

# Savoir compter, savoir conter

## Épisode n° 1 – Des chiffres : point trop n'en faut

 Daniel Temam\*

Les textes que publient les statisticiens contiennent des chiffres : quoi de plus normal, ces textes sont destinés à commenter les données issues de leurs enquêtes et de leurs autres opérations statistiques. Mais il arrive, bien souvent, que les chiffres soient trop nombreux. Dans le pire des cas, les lecteurs se voient proposer des phrases qui ne font que répéter le contenu de tableaux : *Le nombre des salariés concernés a progressé de 2,4 % en 2004, après 6,8 % en 2003, 1,7 % en 2002 et 3,3 % en 2001. Ou encore : En 2004, la progression a été de 8 % dans les hôtels-restaurants, 6,4 % dans le commerce, et 5,5 % dans les services rendus aux entreprises.* Le premier exemple donne une évolution dans le temps, le deuxième compare des taux de croissance par secteurs d'activité (ils sont l'un et l'autre inventés, mais les exemples réels n'auraient pas manqué). Quand ils rencontrent des phrases de ce genre, les lecteurs sont, très rapidement, vaincus par l'ennui. La conséquence est évidente : ils passent à autre chose. Rien ne les oblige en effet à lire les textes qu'ils ont entre les mains. Tout comme chacun d'entre nous, ils reçoivent chaque jour une foule de journaux, magazines, comptes rendus, notes. Et, tout comme chacun d'entre nous, ils commencent à les parcourir, et s'arrêtent dès qu'ils ne sont plus suffisamment intéressés. Comment donc les intéresser ? En commentant l'information chiffrée, en la mettant en perspective. Ainsi, dans le premier exemple, la phrase donne quatre taux de croissance très différents. Il faut expliquer aux



lecteurs les raisons de ces différences. En particulier, ils aimeraient savoir pourquoi la croissance a été particulièrement rapide en 2003. De même, dans le deuxième exemple, les secteurs cités sont en principe ceux où la progression a été la plus forte. Là encore, les lecteurs souhaitent comprendre *pourquoi* la progression a été particulièrement forte dans ces secteurs. Dans les deux cas, et dans tous les exemples analogues, l'information statistique elle-même figurera dans un tableau. Ou, mieux, elle peut être présentée de manière parlante dans un graphique. Inutile de la répéter dans le texte, *s'il ne s'agit que de la répéter*. Le texte ne doit, à aucun moment, être le double d'un tableau. Ceci ne veut pas dire que les textes ne doivent en contenir, mais en nombre restreint. L'auteur doit sélectionner quelques chiffres sur lesquels il souhaite insister, parce que ce sont les plus importants, ou les plus inattendus. Son rôle d'auteur est de les commenter, c'est-à-dire de faire ressortir ce qu'ils apportent. Une autre façon de voir les choses

est de dire que tous les chiffres cités dans le texte doivent être parlants pour les lecteurs. Un petit tour en Corse fournira cette fois un exemple : *La densité moyenne de la Corse s'élève à 29 habitants au km<sup>2</sup>*. Certes, mais 29 habitants au km<sup>2</sup> est-ce peu ou beaucoup ? Les lecteurs géographes le savent, mais ils sont peu nombreux. L'auteur avait donc pris soin d'expliquer que cette densité était particulièrement faible : *La densité moyenne de la Corse s'élève à 29 habitants au km<sup>2</sup>. Ce chiffre la place en dernière position des régions françaises, et fait d'elle l'île méditerranéenne à la densité la moins élevée. À titre de comparaison, la densité de la population de la Sardaigne est de 68 habitants au km<sup>2</sup>, et celle de la Sicile de 198.*

### Faire appel à d'autres sources

Les commentaires gagnent souvent à faire appel à d'autres sources statistiques que la source principale commentée. Dans certains cas, cet appel est même nécessaire si l'auteur ne veut pas tromper ses lecteurs. Un exemple, volontairement pris hors du champ de l'Insee, est particulièrement démonstratif (même s'il n'est pas très gai). Il s'agit de l'évolution du nombre de tués dans les accidents de la circulation. Ce nombre diminue régulièrement en France depuis plusieurs décennies. En 20 ans, entre 1983 et 2003, il a été divisé par deux. Ce chiffre en lui-même est satisfaisant. Les efforts faits dans ce domaine ont porté leurs fruits : limitation de

vitesse, contrôle de l'alcoolémie, amélioration du réseau routier, contrôle technique des véhicules... Au cours de la même période, entre 1983 et 2003, le trafic a fortement augmenté. La baisse du nombre de tués, rapprochée de l'augmentation du trafic, prend alors davantage d'ampleur. L'amélioration est beaucoup plus spectaculaire. Un auteur commentant l'évolution du nombre de tués minore donc les progrès accomplis s'il ne fait pas appel à l'autre source que constitue l'évolution du trafic. Bien entendu, la comparaison n'est pas immédiate, et doit être affinée. L'auteur peut aussi faire appel à des comparaisons internationales. Le nombre de tués pour un million d'habitants en France en 2002 a été de 129, alors qu'il n'a été par exemple que de 60 au Royaume-Uni, ou de 83 Allemagne. Cette fois, les lecteurs concluront qu'il reste, malgré les progrès accomplis, une marge importante de progression. Malheureusement, les auteurs se contentent bien souvent de « leur » source. L'appel à d'autres sources est bien rare, ce qui limite l'intérêt des commentaires, et peut donner, comme dans le cas des accidents de la circulation, des idées fausses. Cette réticence à utiliser d'autres sources est en partie justifiée : les champs sont souvent différents, de même que les concepts utilisés. De plus, la fiabilité de ces sources n'est pas toujours bien connue. Mais ces obstacles n'ont rien d'insurmontable, pour peu que l'on ait la volonté de les franchir.

### Arrondir

Une autre remarque, moins importante, est que les chiffres qui figurent dans les textes sont souvent trop précis. *En 2003, 5 731 personnes ont été tuées dans un accident de la route en France.* En lisant ceci, les lecteurs peuvent croire que le chiffre est connu à l'unité près, ce qui, pour de nombreuses raisons, n'est pas vrai. Un chiffre trop précis risque ainsi de faire illusion sur la précision des données commentées. A l'inverse, les lecteurs plus au



fait du sujet traité auront l'impression que le statisticien surestime la précision de ses chiffres, ce qui n'est pas bon pour sa crédibilité. En fait, arrondir à 5 700 suffit largement. Les lecteurs saisissent plus facilement l'ordre de grandeur d'un chiffre arrondi, qu'ils sont de toute façon plus en mesure de retenir. Ce sont les tableaux qui donneront les chiffres précis fournis par l'enquête, ne serait-ce que parce qu'il est trop compliqué d'arrondir les chiffres d'un tableau tout en respectant l'égalité des totaux. Certains auteurs craignent que les lecteurs soient choqués de lire 5 700 dans le texte et 5 731 dans un tableau. On peut raisonnablement penser qu'il n'en est rien, dans la mesure où les deux chiffres leur paraîtront cohérents. Les lecteurs comprennent parfaitement que 5 700 est l'arrondi de 5 731, et que le chiffre figurant dans le texte est donc un arrondi du chiffre figurant dans le tableau. La règle à appliquer est simple : on arrondit au chiffre le plus proche ; 5 731 s'arrondit ainsi à 5 700 et non à 5 800, parce qu'il est plus près du premier chiffre que du second, ou bien, autre façon de le dire, parce qu'il est inférieur au milieu, soit 5 750. Les lecteurs ne seront choqués que si les deux chiffres, celui du texte et celui du tableau, ne sont pas cohérents, s'ils lisent par exemple 5 600 dans le texte et 5 731 dans un tableau. Ils se demanderaient alors à juste titre quel est le bon chiffre, sans pouvoir répondre.

### Près de, plus de...

Une nouvelle habitude se développe dans les publications de l'Insee : faire précéder les chiffres arrondis d'un « près de », d'un « plus de » ou, plus rarement, d'un « environ ». Certes, cette pratique traduit la volonté de l'auteur de ne pas écrire simplement « un tiers des salariés a telle ou telle caractéristique » alors que ce n'est pas *exactement* un tiers des salariés qui a cette caractéristique. Ce scrupule, pour louable qu'il soit, n'est pas pour autant justifié. Tout d'abord, les lecteurs ne sont pas des imbéciles : ils savent que, sauf hasard extraordinaire, ce ne peut pas être exactement un tiers des salariés. Ils se doutent donc qu'« un tiers » est un chiffre arrondi. Ajouter « près de », ou « plus de » ne leur apporte finalement aucune information supplémentaire : le « près de » ou le « plus de » donne certes le *sens* de l'écart, mais aucune idée sur son *ampleur*. Ces expressions sont donc inutiles, voire nuisibles. Quelquefois, elles aboutissent à conduire les lecteurs à surestimer la précision des chiffres qu'on leur fournit. En lisant « près d'un tiers », ils penseront que l'auteur est sûr que la proportion est inférieure à un tiers. S'ils lisent ensuite dans un tableau une proportion de 32,7 %, ils pourront légitimement en conclure que le chiffre est connu avec une précision meilleure que 0,6 point (33,3 moins 32,7). Or, bien souvent, c'est loin d'être le cas, en particulier si les résultats proviennent d'une enquête par sondage. À vouloir être trop scrupuleux, on finit par tromper les lecteurs. C'est encore plus flagrant quand on lit, ce qui arrive quelquefois, « près de 63 % ». Cette fois les lecteurs ont immédiatement l'idée que le chiffre est connu à 0,5 point près (puisque seuls les nombres compris entre 62,5 et 63,5 s'arrondissent à 63), ce qui n'est généralement pas le cas. Mais il y a plus. En lisant « près d'un tiers des salariés... », et pas seulement « un tiers des salariés », les lecteurs comprennent que l'auteur veut leur suggérer

qu'un tiers, c'est beaucoup. Ou en tout cas que c'est plus que la proportion qu'ils imaginent sans doute. C'est le sens que véhicule l'expression « près de » dans le langage de tous les jours. Mais ce n'est certainement pas ce que veut dire l'auteur ! L'expression ajoute une signification parasite, qui n'est absolument pas recherchée.

### Pour résumer

Les textes comportant des chiffres trop nombreux découragent les lecteurs les mieux attentionnés. La paraphrase des tableaux est en particulier à bannir. Tous les chiffres figurant dans le texte doivent être parlants pour les lecteurs. Mieux vaut arrondir que de donner des chiffres trop

précis. Les « près de », « plus de » et autres expressions équivalentes sont plus nuisibles qu'utiles.

### Prochain épisode...

Le prochain épisode de la rubrique « Savoir compter, savoir conter » s'efforcera de vous expliquer comment rendre votre style plus vivant.

Daniel Temam, X-Ensaë, est chef de la division *Rédaction des publications* au sein du département de l'offre éditoriale de l'Insee, et en particulier rédacteur en chef d'*Insee Première*. Il a été notamment rédacteur en chef d'*Économie et Statistique*, ainsi que du *Rapport sur les comptes de la Nation*. Il a été également journaliste au mensuel *Alternatives Économiques*, et cofondateur et directeur de la publication du magazine *Tangente, l'aventure mathématique*. Par ailleurs, il est depuis de nombreuses années formateur aux techniques rédactionnelles à l'Insee, dans des services statistiques de ministères, des universités et des grandes écoles.

#### Publications :

Auteur de *La nouvelle comptabilité nationale*, Dunod, 1999.

Coauteur de *Sciences économiques, les grandes notions*, Bordas, 1998.

Membre de l'équipe de coordination des *Polytechniciens dans le siècle*, 1994.

Traducteur de *L'économie mondialisée* de Robert Reich, Dunod, 1993.

Auteur et coauteur de plusieurs livres d'initiation au bridge.