les sciences du langage. *Langages* 32(129) : 97-111.
- Rohlfing K, Loehr D, Duncan S, Brown A, Franklin A, Kimbara I, Milde J-T, Parrill F, Rose T, Schmidt T, Sloetjes H, Thies A and Wellinghoff S (2006) Comparison of multimodal annotation tools – workshop report. *Gesprächsforschung – Online-Zeitschrift Zur Verbalen Interaktion* 7 : 99-123.
- Sacks H, Schegloff E and Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4) : 696-735.
- Stallman R (2009) *Free Software Free Society : Selected Essays of Richard M. Stallman*. Boston : GNU Press/Free Software Foundation Inc.
- Stuart B (2008) *Principles of Operating Systems : Design and Applications*. Advanced Topics. Cengage Learning EMEA.
- Sung-Yul Park J and Bucholtz M (2009) Introduction. Public transcripts : entextualization and linguistic representation in institutional contexts. *Text & Talk – An Interdisciplinary Journal of Language, Discourse & Communication Studies* 29(5) : 485-502.
- Thibault P and Vincent D (1988) La transcription ou la standardisation des productions orales. *Linx* 18(1) : 19-32.
- Vincent D (2001) Les enjeux de l'analyse conversationnelle ou les enjeux de la conversation. *Revue québécoise de linguistique* 30(1) : 177-198.