



Chapitre II. Le codage

Laurence Bardin

DANS **QUADRIGE** 2013, PAGES 134 À 149

ÉDITIONS **PRESSES UNIVERSITAIRES DE FRANCE**

ISSN 0291-0489

ISBN 9782130627906

Article disponible en ligne à l'adresse

<https://www.cairn.info/l-analyse-de-contenu--9782130627906-page-134.htm>



CAIRN.INFO
MATIÈRES À RÉFLEXION

Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...

Flashez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour Presses Universitaires de France.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

CHAPITRE II

Le codage

Il faut savoir *pourquoi* on analyse, et l'expliciter, pour savoir *comment* analyser. D'où, ainsi que nous venons de le voir, la nécessité de préciser des hypothèses, de replacer la technique dans un cadre théorique. A moins qu'on ne fasse des *fishing expeditions* comme disent les Anglo-Saxons, c'est-à-dire des analyses exploratoires, « pour voir ». Dans ce cas, c'est le comment (la technique) qui pourra préciser le pourquoi (la théorie). Mais dans les deux cas il y a un lien entre les données du « texte » et la théorie de l'analyste.

Traiter le matériel, c'est le coder. Le *codage* correspond à une transformation — effectuée selon des règles précises — des données brutes du texte. Transformation qui, par découpage, agrégation et dénombrement, permet d'aboutir à une représentation du contenu, ou de son expression, susceptible d'éclairer l'analyste sur des caractéristiques du texte qui peuvent servir d'indices ou comme le dit O. R. Holsti¹ :

« Le codage est le processus par lequel les données brutes sont transformées systématiquement et agrégées dans des unités qui permettent une description précise des caractéristiques pertinentes du contenu. »

1. O. R. HOLSTI, *Content Analysis for the Social Sciences and Humanities*, Addison Wesley Publishing Company, 1969.

L'organisation du codage comprend trois choix (dans le cas d'une analyse quantitative et catégorielle) :

- le découpage : choix des unités ;
- l'énumération : choix des règles de comptage ;
- la classification et l'agrégation : choix des catégories.

I. UNITÉS D'ENREGISTREMENT ET DE CONTEXTE

Quels éléments du texte prendre en compte ? Comment découper le texte en éléments finis ? Le choix des unités d'enregistrement et de contexte doit répondre de façon pertinente (pertinence par rapport aux caractéristiques du matériel, pertinence vis-à-vis des visées de l'analyse).

a) *L'unité d'enregistrement.* — C'est l'unité de signification à coder. Elle correspond au segment de contenu à considérer comme unité de base en vue de la catégorisation et du comptage fréquentiel. L'unité d'enregistrement peut être de nature et de taille très variables. Une certaine ambiguïté règne quant aux critères de distinction des unités d'enregistrement. En effet, certains découpages se font à un niveau sémantique, par exemple, le « thème », d'autres, à un niveau apparemment linguistique, par exemple le « mot » ou la « phrase »¹. Ceci sert de critique à des disciplines dont le caractère scientifique et rigoureux est plus évident. En fait le critère de découpage en analyse de contenu² est toujours d'ordre sémantique, bien qu'une correspondance avec des

1. Mais qu'est-ce qu'une phrase ? Quels sont les critères de définition à retenir ? La phrase est-elle l'unité de sens exprimant une pensée complète ? Est-elle une proposition logique : sujet (ce dont on dit quelque chose) plus prédicat (ce qu'on en dit) ? Ou bien est-ce le critère phonétique qui la définit : arrêts, silence, rupture dans la courbe mélodique (ou les équivalents scripturaux : points, points-virgules) ? (G. MOUNIN, *Clefs pour la linguistique*, Seghers).

2. En analyse du contenu *i.e.* des significations et non en analyse de l'expression, *i.e.* des aspects formels des significations.

unités formelles existe parfois (exemples : mot et mot thème, phrase et unité signifiante).

On peut citer, à titre illustratif, parmi les unités d'enregistrement les plus employées :

- *Le mot* : Certes, le « mot » n'a pas de définition précise en linguistique. Mais pour les usagers de la langue, il correspond à quelque chose. Une précision linguistique peut être apportée si elle est pertinente.

On peut prendre en compte tous les « mots » du texte ou ne retenir que les mots clefs ou mots thèmes (*symbols* en anglais) ; faire la distinction entre mots pleins et mots vides ; effectuer l'analyse sur une catégorie de mots : substantifs, adjectifs, verbes, adverbess... Dans le but d'établir des quotients.

- *Le thème* : La notion de thème, largement utilisée en analyse thématique, est propre à l'analyse de contenu. Berelson définissait le thème comme :

« une affirmation sur un sujet. C'est-à-dire une phrase, ou une phrase composée, habituellement un résumé ou une phrase condensée, sous laquelle un vaste ensemble de formulations singulières peuvent être affectées ».

En fait, le thème est l'unité de signification qui se dégage naturellement d'un texte analysé selon certains critères relatifs à la théorie qui guide la lecture. Le texte peut être découpé en idées constituantes, en énoncés et propositions porteurs de significations isolables. Le thème est

« une unité de signification complexe, de longueur variable ; sa réalité n'est pas d'ordre linguistique mais d'ordre psychologique : une affirmation mais aussi une allusion peuvent constituer un thème ; inversement, un thème peut être développé en plusieurs affirmations (ou propositions). Enfin, un fragment quelconque peut renvoyer (et renvoie généralement) à plusieurs thèmes... » écrit M. C. d'Unrug¹.

1. M. C. d'UNRUG, *Analyse de contenu et acte de parole*, Ed. Universitaires, 1974.

Faire une analyse thématique consiste à repérer des « noyaux de sens » qui composent la communication et dont la présence ou la fréquence d'apparition pourront signifier quelque chose pour l'objectif analytique choisi.

Le thème comme unité d'enregistrement correspond à une règle de découpage (du sens, non de la forme) qui n'est pas donnée une fois pour toutes, puisque le découpage dépend du niveau d'analyse et non de manifestations formelles réglées. Il ne peut y avoir de définition de l'unité thématique comme il y a définition des unités linguistiques.

Le thème est utilisé généralement comme unité d'enregistrement pour des études de motivations, d'opinions, d'attitudes, de valeurs, de croyances, de tendances, etc. Les réponses à des questions ouvertes, les entretiens (entretiens non directifs, ou plus structurés) individuels ou de groupe, d'enquête ou de psychothérapie, les protocoles de tests, les réunions de groupe, les psychodrames, les communications de masse, etc., peuvent et souvent sont analysés sur la base du thème.

Notons qu'une préparation des messages en unités linguistiques (énoncés, propositions, syntagmes) normalisées peut être utile dans certains cas.

- *L'objet ou référent* : Il s'agit de thèmes pivots autour desquels s'organise le discours. Par exemple, les pièces de la maison citées dans une enquête sur l'habitat. Ou les « objets d'attitudes » (cf. l'analyse évaluative d'Osgood), dans une analyse de la presse politique. Dans ce cas, on découpe le texte en fonction de ces thèmes pivots et on groupe autour d'eux tout ce que le locuteur exprime à leur sujet.

- *Le personnage* : L'acteur ou actant peut être choisi comme unité d'enregistrement. Dans ce cas, le codeur repère les « personnages » (être humain ou équivalent tel que animal, etc.) et dans le cas d'une analyse catégorielle les classe en fonction de la grille choisie. Cette grille est généralement établie en fonction des caractéristiques ou attributs

du personnage (traits de caractère, rôle, statut social, familial, âge, etc.). Les œuvres de fiction (films, émissions, roman, bandes dessinées, romans-photos, pièces de théâtre) peuvent être analysées selon leurs personnages. De même que les articles de presse, des manuels scolaires, etc.

Qui, à quel moment ? Avec quel rôle ? Dans quelle situation ? etc. L'unité « personnage » peut être combinée avec d'autres types d'unités.

- *L'événement* : Dans le cas de récits, de narration, il est possible que l'unité d'enregistrement pertinente soit l'événement. Dans ce cas le ou les récits (films, légendes, contes, récits mythiques, articles de presse) seront découpés en unités d'action.

- *Le document* : Le document ou unité du genre (un film, un article, une émission, un livre, un récit) sert parfois d'unité d'enregistrement lorsqu'il peut être caractérisé globalement et dans le cas d'analyse rapide. Il est aussi possible de prendre comme unité d'enregistrement la réponse (à une question ouverte) ou l'entretien, à condition que l'idée dominante ou principale soit suffisante pour la visée recherchée.

En fait, l'unité d'enregistrement est à la croisée d'unités perceptives (mot, phrase, document matériel, personnage physique) et d'unités sémantiques (thèmes, événements, individus), mais il semble difficile, même s'il y a parfois recouvrement, de chercher à faire un découpage de nature purement formelle, dans la majorité des pratiques, du moins pour l'analyse thématique, catégorielle et fréquentielle (celle qui nous sert de base pour ce chapitre).

b) *L'unité de contexte*. — Elle sert d'unité de compréhension pour coder l'unité d'enregistrement. Elle correspond au segment du message dont la taille (supérieure à l'unité d'enregistrement) est optimale pour saisir la signification exacte de l'unité d'enregistrement. Cela peut, par exemple, être la phrase pour le mot, le paragraphe pour le thème.

En effet, dans bien des cas, il est nécessaire de se référer — consciemment — au contexte proche ou lointain de l'unité à enregistrer. Si plusieurs codeurs fonctionnent sur le même corpus, un accord préalable est nécessaire. Par exemple, dans le cas d'analyse de messages politiques, des mots tels que liberté, ordre, progrès, démocratie, société ont besoin du contexte pour être compris à leur juste sens. La référence au contexte est très importante pour l'analyse évaluative et l'analyse de contingence. Selon la taille de l'unité de contexte les résultats sont susceptibles de varier sensiblement. L'intensité et l'étendue d'une attitude peuvent apparaître de manière plus ou moins accentuée selon la taille de l'unité de contexte choisie. Et en matière de co-occurrences, il est évident que le nombre de co-occurrences augmente avec la taille de l'unité de contexte : il est peu probable, par exemple, que l'on puisse trouver des thèmes semblables dans un paragraphe ou quelques minutes d'enregistrement, mais la probabilité augmente dans un texte de plusieurs pages ou dans une émission d'une heure. En général, plus l'unité de contexte est grande, plus les attitudes ou valeurs s'affirment dans une analyse évaluative et plus les co-occurrences sont nombreuses dans une analyse de contingence.

Deux critères président à la détermination de la taille de l'unité de contexte : le coût et la pertinence. Il est évident qu'une unité de contexte large exige une re-lecture de l'environnement plus longue. D'autre part, il existe un optimum au niveau du sens : trop petite ou trop grande, l'unité de contexte n'est plus adaptée, là encore, le type de matériel et le cadre théorique sont déterminants.

De toute façon, il est possible de tester les unités d'enregistrement et de contexte sur de petits échantillons afin d'être sûr d'opérer avec les instruments les plus adéquats.

2. RÈGLES D'ÉNUMÉRATION

Il faut distinguer entre unité d'enregistrement — ce que l'on compte — et règle d'énumération — la manière de compter.

Prenons un exemple. Nous avons un « texte » fini dont l'identification et le découpage ont livré les éléments ou unités d'enregistrement (mots, ou thèmes ou autres unités) suivants :

a, d, a, e, a, b.

Etant donné que la liste de référence, établie sur un ensemble de « textes » ou à partir d'une norme, est *a, b, c, d, e, f*, nous pourrions utiliser divers types d'énumération :

- *La présence (ou l'absence)* : Sont présents les éléments *a, b, d*, et *e* dans ce « texte » précis. Cette présence peut être significative. Elle fonctionne alors comme un indicateur.

Mais l'absence d'éléments (par rapport à une certaine attente) peut, dans certains cas, véhiculer du sens. Ici les éléments *c* et *f* sont absents. En effet, pour certains types de messages ou certains objectifs d'analyse, l'absence est la variable qui importe. Par exemple, l'absence peut manifester des blocages ou refoulements dans des entretiens cliniques, elle peut traduire une volonté cachée dans le cas d'une déclaration publique.

- *La fréquence* : C'est la mesure la plus généralement utilisée. Elle correspond au postulat (valable dans certains cas, non valable dans d'autres) suivant : l'importance d'une unité d'enregistrement croît avec sa fréquence d'apparition.

Dans notre exemple, la fréquence de chaque élément est :

a = 3

b = 1

c = 0

d = 1

e = 1

f = 0.

Une mesure fréquentielle où chaque apparition est affectée du même poids postule que tous les éléments ont une importance égale. Le choix de la mesure fréquentielle simple ne doit pas être automatique. Il faut se souvenir qu'elle repose sur le présupposé implicite suivant : l'apparition d'un item de sens ou d'expression est d'autant plus significative — par rapport à ce qu'on cherche à atteindre dans la description ou l'interprétation de la réalité visée — que cette apparition est répétée avec une plus grande fréquence. C'est donc la régularité quantitative d'apparition qui est considérée comme significative. Cela suppose que chaque item a la même valeur. Ce qui n'est pas toujours le cas.

• *La fréquence pondérée* : Si l'on suppose que l'apparition de tel élément a plus d'importance que tel autre, on peut avoir recours à un système de pondération. Par exemple, si l'on considère que l'apparition de *b* et *d* a un poids double de celle de *a*, *c*, *e* et *f*, on affectera chaque élément de coefficients au moment du codage.

Soit la pondération suivante :

$$a = 1$$

$$b = 2$$

$$c = 1$$

$$d = 2$$

$$e = 1$$

$$f = 1.$$

Cela donne les résultats suivants :

$$a = 3 \times 1 = 3$$

$$b = 1 \times 2 = 2$$

$$c = 0 \times 1 = 0$$

$$d = 1 \times 2 = 2$$

$$e = 1 \times 1 = 1$$

$$f = 0 \times 1 = 0.$$

Nous obtenons donc des résultats différents de ceux obtenus par la mesure à fréquence non pondérée.

La pondération peut correspondre à une décision prise *a priori*, mais elle peut aussi traduire les modalités d'expression ou l'intensité d'un élément.

• *L'intensité* : Soit dans notre exemple trois degrés (correspondant à des variations sémantiques ou formelles au sein d'une même classe), dans l'apparition d'un élément :

$$a_1, a_2, a_3 - b_1, b_2, b_3, \text{ etc.}$$

et l'affectation d'une note différente selon la modalité d'expression :

$$\begin{array}{ll} a_1 = 1 & b_1 = 1 \\ a_2 = 2 & b_2 = 2 \\ a_3 = 3 & b_3 = 3, \text{ etc.} \end{array}$$

Soit le « texte » :

$$a_1, d_3, a_3, e_1, a_3, b_1.$$

La mesure sera :

$$\begin{array}{l} a = 7 (1 + 3 + 3) = 7 \\ b = 1 \\ c = 0 \\ d = 3 \\ e = 1 \\ f = 0. \end{array}$$

La mesure de l'intensité avec laquelle apparaît chaque élément est indispensable dans l'analyse des valeurs (idéologies, tendances) et des attitudes. Si nous rencontrons les quatre énoncés suivants dans une étude de la presse chinoise des années 60, il faut pouvoir différencier l'intensité des positions correspondantes :

1. « *Nous pourrions trouver nécessaire de désapprouver la politique de Khrouchtchev.* »

2. « *Nous devrions dénoncer amèrement la politique de Khrouchtchev.* »

3. « *Nous allons bientôt commencer à dénoncer la politique de Khrouchtchev.* »

4. « *Nous avons quelquefois été en désaccord avec la politique de Khrouchtchev dans le passé* »¹.

Pour faciliter le jugement du degré d'intensité à coder, on peut s'appuyer, comme le suggérait Osgood, sur des critères précis : intensité (sémantique) du verbe, temps du verbe (conditionnel, futur, impératif...), adverbes de modalités, adjectifs et attributs qualificatifs...

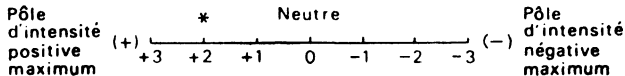
• *La direction* : La pondération de la fréquence traduit un caractère quantitatif (intensité) ou qualitatif : la *direction*. La direction peut être favorable, défavorable ou neutre (éventuellement ambivalente) dans le cas d'une étude de favorabilité/défavorabilité. Mais les pôles directionnels peuvent être de nature diverse : beau/laid (critère esthétique), petit/grand (taille), etc.

On affecte les éléments du texte d'un signe (indice qualitatif) ou d'une note.

$$a_+, d_0, a_+, e_-, a_+, b_+$$

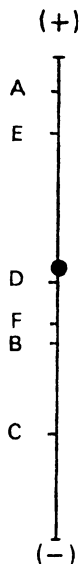
(par exemple : + = positif, — = négatif, 0 = neutre, ± = ambivalent).

Osgood a, dans l'analyse évaluative, eu recours à des *échelles bipolaires* pour coder l'*intensité* et la *direction*. La pondération des mesures fréquentielles l'a amené ensuite à la représentation des résultats sous forme de *profils*.



Echelle bipolaire à sept points (ou degrés), pour un exemple où *a* est de direction positive et d'intensité 2 (a_{+2}). Indiqué par le signe *.

1. Exemple donné par O. R. Holsti.



Profil. Le profil traduit l'ensemble des fréquences pour chaque élément. Par exemple, ici, A (ensemble des unités enregistrées pour un corpus donné) et E sont très positifs, D neutre, F et B légèrement négatifs et C assez négatif.

- *L'ordre* : L'ordre d'apparition des unités d'enregistrement (par exemple dans un entretien ou un récit) peut être l'indice pertinent. Si *a* arrive au premier rang et *d* au second, si *a* précède *d*, ou *d* succède à *a*, cela peut avoir une signification plus importante (dans le cadre de l'inférence) que la fréquence. Ou bien fréquence et enchaînement des unités d'enregistrement peuvent se combiner dans les mesures. Si des constantes dans l'ordre de succession des éléments sont mises à jour (par exemple si $a \rightarrow d \rightarrow a$ revient avec une fréquence significative), cela peut être utile à savoir.

- *La co-occurrence* : La co-occurrence est la présence simultanée de deux ou plusieurs unités d'enregistrement dans une unité de contexte.

Il y a deux possibilités pour décider de l'unité de contexte :

- On choisit celle-ci en nombre d'unités d'enregistrement : par exemple en découpant le texte par 3 (ou 4, ou 5, ou n'importe quelle autre taille) unités :

soit le « texte »	$a, d, a, c, a, b, e, e, e, \dots$
après découpage	$a, d, a/c, a, b/e, e, e/\dots$

Résultats : 2 éléments a dans la première unité
 3 éléments e dans la troisième unité

- On décide du nombre d'unités d'enregistrement antérieures et/ou postérieures à une unité donnée servant de pivot : par exemple, un mot précédent et un mot suivant chaque substantif, ou deux thèmes antérieurs à un thème choisi, etc.

La mesure de co-occurrence (analyse de contingence) rend compte de la *distribution* des éléments, et de l'*association* de ceux-ci. La distribution des éléments peut être un point de connaissance significatif. Par exemple, deux « textes » présentent le même nombre d'éléments a , mais dans le premier, ils sont dispersés dans tout le texte et dans le second, massés dans un certain passage.

L'usage de l'association comme indicateur repose généralement sur le postulat que des éléments associés dans une manifestation langagière sont (ou seront) aussi associés dans l'esprit du locuteur (ou du destinataire). Des modalités qualitatives différencient éventuellement la nature de la co-occurrence :

Association (l'élément a apparaît avec l'élément b).

Equivalence (l'élément a ou l'élément d apparaissent dans un contexte identique. On peut, peut-être, déduire un caractère d'équivalence ou de substitution).

Opposition ou exclusion (l'élément a n'apparaît jamais avec l'élément c).

Signalons enfin que la proximité d'occurrence peut être mesurée : si a est à trois unités d'enregistrement ou deux minutes de distance de b , cela n'a pas la même importance peut-être que s'il est à sept unités d'enregistrement ou quatre minutes et demie de b .

Remarques. — Tout choix d'une règle (ou de plusieurs règles) d'énumération repose sur une hypothèse de correspondance entre la présence, la fréquence, l'intensité, la distribution, l'association de la manifestation langagière et

la présence, la fréquence, l'intensité, la distribution, l'association de variables non langagières inférées. Il convient de chercher la correspondance la plus pertinente.

— Une variable d'inférence peut parfois se manifester de diverses façons. Il est possible de l'atteindre par le biais d'indices différents ou complémentaires. Par exemple, en analyse de presse, la surface des articles, la taille des titres ou la fréquence des événements décrits sont peut-être trois moyens de codage et d'énumération aptes à rendre compte de la même réalité.

— On a essayé d'utiliser des systèmes d'énumération applicables à un matériel continu (mesures d'espace, de temps) ou gradué (mesures de couleur). Sauf cas particuliers, la précision de la mesure est plus apparente que réelle. Le comptage d'une unité d'enregistrement par minute ou par centimètre carré est peut-être encore plus artificiel que le découpage d'un texte par phrases ou par paragraphes. Néanmoins, il arrive que des mesures de ce type soient adaptées ou seules possibles. S'il est démontré par des expériences antérieures que l'analyse des slogans publicitaires permet d'arriver aux mêmes résultats que l'analyse du texte correspondant, à condition de pondérer ces slogans en fonction de leur taille, il sera plus rapide d'utiliser la première mesure. S'il est confirmé, par comparaison avec les résultats obtenus par d'autres tests de personnalité, que le découpage en carrés de la surface d'un test de village est une méthode précise, il convient de l'employer. Si l'analyse, par séquences minutées, d'une émission de type narratif donne de bons résultats en fonction de l'objectif, il faut la pratiquer.

3. ANALYSE QUANTITATIVE ET ANALYSE QUALITATIVE

Procédures « quantitatives » et procédures « qualitatives » ont alimenté un débat passionné aux Etats-Unis dans les années 50. Certains plaçaient le caractère quantitatif dans leurs définitions de l'analyse, d'autres défendaient la validité d'une analyse « qualitative ».

A. L. George¹, lors du premier congrès réunissant les analystes, essaya de préciser les caractéristiques de l'une et l'autre méthode. Dans la mesure où « l'analyse de contenu est employée comme un instrument de diagnostic pour faire des inférences spécifiques ou des interprétations causales à propos d'un aspect quelconque de l'orientation comportementale du locuteur », elle n'est pas obligatoirement quantitative dans sa démarche comme on le soutenait jusqu'alors (sous l'influence de Berelson notamment).

L'approche quantitative est fondée sur la *fréquence* d'apparition de certains éléments du message. L'approche non quantitative a recours à des indicateurs non fréquentiels susceptibles d'autoriser des inférences ; par exemple la *présence (ou l'absence)* peut être un indice aussi (ou plus) fructueux que la fréquence d'apparition.

Quelle est l'évolution de la fréquence d'apparition du mot « *patrie* » dans les manuels d'histoire depuis une cinquantaine d'années ? Le mot *patrie* est-il présent ou absent des manuels d'histoire en 1975 ? Dans un contexte donné, par exemple les discours d'un homme politique, l'apparition d'un mot non attendu ou propre à l'opposition, une phrase plus tempérée ou plus restrictive que les propositions habituelles sur la question, peuvent fonctionner comme indice de poids, si on ne les noie pas dans un décompte fréquentiel.

Approche quantitative et qualitative n'ont pas le même terrain d'action. La première obtient des données descriptives par une méthode statistique. Elle semble, grâce au décompte systématique, plus précise, plus objective, plus fiable et fidèle, car l'observation y est davantage contrôlée. Rigide cependant, elle est utile dans les phases de vérification des hypothèses. La seconde correspond à une procédure plus intuitive mais aussi plus souple, plus adaptable à des indices non prévus ou à l'évolution des hypothèses. Elle est à employer lors des phases de formation des hypothèses. Elle permet de suggérer des relations possibles entre un indice du message et une ou plusieurs variables du locuteur (ou de la situation de communication).

L'analyse qualitative présente certaines caractéristiques particulières. Elle est surtout valable pour faire des déductions spécifiques à propos d'un événement, d'une variable d'inférence précise et non pour des inférences générales. Elle peut fonctionner sur des corpus réduits et établir des catégories plus discriminantes puisqu'elle n'est pas liée comme l'analyse quantitative à des catégories donnant lieu à des fréquences suffisamment élevées pour que les calculs soient possibles. Elle pose des problèmes au niveau

1. A. L. GEORGE, Quantitative and qualitative approaches to content analysis, in I. de SOLA POOL, *Trends in content Analysis*, 1959.

de la pertinence des indices retenus puisqu'elle sélectionne ces indices sans traiter exhaustivement tout le contenu. Il y a des risques de laisser des éléments importants ou de prendre en compte des éléments non significatifs. La compréhension exacte du sens est capitale dans ce cas. De plus, jouant sur des éléments isolés ou des fréquences faibles, elle voit s'accroître le risque d'erreur. D'où l'importance du contexte. Contexte du message mais aussi contexte extérieur à celui-ci : quelles sont les conditions de production, autrement dit, qui parle à qui, dans quelles circonstances, quel est le moment et le lieu de la communication, quels sont les événements antérieurs ou parallèles ? D'autre part, plus que l'approche quantitative, fixée, l'approche qualitative, évolutive est confrontée avec le risque de « circularité ». Les hypothèses formulées au début peuvent être influencées en cours de route par ce que l'analyste comprend de la signification du message. Là, plus qu'ailleurs, il faut relire le matériel, faire alterner relecture et interprétations, se méfier de l'évidence (n'y a-t-il pas une « évidence » contraire ?) fonctionnant par approximations successives. Souple dans sa démarche, l'analyse qualitative doit être souple dans l'usage de ses indices. Particulièrement en propagande ou en psychothérapie où les conditions de production changent parfois brusquement, les manifestations par la communication d'une même réalité pouvant se modifier rapidement. Autrement dit, les indices sont instables et une résistance au changement de la part de l'analyste est d'autant plus néfaste que la procédure qualitative fonde son interprétation sur peu d'éléments.

Enfin, précisons que l'analyse qualitative ne rejette pas toute forme de quantification. Ce sont les indices qui sont retenus de manière non fréquentielle. Mais on peut faire appel à des tests quantitatifs : par exemple, l'apparition d'indices similaires dans des discours semblables.

En conclusion, on peut dire que ce qui caractérise l'analyse qualitative est que « l'inférence — chaque fois qu'elle est faite — est fondée sur la présence de l'indice (thème, mot, personnage, etc.), non sur la fréquence de son apparition, dans chaque communication individuelle ».

La discussion approche quantitative/approche qualitative a marqué un tournant dans la conception de l'analyse de contenu. Dans la première moitié du xx^e siècle, ce qui marquait la spécificité de ce type d'analyse était la *rigueur*, donc la *quantification*. Par la suite, on a compris que le propre de l'analyse de contenu est l'*inférence* (variables inférées à partir de variables d'inférence au niveau du message), que les modalités d'inférence soient fondées sur des indicateurs quantitatifs ou non.

Il est évident que la nature du matériel joue dans le choix du

type de mesure. On peut, par exemple, faire la distinction entre des messages normalisés et des messages singuliers. Les premiers correspondent à un corpus constitué de messages provenant de locuteurs différents. Par exemple : réponse à des questions ouvertes organisées en prévision du codage, avec tout ce que cela implique de standardisation, de nivellement, de conformation ; dans ce cas, le type d'investigation prépare et oriente un type d'analyse fondé sur la quantification dans une situation normalisée. Les seconds sont des messages provenant d'un seul émetteur ou de plusieurs, mais irréductibles à la normalisation (singularité de l'expression, de la situation dans les conditions de production, de la finalité dans l'objectif de communication). C'est par exemple le cas d'un entretien non directif se présentant comme un tout, un système structuré selon des lois qui lui sont propres, donc analysable en lui-même ou incomparable.

Il est nécessaire parfois de se détacher de la toute-croyance sociologique en la signification de la régularité. L'événement, l'accident, le rare ont quelquefois un sens très fort qu'il ne faut pas étouffer.