HAL
open science

Splendeurs et misères de la transcription

Berthille Pallaud

To cite this version:

Berthille Pallaud. Splendeurs et misères de la transcription. Colloque "Langage et Psychanalyse", 2001, Aix-en-Provence, France. pp.59-73. hal-00136752

HAL Id: hal-00136752

https://hal.science/hal-00136752v1

Submitted on 19 Apr 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Berthille Pallaud

Langues Méditerranéennes 2003
p 59-72.

Splendeurs et misères de la transcription

INTRODUCTION

« *Il apparaît qu'une transcription fiable se doit d'être une transcription critique* »
Phrase non pas de linguiste mais du psychanalyste Jean Allouch.

La pratique de la transcription d'énoncés enregistrés ne peut se faire sans que cette pratique entraîne une réflexion sur l'analyse de discours que ne peut manquer de convoquer la transcription pour avoir lieu. C'est pourquoi, je propose d'envisager à travers l'activité de transcription d'énoncés enregistrés :

– la transcription d'énoncés : quelle écoute ? ;
– les hypothèses de l'écoute pour la transcription ;
– les difficultés liées à l'identification et à la transcription de lapsus.

Je ferai référence à l'expérience de transcription acquise dans le cadre du Groupe aixois de recherche en syntaxe que dirigeait Claire Blanche-Benveniste jusqu'en 1999 et qui était chargé d'établir et d'analyser des corpus de français parlé.

J'inclurai également certaines des remarques développées par le groupe *stécriture* de l'École de psychanalyse lacanienne (EPL) dans son livre *Le transfert dans tous ses errata* suivi de *Pour une transcription critique des séminaires de Jacques Lacan* (EPL, 1996).

Berthille Pallaud, UMR CNRS *6057, « Parole et langage », Département de linguistique française, université de Provence, Aix-Marseille 1, 29 av. R. Schuman, 13621 Aix cedex.*

TRANSCRIPTION D'ÉNONCÉS : QUELLE ÉCOUTE ?

C'est un paradoxe maintes fois souligné : on ne peut travailler sur le texte oral sans passer par sa trace écrite. Il est vrai qu'actuellement les nouvelles technologies permettent la saisie numérique de la source orale. Il est alors possible de la confronter indéfiniment, point par point, grâce à des procédures d'alignements diverses, au texte écrit résultant de sa transcription. Dans notre domaine linguistique, la transcription d'un énoncé a pour objectif l'établissement de corpus dont les recueils constituent actuellement de véritables ressources langagières. Suivant les cas, la transcription peut être phonétique ou orthographique.

Transcrire orthographiquement un énoncé revient à interpréter l'énoncé oral (Blanche-Benveniste *et al.*, 1990) et donc à identifier des morphèmes. Lorsque Lacan mentionne la transcription, en l'opposant à l'écriture il parle de la transcription phonétique : « L'écriture, loin d'être transcription, est un autre système un système auquel éventuellement s'accroche ce qui est découpé dans un autre support, celui de la voix » (1968, p. 328).

Dès la phase de transcription orthographique, l'établissement du texte de l'énoncé repose sur une écoute autrement dit sur une analyse de discours implicite et particulière. Quelles en sont les caractéristiques ?

1. Une transcription orthographique vise à l'identification des morphèmes et établit donc la signification de l'énoncé mais *non son sens.* Elle n'a pas de visée discursive. Elle ne lève pas les ambiguïtés discursives. Seuls les locuteurs au cours de leur énoncé peuvent préciser, éventuellement, quelque sens de leurs propos :

L1 (journaliste) – la nature nous parle mais nous ne l'écoutons pas disait Victor Hugo

L2 (Jacques Chirac) – je ne suis pas sûr qu'il pensait à l'écologie.

Elle ne se préoccupe pas de signaler ou lever l'ambiguïté qui peut apparaître dans un énoncé comme celui-ci : « Oublions cette époque où les Français ne s'aimaient pas. »

Ce type d'énoncé ne recèle aucune difficulté pour établir la signification de chacun de ses éléments et donc sa transcription. Il n'en est pas de même pour l'établissement de son sens qui impose la connaissance des circonstances de son apparition, en l'occurrence les années 1940 en France. Le locuteur rappelait, à propos du procès Touvier, le peu de réactions publiques à l'encontre des lois anti-juives qui touchaient pourtant des Français. L'interlocuteur, lui, répondait en parlant de « la fracture sociale » apparue dans les années 1980 et dénoncée, il y a peu, par le chef de l'État français.

2. Si la transcription ne précise pas le sens de l'énoncé qui reste à la charge du lecteur, elle n'identifie pas non plus des *actes de langage* (au sens de Searle, 1972). Elle s'attache à la signification de l'énoncé et non à sa valeur d'énonciation (encore moins à l'identification sociale ou individuelle de cette valeur d'énonciation). *L'intention* individuelle, si elle assigne à l'acte d'énonciation sa valeur propre, est ignorée du transcripteur. La transcription ne distingue pas entre ce qui est intentionnel ou non dans les propos du locuteur. Cette affirmation est posée par Claire Blanche-Benveniste dès le début des recueils de français parlé en 1977 : « Au risque de choquer des habitudes bien ancrées, nous ne distinguerons pas entre les phénomènes qui paraissent créés par la volonté des locuteurs et ceux qui paraissent leur échapper ; nous traiterons de la même façon des phénomènes apparemment involontaires comme bredouillages, hésitations, maladresses, reprises, et d'autres qui semblent intentionnels comme : répétitions intensives, variations stylistiques et autres. On verra que ce point de vue, s'il néglige la part d'intention du locuteur, sur laquelle nous ne pensons pas pouvoir porter de jugement, a l'avantage de suivre une ligne d'analyse grammaticale unifiée » (Blanche-Benveniste *et al.*, 1990, p. 20).

3. Cette position est d'autant plus légitimée que l'énoncé prononcé *n'est pas adressé au transcripteur* qui est ignoré du locuteur (ce qui ne veut pas dire que le locuteur ne sache pas qu'il y aura transcription de ce qu'il dit) [1]. Le locuteur tiers ou partenaire est celui qui enregistre la conversation, l'entretien ou le débat. Le transcripteur est très rarement celui qui a fait l'enregistrement. D'ailleurs, l'établissement de la transcription d'un même enregistrement mobilise plusieurs personnes aussi bien pour la transcription que pour sa vérification. Il est évident que cet énoncé ne leur est pas adressé.

4. Même si les propos du locuteur ne sont pas adressés au transcripteur, la transcription portera les traces des relations transférentielles que ce dernier établit lors de l'établissement du texte. On retrouvera tout ce que le transcripteur *imagine* sur le locuteur au cours de sa transcription, qui en sera le reflet, particulièrement en ces points où « le texte fuit par certains bouts » suivant la jolie formule de Claire Blanche-Benveniste et Colette Jeanjean (1987, p. 112).

Dans l'exemple suivant, le locuteur qui parle évoque son travail dans un foyer de semi-liberté pour jeunes filles en réinsertion scolaire : « Le fait qu'il

1. La législation actuelle sur la propriété de la voix veille à ce que le locuteur puisse donner son accord sur le devenir de ses paroles enregistrées.

y ait niveau /*bac*, bas et que / peut-être ça pose aussi un problème » (Strasbourg C2, 4, 7).

Pour le transcripteur, il était peu probable que le mot soit *bac* et pourtant c'était bien le cas. Il est vrai que, dans la majorité des cas, à l'inverse de l'exemple précédent, il est impossible de trancher entre deux solutions entendues : « Si j'entends aussi bien "a" que "b" dans une séquence de français parlé, je ne saurai donc jamais ce que le locuteur a "dit en vrai." Il semble bien que la réponse soit "non", même si on améliore les appareils d'enregistrements, et l'oreille de celui qui écoute, même si l'on met les gestes et les mimiques et même si l'on questionne le locuteur sur ce qu'il a voulu dire. Le texte authentique fuit, du moins par certains bouts » (Blanche-Benveniste et Jeanjean, 1987, p. 112).

L'*écoute* est donc une activité interprétative qui ne vise pas qu'à une stricte correspondance phonologique lorsque le transcripteur établit le texte. Ce faisant, il peut même sembler faire preuve d'erreurs perceptives plus nombreuses que celles suscitées par les difficultés d'ordre purement perceptif. L'oreille qui écoute n'est pas un traître mais elle est surtout asservie à la recherche de signification. Par là, il n'est pas rare qu'elle semble s'écarter de la source orale. Ce fait a pour conséquence qu'il est vain de prétendre déterminer avec certitude quand la source auditive est seule responsable de l'hésitation perceptive du transcripteur : « Phonéticiens et psychologues s'accordent à dire que la perception est un processus actif qui s'appuie sur la compréhension des énoncés » (Blanche-Benveniste et Jeanjean, 1987, p. 103).

L'écoute du transcripteur produit donc un énoncé et une énonciation qui seront *aussi* les siens au moins en certains bouts : une sorte de maillage non pas inter-locuteurs mais *inter locuteur-transcripteur*, un genre de *texte collectif*. Plus la transcription d'un énoncé est revue, réécoutée, plus le texte initial se surcharge. Cela fait le désespoir de la linguistique informatique qui, pour l'automatisation de ses études sur de grands corpus, a besoin de réduire les équivocités et surtout les hésitations interprétatives.

Certains passages de l'énoncé enregistré sont difficiles ou même impossibles à décrypter : cela peut tenir à plusieurs causes toujours intimement liées.

a) Il peut s'agir de difficultés liées à l'enregistrement bien que les nouvelles technologies et les possibilités actuelles de traiter le signal facilitent beaucoup, de ce point de vue, la tâche du transcripteur.

b) Les difficultés peuvent être liées aux *facilités de prononciation* (Gadet, 1989) notamment dans des formules comme : « Ça a été vu, tu as vu » rarement prononcé comme telles. Ces passages sont à l'origine de divergences dans les

conventions de transcription et la *notion de fidélité* qui est au cœur de ces divergences soulève de nombreuses questions. Notamment : à quoi être fidèle ? L'établissement d'un texte total qui rende compte de l'énoncé dans tous ses états est-il possible ?

Si on veut rendre compte des variations dans la prononciation, où s'arrêter ?

c'a été, t'as vu, p'tit, maint'nant, pask

MUSIQUE *16,1 par exemple moi quand j'écris un texte ben ch- je sais pas je le sens euh des fois je le sens vrai*

Si les exemples précédents essaient de rendre compte des phénomènes de prononciation, la transcription orthographique privilégie le rétablissement des morphèmes et de la syntaxe ce qui entraîne *une infidélité au niveau de la prononciation* :

ça a été, tu as vu, petit, maintenant, parce que

MUSIQUE *16,1 par exemple moi quand j'écris un texte ben j- je sais pas je le sens euh des fois je le sens vrai*

Cette stratégie de transcription a été choisie dès le début par le Groupe aixois de recherche en syntaxe pour ses études syntaxiques. Elle a été adoptée pour la constitution de corpus dans le cadre des nouvelles technologies liées aux ressources linguistiques. Les raisons de son adoption tiennent au moins autant à des considérations techniques que linguistiques [2].

Le constat est que les objectifs visés par les transcripteurs conduisent à des choix dans les *fidélités prioritaires*. Les nouvelles technologies en permettant la superposition des supports des énoncés oraux (son, transcription orthographique) et des analyses (prosodique, phonétique, morphologique et syntaxique, pragmatique par exemple) facilitent l'accès aux différents aspects de l'énoncé et donc aux divers objectifs d'analyse. C'est ce que s'est proposé de faire, dès ses débuts en 1981, le projet américain d'études du langage de l'enfant CHILDES [3] (Macwinney, 1991) et que poursuivent notamment les équipes aixoises Parole et Langage et DELIC.

c) Parfois, *les conditions d'écoute* provoquent une hésitation perceptive c'est-à-dire des solutions de transcriptions concurrentes. Cette hésitation qui rend compte, elle aussi, de l'activité d'énonciation du transcripteur n'est restituée que si les conventions de transcription adoptées prévoient la possibilité de *multi-transcriptions* :

mais je / v*eux*, v*ais*/ répondre à Jean (Navale, 3é, 3)

---

2. Le traitement des grands corpus impose l'élimination des variations équivoques et donc les trucages orthographiques inventés pour rendre compte de la diversité de prononciation.

3. *Child Language Data Exchange System* : principes de transcription CHAT et logiciels d'analyse automatique CLAN.

SCALONE2 32,8 on / parl*e*, parl*ait* / de qua- de trente-neuf pas
AIDESD 3,3 moi je viens à la salle pour euh quand j'ai f- euh enfin */le, un/* prof nous donne un exercice
JEANNE 1,8 pendant trois mois dans une société qui utili- qui qui utilisait */des, les/* sous-produits

Les équipes travaillant sur des corpus oraux ont le plus souvent décidé de conserver dans la transcription les « multi-écoutes ». D'autres, cependant, ne les mentionnent pas ou seulement partiellement (Francard et Pérronnet, 1989 ; Mouchon et Fillol, 1980 ; Thibault et Vincent, 1990 ; Vion, 1990).

Les psychanalystes qui ont contribué à la transcription critique du séminaire de Lacan sur *Le transfert* (EPL, 1996) ont souligné la nécessité du recours aux multi-transcriptions dans le contexte de la distinction entre ce qui est faute et ce qui est variante : *La faute dès lors est une variante dépourvue d'autorité, non retenue comme variante* (p. 224).

Exemple de faute retenue par EPL (1996, p. 19) et qui conduit à remplacer un ou des éléments par d'autres :

« p. 14-26
On doit se demander par quels moyens opérer honnêtement avec les désirs
On doit se demander par quels moyens opérer honnêtement avec le désir
Il s'agit du désir dans sa singularité et non d'une éventuelle liste des désirs. »

La variante signalée (EPL, 1996, p. 20) correspond à ce que les linguistes, qui travaillent sur corpus et font de la transcription, rangent dans le cadre des multi-transcriptions :

« p. 26-15
J'attendrai *que cette année, nous sachions entre qui et qui nous sommes*
J'entendrai *que cette année, nous sachions entre qui et qui nous sommes*
Variante possible. Autre variante : J'attendrais. »

Une étude sur les modifications apportées à des transcriptions, lors de la phase de vérification (Pallaud, 2002), montre que les erreurs du transcripteur ne se produisent pas au hasard. Le quart des corrections proposées consiste en des remplacements partiels d'énoncés. Les oublis sont assez fréquents et beaucoup plus nombreux que les ajouts erronés : 35 % contre 5 %. Il est, donc, beaucoup plus courant d'oublier des éléments de l'énoncé que de lui en ajouter ou d'en remplacer.

On retrouve les oublis dans l'édition critique de « Le transfert » de Lacan par EPL (1996, p. 19)

« p. 16-14
[...] Socrate ainsi mis à l'origine, disons-le tout de suite, du plus long transfert
[...] qu'ait connu l'histoire.

[...] qu'ait connu l'histoire de la pensée
de la pensée avait été supprimé. »

LES HYPOTHÈSES DE L'ÉCOUTE POUR LA TRANSCRIPTION

L'écoute de celui ou celle qui transcrit un énoncé s'appuie sur deux hypothèses concernant le locuteur enregistré : le locuteur est *compétent* dans sa langue et son énoncé est *cohérent*. Or, qu'en est-il de cette compétence et de cette cohérence lorsque le locuteur est un enfant, une personne âgée, handicapée ou même aphasique ? On voit que les questions susceptibles d'interpeller la pratique de transcription ne manquent pas de surgir dès lors qu'on sort d'un locuteur idéal.

La cohérence textuelle est à envisager à tous les niveaux d'analyse que ce soit phonétique, phonologique, morphologique (nominale aussi bien que verbale : *je* est en général suivi d'un verbe sauf si l'énoncé est interrompu après ce pronom) ou même syntaxique.

Le niveau argumentatif est celui qui est le plus souvent sous-entendu lorsqu'il s'agit de cohérence. Un discours incohérent est avant tout un discours auquel il est impossible d'attribuer un sens. La solution de transcription adoptée *respecte la cohérence du discours* (cohérence dont le locuteur est crédité *a priori*) et veille à ce qu'elle ne suscite pas de contradiction interne tant au point de vue de l'information que de son organisation (par exemple dans les discours narratif ou explicatif) :

le fait qu'il y ait niveau /bac, bas et que/ peut-être ça pose aussi un problème (StrasbourgC2, 4, 7)

De même, le transcripteur peut disposer d'informations sur le cadre de l'entretien enregistré et interpréter un passage en accord avec ces données extra-textuelles comme c'est le cas dans l'exemple suivant ou seule la première solution est correcte.

*elle travaille sur la restauration des /*lésions, *liaisons/ (Strasbourg41c, 3, 11*

Le cadre où cet énoncé fut prononcé est celui d'un laboratoire de neurobiologie où le locuteur travaille comme agent technique.

Comme le souligne Carter-Thomas (2000), *la plupart des aspects relatifs à la perception de la cohérence d'un texte sont étroitement dépendants de la perception individuelle d'un lecteur dans une situation particulière. Ces aspects sont pour ainsi dire invisibles à la surface du texte.*

La transcription des séminaires de Lacan a souvent eu recours à la cohérence textuelle pour arrêter un choix de transcription, qui, parfois, est loin d'être sans conséquence. Ce recours est parfois explicite EPL (1996, p. 20) :

« *Alors, nous allons repartir à procéder du* A *[...]*
Alors, nous allons repartir à procéder du a *[...]* »
Le paragraphe suivant critique l'intersubjectivité, « la relation berger-bergère ». Il s'agit de *a* au sens de petit autre.

Comme le rappelaient Blanche-Benveniste et Jeanjean (1987), s'il n'y avait pas ces contraintes phonologiques, morphologiques et syntaxiques, une formule comme [jèmalopié] pourrait être transcrites de 2 000 façons différentes (Mariani, 1980) alors que seules deux solutions sont orthographiques : *j'ai mal au(x) pied(s).*

Malgré ces contraintes, l'énoncé oral offre au transcripteur des occasions de multi-transcriptions qui tiennent à deux sortes de facteurs.

Hésitations sur les frontières des mots

Le premier facteur tient à des hésitations sur les frontières des mots c'est-à-dire sur la segmentation de l'énoncé oral en unités de signification. Les frontières de mots sont régies par des principes qui gouvernent les successions de sons admises dans la langue. Ces principes sont prosodiques et concernent trois sortes de paramètres. Les schémas rythmiques de durée aux frontières des unités de sens se caractérisent par un allongement ou un raccourcissement de la durée des voyelles et des consonnes. Par exemple, les études de psycholinguistes (Christophe *et al.*, 1995) montrent que, dès quatre jours, les nourrissons sont sensibles aux indices qui marquent une frontière de mots : allongement de la voyelle terminale et allongement de la consonne initiale. Ils s'avèrent capables de distinguer l'élément ma ti dans la séquence *schéma tigré* de l'élément mati dans le mot *mathématicien*. Les modulations de la hauteur de la voix ainsi que des variations d'intensité sur les syllabes concourent également à fractionner l'énoncé oral en unités signifiantes.

Difficultés liées à l'homophonie

Il reste que certains passages présentent, pour le transcripteur, en dépit de ces indices formels ou d'indices sémantiques, des ambiguïtés liées à l'homophonie (et ce malgré le respect de la cohérence textuelle) :
– *homophonie au niveau des mots ou amorces de mots,* parfois signalés par les locuteurs eux-mêmes :
– *les journalistes dont les revendications* étiques sans h *ne sont même plus entendues rêvent d'un autre monde (Fr. Culture, commentateur, 27/11/96)*

– *on ne sait pas là-dedans qui c'est qui sait. Ca peut s'écrire de deux façons :* qui c'est qui sait, qui sait qui c'est *(Lacan, séminaire d'un autre à l'Autre,* 1968-1969, p. 375).

Freud a montré, dans un cas de perversion, que l'homophonie peut dépasser le cadre d'une langue. Un patient d'origine anglaise ne pouvait tomber amoureux que de femmes ayant « ein *Glanz* auf der Nase » (un brillant sur le nez). Cette contrainte céda du moment où il put évoquer le souvenir de « a *glance* » (un regard) vers sa mère nue. L'inscription, au sens littéral, de cette signification seconde (qui plus est dans une langue seconde) avait cheminé dans la vie du patient en masquant la première écriture et donc le souvenir.
– *homophonie sur plusieurs mots.* « Même si une écriture alphabétique comme le français est fondamentalement phonographique (et a donc vocation à "noter" la prononciation), il est bien connu que le passage d'un mot à l'autre ne respecte nullement le découpage en syllabe, si bien qu'il est fréquent qu'une même syllabe (phonique) se trouve à cheval entre deux mots : *il arrive enfin* réalisé /i-la-ri-ve en-fin/ » (Sabio, 2000).

il restait plus /que, qu'eux/ *les vieux docteurs (Barbara, 9,14)*
FEMMES 3,14 *et bon les* /ren- ses, rensei-/ *euh ses rapports avec mon père euh sont très espacés*
FEMMES 5,1 *je n'étais pas allée aux vêpres et comment el- enfin ben oui elle* /l'a su, a su/

Cette difficulté peut donner lieu à des calembours ce dont témoignent les enfants en cours d'acquisition de l'écriture. Ils passent par une phase de transcription qui rend bien compte de ce phénomène de sur- ou sous-segmentation en mots écrits. On lui doit les écrits suivants :

le nours, des sanfans, tou lmonde

C'est là une observation que tout enseignant d'école primaire (voire même de collège) peut faire. La publicité comme les média ont recours au calembour dans leurs titres :

« *Les maudits mots dits dimanche* ». *Emission de France Inter*

Certains poètes et écrivains humoristiques ont poussé l'emploi du calembour jusqu'à la construction d'holorimes :

Étonnamment monotone et lasse
Est ton âme en mon automne, hélas (Louise de Vilmorin)

Que de jeunes enfants en cours d'acquisition de l'écriture ou des adultes en difficultés révèlent leur assemblage hétéroclite de « bout d'écrits » pour répondre à la demande d'écriture d'un énoncé oral ne fait pas oublier que cette « écriture » inscrit pour le Sujet une signifiance qui lui est propre. Ce n'est pas seulement un jeu de correspondance « fantaisiste » entre un entendu et des lettres. La correspondance est aussi fantasmatique comme le

montre, par exemple, Dolto (1981) qui décrit l'apparition des fantasmes chez des enfants à partir d'un entendu « transcrit » à leur manière où les mots de *salle d'attente* et de *brasero* ont été intégrés dans une histoire inquiétante d'Indiens, de *tentes* et de *bras à zéro*.

C'est ce que révèle également l'écoute clinique suivante dont la clé repose sur une segmentation différente d'un énoncé. Bergès (1991, p. 34) décrit un échange qu'il eut avec une patiente venue le trouver dans une angoisse très grave :

> « *Elle m'explique que cela a commencé alors qu'elle se reposait dans sa maison et pour préciser le début de la chose, elle emploie une phrase au présent :* "J'entends des pas *dans le jardin." C'est à ce moment que je souligne la proposition en la répétant dans une équivocité complète en français :* "J'entendais pas *dans le jardin" qui devient ainsi une forme verbale au passé. Et c'est dans la scansion des mots et non dans leur phonétique que j'interviens. Dès lors, la crise s'apaise, les voix cessent.* »

Hésitations sur la segmentation en propositions

Le deuxième facteur tient à des hésitations sur la segmentation en propositions. Les indices prosodiques ne sont pas toujours présents qui permettent, pour certains constituants syntaxiques (compléments locatifs et temporels, adverbes, les segments flottants associés), de décider de leur caractère enchaîné ou détaché par rapport au constituant précédent (Sabio, 1995). Or dans ces cas, les marques prosodiques conditionnent très directement l'interprétation syntaxique, et leur absence dans l'énoncé oral peut faire obstacle à la compréhension du texte et rendre l'analyse syntaxique indécidable, ou tout au moins hasardeuse. Pour donner un exemple simple, la séquence suivante, empruntée à Sabio :

> « J'ai énormément de mal non pas à me réveiller je me réveille *immédiatement* je me lève et je me dis il faut que je me prépare » *(Rhumato, 222)*.

Il a été impossible de décider si l'adverbe *immédiatement* dépend du verbe le précédant ou de celui qui le suit.

Lorsque l'énoncé est dit d'une voix monocorde, non intonée, comme c'est le cas des voix de synthèse au début des synthétiseurs de voix vocale, même si l'identification des mots est possible, l'impression de discours insensé s'impose du fait même de ces indécidables dans l'interprétation syntaxique. Ce type de phrasé qui peut être également observé dans certains troubles aphasiques ou chez certains psychotiques perturbe sérieusement la communication mais n'empêche nullement la transcription des énoncés. C'est par exemple le cas de ce petit garçon qui, au début de sa psychothérapie produisait un flot de paroles ininterrompu, sans pauses et sans intonation, et dont la séquence suivante est restée énigmatique (Mège-Morin et Pallaud, 2002) :

> *J'ai pas écrit le mot moi* (*le mot : moi* ou *le mot, moi* ?)

LES DIFFICULTÉS LIÉES À L'IDENTIFICATION ET LA TRANSCRIPTION DE LAPSUS

Les énoncés oraux présentent des phénomènes d'achoppement qu'on peut ranger en trois catégories : les erreurs sur les mots (communément nommées lapsus), les amorces de mots et les phénomènes de bribes ou d'interruption.

La catégorie des bribes est délicate à appréhender pour le linguiste car elle est très liée au sens à travers la rupture de cohérence sémantique ou syntaxique.

Dans les corpus du Groupe aixois de recherche en syntaxe, les amorces de mots sont appréhendables de façon univoque car elles sont signalées par une marque de convention précise : Elles sont donc identifiées dans le texte transcrit d'abord par leur *matérialité* : *un bon film s'il est bon il* mar- *eh ben il ne marchera pas (Radio, 02-04 1998)*.

Dans la très grande majorité des cas, ces phénomènes étant identifiés comme des « réductions » de mots qui ne sont qu'amorcés, leur transcription est orthographique et atteste ainsi de cette interprétation. C'est en tout cas la règle de transcription qui est donnée :

– on transcrira : « C'*est* mau- *c'est de très mauvais goût.* »
– et non : « C'*est* mo- *c'est de très mauvais goût.* »

Il reste cependant que si cette interprétation est hautement probable, elle ne peut être posée avec certitude. Le locuteur avait peut-être commencé à dire « *c'est* moche », terme qu'il a refusé et interrompu. « La référenciation lexicale est donc seulement une question de probabilité, liée à la cohésion sémantique du texte » (Jeanjean, 1984, p. 172). Côté transcripteur, les amorces demeurent donc une interprétation et côté locuteur, qu'elles soient laissées inachevées, complétées ou corrigées, elles ne sont pas à l'abri des effets de « lapsus » et peuvent constituer d'authentiques faits autonymiques au sens où l'entend Authier-Revuz (1995).

En fait, les amorces de mots, à l'oral, soulèvent les mêmes problèmes d'identification que les erreurs de mots qu'il s'agisse de substitution de mots, de morphèmes (temps, négation, genres, etc.) ou de néologismes.

Dans un grand nombre de cas, l'attribution du terme « erreur » souvent employé à propos de ces achoppements trouve sa justification dans les autocorrections et marques de repentirs, d'excuses ou de surprise diverses qui les accompagnent parfois. Il est à noter que, souvent, seules ces remarques explicites de correction signalent les hésitations du locuteur. En l'absence de ces

« marques énonciatives ou métalinguistiques », ces énoncés ne seraient pas signalés comme comportant une irrégularité (corpus Pallaud, 94, 42) :

– elle veut *se lever* elle veut *sortir* bon en voilà encore un de lapsus ;

– le catéchisme nous a appris que si on a un pouvoir sur les *hommes* on ne doit pas en abuser oui enfin il s'agit d'un pouvoir sur les *animaux* + *vous voyez que je me trompe.*

Au stade de la transcription il n'est pas nécessaire de s'assurer que l'erreur constatée relève d'une mauvaise articulation (échanges ou télescopages de phonèmes), d'un défaut de savoir, ou encore d'hésitations « cognitives ». Pourtant, c'est au niveau de l'explication de l'erreur que les théories s'affrontent.

La description la plus courante, adoptée par certains linguistes, met en cause l'élaboration de l'énoncé sur les axes syntagmatique et paradigmatique. C'est le cas, par exemple, des études récentes sur les lapsus (Rossi et Peter-Defage, 1998, et Arnaud, 1997) : « Selon le critère contextuel, chaque type de lapsus est soit syntagmatique, si l'origine est présente dans le contexte de l'énoncé..., soit paradigmatique, si l'origine est absente de ce contexte et si l'erreur est provoquée par un phénomène associatif quelconque » (Rossi et Peter-Defare, p. 42).

Erreur par anticipation d'un phonème :

*mais c'était quoi un /mate*leau, *mate*lot/ *à* eau *[rires] un matelas à eau* (Pallaud 129)

Monsieur Toubon vous mentez *vous euh vantez les* mérites *de la concertation au sein des partis de droite* (Corpus Pallaud, 200).

Erreur par persévération d'un phonème :

Uderzo veut récupérer les droits Uderzo et [gosino] euh Goscinny *(rires) chez Dargaud* (Pallaud, 77).

Deux sortes de lapsus sont ainsi distingués selon que leur origine est attribuée à l'élaboration de la parole sur l'axe syntagmatique ou sur l'axe paradigmatique. Il est à noter que l'identification des lapsus dits « d'origine paradigmatique » n'est pas du même ordre que les autres types de lapsus. En effet, la catégorie « paradigmatique ou non contextuelle » n'est établie que sur des critères « négatifs » : le lapsus est déclaré « paradigmatique » lorsqu'il est impossible de trouver dans le contexte antérieur ou ultérieur l'origine de l'altération : exemple n° 55 (Rossi et Peter-Defare, 1998, p. 36) :

il se trouve qu'on (dépasse), qu'on *dépense* énormément de pognon [247].

« *Le lien sémantique entre "dépasse" et "dépense" va de soi car on dépasse son budget à "dépenser" trop d'argent, quant à la parenté phonologique elle est évidente. (Ce lapsus) est une substitution de mots qu'*aucune origine phonologique dans le contexte *ne peut expliquer ici.* »

Ce type de lapsus a ainsi un statut de « reste ». Le raisonnement est, donc, quelque peu circulaire puisque les auteurs posent en prémisse que les erreurs pouvant se décrire par des déplacements sur l'axe syntagmatique s'expliquent par ces perturbations sur l'axe syntagmatique et n'ont rien à voir avec une élaboration sur l'axe paradigmatique. L'insistance est forte sur cette hypothèse d'une origine contextuelle du lapsus. Elle se retrouve même dans les recommandations pour le recueil des données. « Nous n'avons retenu que les lapsus dont nous étions sûrs et pour lesquels nous pouvions obtenir un contexte suffisant où pouvait être identifiée l'origine éventuelle » (Rossi et Peter-Defare, 1998, p. 17).

Que les contraintes de la langue pèsent sur la formation des erreurs de langage n'évacue pas pour autant l'intervention d'une pensée inconsciente. C'est ce qu'a toujours soutenu Freud (1901) lorsqu'il s'est intéressé entre autres aux lapsus et en a débattu vigoureusement avec Meringer et Mayer (respectivement linguiste et psychiatre). Il se hasarde à transcrire les lapsus de ses patients, non à l'aide des hypothèses ou idées qui lui viennent à l'esprit lorsqu'il les écoute, mais au seul appui des libres associations de ces patients. Par exemple, Freud rapporte la substitution suivante faite par un patient :

« Ja das *draut... dauert* vielleicht noch einen Monat ». Le mot « draut » avec un *r* me paraît inexplicable, la lettre *r* du mot correct « dauert » n'ayant pas pu produire un effet pareil. J'attire sur ce fait l'attention de R. v. S. qui m'explique aussitôt qu'en parlant il pensait : « das ist eine *traurig* Geschichte ». Il avait donc pensé à deux réponses qui se sont fondues en une seule par l'intermédiaire de deux mots (*draut* provenant de la fusion de *dauert* et de *traurig*) (p. 66).

Freud propose un cadre à son écoute et, donc, à la transcription de ce qu'il entend. Il s'agit du contexte des libres associations du patient. Ce contexte a les dimensions de tout ce qui est prononcé par le patient. Pour le linguiste, le cadre est à la fois plus restreint et plus illimité. Plus restreint car nous avons vu que les solutions de transcription tendent à respecter une cohérence textuelle limitée le plus souvent à l'entourage immédiat du lapsus. Plus illimité puisque cette cohérence est affaire d'appréciation individuelle propre à chaque transcripteur. La transcription reste donc ainsi liée aux associations d'idées du transcripteur qui, elles, peuvent varier à chaque relecture. « Le texte authentique est un mythe rassurant, il n'est pas vérifiable » (Blanche-Benveniste et Jeanjean, 1987, p. 112).

Le linguiste ne peut distinguer parmi les nombreuses erreurs de langage que tout locuteur produit en parlant ce qui serait lapsus ou non, sans quitter

le terrain linguistique. Inversement, tous les achoppements interpellent-ils le psychanalyste ? Il est évident qu'ils ne donnent pas tous matière à écoute et donc à « transcription ». L'achoppement qui reçoit le qualificatif de lapsus est donc bien pris dans un autre discours théorique.

BIBLIOGRAPHIE

ARNAUD, P.J.L. 1997. « Les ratées de la dénomination. Typologies dans les lapsus... » dans Boysson et Thoinon (sous la direction de), *La dénomination*, Lyon, PUL, 307 p.

AUTHIER-REVUZ, J. 1995. *Ces mots qui ne vont pas de soi. Boucles réflexives et non coïncidences du dire*, Paris, Larousse.

BERGÈS, J. 1991. « Face aux limites de la psychanalyse », dans *L'analyse et l'analyste. 1er colloque interassociatif de psychanalyse*, Ed. Solin, Parsi, Arles, p. 31-35.

BLANCHE-BENVENISTE, C. ; JEANJEAN, C. 1987. *Le français parlé. Transcription et édition*, Didier Erudition, Paris.

BLANCHE-BENVENISTE, C. ; BILGER, M. ; ROUGET, C. ; VAN DEN EYNDE, K. 1990. *Le français parlé. Études grammaticales*, Paris, Éditions du CNRS.

CARTER-THOMAS, S. 2000, *La cohérence textuelle. Pour une nouvelle pédagogie de l'écrit*, Paris, L'Harmattan, coll. « Langue et parole ».

CHRISTOPHE, A. ; DUPOUX, E. ; BERTONCINI, J. ; MEHLER, J. 1995. « Do infants perceive word boundaries ? An empirical study of the bootstrapping of lexical acquisition », *J. of the Acoustical Society of America, 3*, 1570-1580.

DOLTO, F. 1981. *Au jeu du désir*, Paris, Le Seuil.

EPL 1996. *Le transfert dans tous ses errata* suivi de *Pour une transcription critique des séminaires de Jacques Lacan*, Édition École de psychanalyse lacanienne, Paris.

FRANCARD, M. ; PÉRONNET, L. 1989. « La transcription de corpus oraux dans une perspective comparative. La démarche du Projet PLURAL », dans Ouellon, C. (sous la direction de) *La description des langues naturelles en vue d'applications informatiques*, Centre international de recherche sur le bilinguisme (publication K-10), Québec, 295-307.

FREUD, S. 1901. *Psychopathologies de la vie quotidienne*, Paris, Payot.

GADET, F. 1989. *Le français ordinaire*, Paris, A. Colin.

JEANJEAN, C. 1984. « Les ratés c'est fa- fabuleux », Étude syntaxique des amorces et des répétitions, *Linx, 10, Syntaxe et discours*, p. 171-177.

LACAN, J. 1968. *D'un autre à l'Autre*, texte polycopié d'une transcription.

MACWHINNEY, B. 1991. *The CHILDES project : tools for analyzing talk*, L ; Erlabaum Associates, Hillsdale USA.

MÈGE, J. ; PALLAUD, B. 2002. Textes et scribes dans la cure de quatre enfants psychotiques, dans Actes du colloque : *L'être en question*, Journées de L'AFORESH, Cannes, 13-14 juin 1998.

MOUCHON, J. ; FILLOL, F. 1980. *Pour enseigner l'oral*, Paris, Cédic.

PALLAUD, B. 2002. « Erreurs d'écoute dans la transcription de données orales », Actes du colloque à Tours, 8-9 décembre 2000, *Transciption de la parole normale et pathologique. Parole.*

ROSSI, M. ; PETER DEFARE, E. 1998. *Les lapsus ou comment notre fourche a langué*, Paris, Ed. PUF.

SABIO, F. 2000. « Les difficultés de la notion de *mot* : l'exemple des liaisons graphiques dans les textes d'enfants », *Linx, 42*, p. 119-130.

SEARLE, J.R. 1972. *Les actes de langage. Essai de philosophie du langage*, Paris, Herman.

THIBAULT, P. ; VINCENT, D. 1990. *Un corpus de français parlé. Montréal 84 : historique, méthodes et perspectives de recherche*, Recherches sociolinguistiques/1, Bibliothèque nationale du Québec.

VION, R. 1992. *La communication verbale. Analyse des interactions*, Paris, Hachette.

Résumé

L'objet de cet exposé est de développer quelques réflexions sur la pratique de transcription d'énoncés enregistrés. Elles s'appuient sur la pratique de transcription notamment au sein du Groupe aixois de recherche en syntaxe et des remarques critiques publiées par l'École de psychanalyse lacanienne à l'occasion d'une transcription critique d'un séminaire de Lacan, « Le transfert ». Que ce soit pour le psychanalyste ou pour le linguiste, la transcription orthographique repose sur une analyse de discours. Dans le cas d'établissement de corpus, cette analyse de discours reste implicite et présente des particularités qui lui sont propres mais peuvent néanmoins être confrontées à l'écoute du psychanalyste.

Mots clefs

Transcription, corpus oraux, écoute psychanalytique, lapsus.

TRIUMPHS AND TRIALS OF TRANSCRIPTION

Summary

Some considerations are developed on the practice of transcription of recorded interview. They lean on the practice of transcription notably within the Groupe Aixois de Recherche en Syntaxe and the critical remarks published by the Ecole de Psychanalyse Lacanienne which took in charge a critical transcription of a seminar of Lacan, « Le transfert ».

For the psychoanalyst and even for the linguist, the orthographic transcription is resting on a discourse analysis. In the case of establishment of corpus, this discourse analysis remains implicit and presents peculiarities which are particular to this type of analysis but can be nevertheless confronted with the way the psychoanalyst listen to his patient.

Key words

Transcription, oral corpora, psychoanalysis' listening, slips of the tongue.