

Compte-rendu de l'ouvrage de
Norbert Dittmar :
*Transkription. Ein Leitfaden mit Aufgaben
für Studenten, Forscher und Laien*¹

François MULLER
Université de Paris

Norbert Dittmar est professeur de linguistique allemande et de linguistique générale au Département de Lettres et Sciences Humaines de l'Institut de Philologie allemande et néerlandaise de l'Université Libre (FU) de Berlin. Il a publié déjà plusieurs ouvrages de sociolinguistique (*Soziolinguistik*, 1973 ; *Grundlagen der Soziolinguistik* [Fondements de la sociolinguistique], 1997 ; *The discourse marker also in L2 acquisition*.)

Transkription est une présentation générale des systèmes de transcription en usage pour la transcription de la langue orale à des fins d'archivage et de recherche. La partie centrale de cette étude est constituée par la présentation détaillée des principaux systèmes de transcription actuellement utilisés en Allemagne par les sociolinguistes et plus généralement par les chercheurs en sciences humaines (ethnologie, didactique, apprentissage des langues, etc.) en vue de la représentation graphique de systèmes de signes oraux et non verbaux dans la communication de la vie quotidienne et professionnelle. Le souci constant de concilier un exposé théorique de haut niveau et un ensemble d'informations pratiques, notamment informatiques (CD, programmes et sites Internet) constitue sans conteste le mérite et la grande originalité de l'ouvrage, qui comprend aussi un glossaire des principaux termes techniques utilisés, ainsi qu'une bibliographie bien documentée. L'ouvrage

¹ Norbert Dittmar. *Transkription. Ein Leitfaden mit Aufgaben für Studenten, Forscher und Laien*. - Qualitative Sozialforschung 10 - Lesken + Budrich, Opladen, 2002. - 247 pages. ISBN 3-8100-2650-6 - 15,50 euros.

François Muller

certain nombres d'exercices pratiques dont les solutions se trouvent... sur la « home page » de l'auteur (http://www.userpage.fu-berlin.de/~nordit/HP_LINK:Transkriptionsbuch).

L'ouvrage débute par une réflexion sur le besoin – ou la nécessité – de vouloir ou de pouvoir disposer de traces écrites de ce qui a été dit, car les hommes savent depuis la plus haute antiquité que si les paroles volent, les écrits, eux, restent. Cette réflexion sur le primat du visuel sur l'auditif sert d'introduction générale à l'ensemble du texte ; elle porte sur la constitution de la mémoire collective, sur le haut degré d'authenticité dont bénéficie a priori l'écrit par rapport à l'oral et sur l'aspect contraignant, voire contractuel, du « document » écrit.

La transcription est le prérequis indispensable à toute étude scientifique de la communication orale, car elle fournit sur un support matériel (papier) ou virtuel (informatique) la matière première de la langue parlée dans ses multiples manifestations. L'auteur cherche ensuite à cerner de plus près les spécificités de la langue parlée en les opposant à celles de la langue écrite et esquisse à grands traits l'état actuel de la recherche en ce domaine.

Le chapitre suivant cherche à établir dans quelle mesure et dans quelles conditions une transcription peut être considérée comme une démarche scientifique. En d'autres termes, comment établir la validité de la pratique du transcripteur ? En effet, quelles que soient par ailleurs leurs orientations théoriques, toutes les disciplines des sciences humaines ont impérieusement besoin d'une théorie de la pratique sociale qui inclut les systèmes de transcription eux-mêmes. Mais comme une telle théorie ne peut se fonder que sur de vastes corpus de comportements communicatifs - au sens large - existant réellement, la constitution de cette documentation de processus communicatifs devient un problème épineux. Arrivé à ce point, on s'attend à ce que l'auteur évoque alors le problème bien connu des spécialistes d'herméneutique en général et de la critique des textes en particulier, c'est-à-dire ce que, à la suite de Schleiermacher, on a appelé la question du cercle herméneutique. En effet, depuis qu'Ehlich et Rehbein (1976) ont montré que toute transcription est aussi une interprétation, la problématique du cercle herméneutique s'imposait plus ou moins, ne serait-ce qu' en filigrane. Mais il n'en est rien, et c'est un peu dommage, car le rapprochement eût été intéressant, tant pour la problématique de la transcription elle-même que pour la contribution, même réduite à un éclairage nouveau, que cette problématique aurait pu apporter à la compréhension générale de la question du cercle herméneutique.

Le quatrième chapitre pourrait s'intituler : « Dis-moi comment tu parles et je te dirai qui tu es. » Dans l'impossibilité de pouvoir rendre compte de la charge affective, voire émotive – ce que l'auteur appelle « l'aura » – que recèlent certaines voix, on en est réduit à restituer leur couleur, c'est-à-dire leurs variétés sociologiques et/ou dialectales. C'est ainsi que Jane Austen (*Emma*, 1806), mais aussi Fontane, Tucholsky, et en France Balzac (on pourrait ajouter Maupassant) ont cherché à décrire le plus fidèlement possible les écarts essentiellement (mais pas uniquement) phonétiques par rapport à la langue standard de leur époque. Signalons que l'auteur s'est ici limité aux romans, omettant les parlers hauts en couleurs des servantes et valets de Molière, près d'un siècle et demi avant Jane Austen. Cette transcription peut être qualifiée de « littérale », car elle se contente d'utiliser les « lettres » de l'alphabet, sans recours à aucun signe diacritique. Klein et Schütte (*Transkriptionsrichtlinien für die Eingabe in DIDA*, 2000) considèrent ce système élémentaire de transcription comme une variante de la « translittération ». D'autres parlent ici de « dialecte visuel ». J.A. Edwards (*Transcription of Discourse*, 1992) est très critique pour ce sabir qu'elle appelle « an impressionistic extension of English spelling. » Sans aller aussi loin, N. Dittmar, tout en ayant conscience que la translittération est un pis-aller, c'est-à-dire un compromis acceptable entre la transcription orthographique classique et les transcriptions phonétiques dites « scientifiques », regrette tout de même que la translittération n'ait pas fait l'objet d'une uniformisation sur le plan international, comme ce fut le cas pour l'alphabet phonétique international de L'API (Association Phonétique Internationale), sur lequel l'auteur s'étend longuement. Le chapitre se termine par une brève présentation de deux autres systèmes de transcription conçus spécialement en vue du traitement informatique. Le premier, le PDL (*Pidgin- Deutsch-Lautschrift*), a été élaboré en 1974 par des chercheurs de l'université de Heidelberg. Le principe en est le suivant : pour ne pas avoir à recourir à la multitude de signes et de symboles de l'alphabet phonétique international, ceux-ci ont été remplacés par des lettres latines, ce qui facilite grandement l'implémentation. SAMPA (*Speech Assessment Methods Phonetic Alphabet*) a été mis au point dans les mêmes buts, mais par des chercheurs anglo-saxons (cf. Gibbon, Moore, Winski : *Handbook of Standards and Resources for Spoken Language Systems*, 1997). Pour le codage, SAMPA, à la différence de PDL, a recours au code ASCII.

Le coeur du livre, au dire même de l'auteur, est le grand chapitre cinq, qui compte une centaine de pages, soit près de la moitié de l'ouvrage et qui présente sept systèmes différents de transcription. En prolégomènes à cette

François Muller

présentation exhaustive, N. Dittmar cite les six maximes formulées par J.W. Du Bois (*Transcription design principles for spoken discourse research*, 1991, et J.W. Du Bois *et alii* : *Discourse transcription* - Santa Barbara papers in linguistics, vol. 4, 1992) :

1. Définis des catégories de transcription qui soient parfaitement adaptées aux buts de ta recherche.
2. Fais en sorte que ton système soit simple (c'est-à-dire aussi facilement déchiffrable que possible).
3. Choisis des caractères stables et bien définis.
4. Applique le principe d'économie dans l'inventaire des signes.
5. Module ton système afin qu'il puisse s'adapter à différentes fonctions.
6. Organise ton système en vue d'une application informatique la plus simple possible.

L'auteur, en accord avec A. Deppermann (*Gespräche analysieren. Eine Einführung in konversationsanalytische Methode*, 1999), insiste sur un point qui lui tient manifestement à coeur, car il y revient à plusieurs reprises, c'est-à-dire sur la première maxime : trop de chercheurs utilisent en effet des systèmes trop élaborés qui de ce fait compliquent inutilement le travail en aval. Il oppose ainsi une transcription élémentaire (*Basistranskript*) à une transcription fine (*Feintranskript*).

Avant de passer à la description proprement dite des sept systèmes retenus, l'auteur présente une grille de lecture qui lui servira à les évaluer. Pour ce faire, il a retenu sept critères :

1. Le « design » (c'est-à-dire l'architecture ou la configuration générale), qui comprend trois points :
 - 1.1. la présentation d'ensemble : Soumise à la linéarisation de l'énoncé, la transcription est réduite à choisir entre (a) la présentation sous une forme analogue à celle d'une partition musicale, (b) l'agencement séquentiel et (c) la présentation sous forme de colonnes. Aucune des trois façons de procéder n'est pleinement satisfaisante, bien que la première soit la plus aisée non seulement à mettre en oeuvre, mais aussi à exploiter.
 - 1.2. La translittération proprement dite : quel est l'inventaire des signes retenus ?

- 1.3. l'en-tête, dans laquelle sont données des informations générales concernant les conditions dans lesquelles a été faite la transcription.
2. Les contributions des divers participants, ce qui implique les données verbales, mais aussi communicationnelles et prosodiques (Identification des locuteurs, informations sur le début et la fin de la communication, les chevauchements, les enchaînements, etc.)
3. Unités énonciatives. Les unités minimales de communication sont : les mots, les limites de l'énoncé (début / fin), le fragment de mot (interruption), réactions de ou des interlocuteur(s). L'auteur rappelle qu'il n'est pas toujours aisé de segmenter le flux verbal, surtout chez des enfants ou des locuteurs qui maîtrisent mal la langue. Mais cette première étape n'en est pas moins décisive, car elle constitue un encodage minimal incontournable.
4. Phénomènes prosodiques : l'auteur en a retenu neuf : la structure intonative des mots, l'intonation des syntagmes, l'accent de mots, l'allongement des sons, l'intensité (*forte* VS *piano* ; *crescendo* VS *diminuendo*), le tempo (*allegro* VS *lento*) et l'accélération (*accelerando* VS *rallentando*), le rythme (combinaison de la hauteur, de l'intensité et de la durée : *staccato* ; *glissando*, *legato*), manifestations physiques qui accompagnent la parole (tousotements, rires, soupirs, etc.), et enfin les pauses (silences, hésitations, vocalisations).
5. Manifestations non verbales (proxémique, motrice, mimique, gestique, etc.) : elles ne sont pas toujours prises en compte dans la transcription.
6. Commentaires (point de vue du transcripteur) : l'auteur estime qu'il est parfois important de signaler tous les faits saillants qui peuvent surgir dans la communication et qui font sens, d'une manière ou d'une autre.
7. Signes spéciaux (« Extras ») : N. Dittmar en a sélectionné huit : la restitution phonétique, les noms propres, les traductions, les changements de langues, le signalement des passages non transcrits, les signes prévus pour autre chose que la transcription de la communication, les signes diacritiques optionnels, les signes conventionnels pour les juxtapositions de plusieurs signes particuliers.

François Muller

L'auteur entame ensuite l'examen des sept systèmes retenus : l'analyse de conversation (KA), HIAT, DIDA, le système de Du Bois, GAT, CHAT et CHILDES. Les diverses présentations suivent à peu près toutes le même canevas : historique, l'idée directrice, présentation des signes conventionnels, et se terminent toujours par un exemple de transcription.

1. L'analyse de conversation (KA)

L'auteur rappelle qu'à la fin des années soixante et au début des années soixante-dix, deux concepts étaient particulièrement répandus dans le milieu des chercheurs en sociologie: l'ethnométhodologie et « l'interaction sociale » (Goffman). Les sociologues de l'époque s'étaient fixé pour objectif d'étudier les pratiques sociales des individus dans leur vie quotidienne essentiellement sinon exclusivement sous l'angle de leurs interactions verbales. Pour ce faire, ces chercheurs ne travaillaient pratiquement que sur des transcriptions. L'accent était mis sur la structure séquentielle des contributions des différents participants. Une attention particulière était de ce fait accordée aux transitions d'une contribution à l'autre, ainsi qu'aux fonctions expressives de la parole (opposition : voix forte VS faible, mise en exergue d'un mot ou d'une expression, allongement des voyelles, aspiration, pauses, etc.). L'analyse de conversation a réussi à s'imposer en Europe essentiellement grâce à Levinson (1983). L'appréciation générale portée par l'auteur sur l'analyse de conversation est largement positive : les catégories sont bien définies, les signes sont pour la plupart alphabétiques, donc aisément utilisables. C'est ce qui explique que de nos jours encore, un grand nombre de spécialistes en communication, mais aussi de linguistes et de sociologues continuent à se servir des conventions établies par l'analyse de conversation. Une réserve cependant : l'implémentation du système tel qu'il se présente depuis le début des années quatre-vingt-dix se fait attendre.

2. HIAT

Avec « le tournant pragmatique » (Austin, Searle, Levinson) dans les années soixante-dix, l'activité langagière devint objet d'investigations

scientifiques. On se mit à archiver électroniquement les actes de paroles pour en étudier le fonctionnement. Les promoteurs de cette démarche nouvelle étaient en Allemagne Ehlich et Rehbein. Ils appréciaient le système de transcription de l'analyse de conversation, mais ils avaient rapidement discerné ses faiblesses ou tout simplement ses limites. C'est surtout l'agencement purement séquentiel qu'ils jugeaient peu opérationnel. C'est pourquoi ils optèrent pour la transcription sous forme de partition musicale, plus à même, selon eux, de rendre compte de l'interaction verbale quasi polyphonique des divers participants. Abandonnant la segmentation du discours en unités, ils préconisent des « marqueurs du flux verbal ». – Si l'on excepte DIDA (IdS, Mannheim), HIAT est sans conteste le système le plus utilisé en Allemagne, notamment par les pédagogues, les linguistes et les sociologues, ainsi que par les transcrip-teurs institutionnels (à l'école, à l'hôpital, au tribunal). Cinq raisons majeures expliquent, selon N. Dittmar, ce large succès :

- la présentation d'ensemble sous forme de partition offre une synopsis très claire, qui se laisse aisément déchiffrer. En effet, une portée étant attribuée à chaque participant, la synchronicité et la diachronicité de chaque contribution sont fidèlement documentées ;
- le système est très flexible : la transcription « primaire », qui fait abstraction de toute segmentation fine, peut être réalisée avec une grande économie de signes ; c'est la raison pour laquelle on l'appelle « semi-interprétative » ;
- les auteurs du système ont eu un souci constant, et ce dès la première version, d'apporter des solutions concrètes à tous les problèmes pratiques qui pouvaient se poser (par ex. pour les corrections) ;
- de même, les auteurs ont été très attentifs à l'implémentation de leur système, tant sur Macintosh que sur MS-DOS ;
- l'implémentation initiale, fondée sur des symboles essentiellement musicaux, a été progressivement abandonnée au profit d'un système plus convivial.

3. DIDA

DIDA (Diskurs-DAten-Bank) a été développé au début des années quatre-vingt-dix à l'Institut de la langue allemande (IdS) de Mannheim. Il s'agissait de mettre au point un système conforme aux dernières avancées

François Muller

de l'analyse du discours et de la conversation, afin d'exploiter par l'informatique les nombreux corpus que l'IdS avait accumulés au cours des dernières décennies. Au départ, ce système n'avait été conçu qu'à l'usage interne de l'IdS ; mais par la suite il rendit aussi d'appréciables services à d'autres chercheurs en sciences humaines. C'est ainsi par exemple que le corpus de la langue parlée dans la ville de Mannheim (cf. Kallmeyer : *Kommunikation in der Stadt*, 1994) a été recodé en DIDA. On notera que le système prend en compte des paramètres comme le tempo du flux verbal, ainsi que l'intensité de la voix. Parmi les nombreux mérites du système, N. Dittmar souligne essentiellement l'excellente iconicité des signes utilisés et l'existence d'une multiplicité de programmes d'application spécifiques, comme par exemple PRAAT pour l'étude des phénomènes prosodiques.

4. Le système de Du Bois

Le système de Du Bois (DT : Discourse Transcription) se situe dans le droit fil des réflexions initiées dès 1979 par E. Ochs. Par rapport à ses prédécesseurs et concurrents, DT présente essentiellement trois avantages :

- Du Bois et son équipe ont formulé des exigences précises en matière de lisibilité, d'iconicité et d'implémentation ;
- les auteurs prennent grand soin de distinguer la transcription proprement dite, qui est avant tout interprétative, de l'encodage, qui doit se faire en fonction du but recherché ;
- ils établissent ensuite une seconde distinction (toute scalaire) entre transcription « large », à caractère purement descriptif, et transcription « étroite », à visée plus scientifique.

5. GAT

GAT (*Gesprächsanalytisches Transkriptionssystem*), créé en 1997 par des linguistes (Selting, Auer), s'est fixé comme objectif de faire la synthèse des systèmes existants afin de mettre à la disposition des chercheurs de langue allemande un système unifié de transcription. On facilite ainsi les échanges de gros corpus non seulement au sein d'une même discipline, mais surtout entre les différentes disciplines des sciences humaines, ce qui était aussi l'un des buts des créateurs de GAT. Ce système a été conçu

avant tout pour l'étude de la langue courante dans le cadre d'une approche pragmatique de la communication. Le « design » présente les caractéristiques suivantes :

- notation séquentielle ;
- tout doit être écrit en minuscules (ce point n'est pas sans importance en allemand, où tous les noms, y compris les noms communs, s'écrivent avec une majuscule) ;
- dans l'en-tête (*Transkriptionskopf*) doivent figurer un certain nombre d'informations précises, notamment sur les conditions matérielles de la transcription.

GAT a développé de façon satisfaisante les conventions de l'analyse de conversation (KA). Il préconise par ailleurs de se fonder sur les unités prosodiques et non sur les unités énonciatives.

GAT est utilisé surtout par les linguistes, notamment par les spécialistes de l'analyse de la langue orale, qui utilisaient antérieurement KA.

6. CHAT et CHILDES

CHAT (*Codes for Human Analysis of Transcripts*) a été mis au point dans le cadre du projet CHILDES (Child Language Data Exchange System), sur lequel travaillent actuellement en Allemagne plus de quatre-vingt groupes de recherche. Le but de l'opération est de constituer une banque de données sur l'acquisition de la langue maternelle et de mettre à la disposition des chercheurs un système de transcription unifié. CHILDES comprend depuis 1991 trois éléments :

- un système de transcription et d'encodage (CHAT) ;
- une banque de données ;
- une série de programmes (CLAN) pour l'analyse des transcriptions.

CHAT présente les caractéristiques suivantes :

- Il ne se limite pas à la seule transcription des sons, mais note aussi les informations paraverbaux et non verbales ;
- toute nouvelle énonciation est notée sur une nouvelle ligne (même si le locuteur reste le même) ;

François Muller

- les informations complémentaires sont consignées dans un commentaire distinct de la transcription elle-même ;
- un système de notation très riche est mis à la disposition des transcrip-teurs pour qu'ils puissent, dans un souci d'authenticité maximale, noter toutes les particularités des différentes langues.

CHAT a le grand mérite, aux yeux de l'auteur, de bien séparer les tâches : ainsi l'observation (c'est-à-dire la transcription proprement dite) ne doit pas être confondue avec l'analyse (ou l'exploitation) scientifique des données.

Dans les trois derniers chapitres, beaucoup plus brefs, N. Dittmar traite successivement des informations non verbales (proxémique, cinésique, gestique, mimique et orientation du regard), ainsi que de l'insertion d'images dans les transcriptions. (chapitre 6). Ensuite il aborde les problèmes informatiques (logiciels, Internet comme source d'informations et d'accès aux grands corpus) (chapitre 7). Et il clôt son propos par des conseils très pratiques (par ex. sur le choix du système de transcription), conformément à l'esprit de l'ensemble de l'ouvrage, qui a privilégié tout au long de ses 247 pages une approche très conviviale de problèmes éminemment techniques, ce qui n'exclut évidemment pas une grande rigueur et une remarquable clarté dans la présentation de la matière elle-même.