

L'ANALYSE STATISTIQUE COMME DEMARCHE HEURISTIQUE ET
COMME PROCEDURE DE DECISION DANS L'ETUDE DU STYLE

ANDRÉ RAPHAEL

A la fin de son important ouvrage *Shakespeare's Imagery and What It Tells Us*,¹ Caroline Spurgeon donne, pour chacune des pièces de Shakespeare, le nombre d'images et le nombre de vers ou de lignes. Nous rappelons ces données aux colonnes (1) et (2) de notre tableau A. On peut y voir, par exemple, que la pièce qui comporte le plus d'images est *Troilus and Cressida*, et que celle qui en comporte le moins est *The Comedy of Errors*. On ne peut guère aller au-delà de ce genre de constatation. En effet, la colonne (2) est là pour rappeler que ce n'est pas le nombre d'images qui est pertinent, mais bien le rapport entre ce nombre et le nombre de vers ou lignes de la pièce. En d'autres termes, il aurait fallu faire apparaître dans une troisième colonne le résultat de la division du nombre d'images par le nombre de vers ou lignes, et, ce faisant, donner pour chacune des pièces ce qu'on pourrait appeler un "indice de densité d'images". C'est ainsi que, pour la première partie de *Henry VI*, nous aurions:

$$152 : 2676 = 0,056.$$

L'ensemble des indices de densité ainsi obtenus serait déjà plus instructif que les données apparaissant dans les colonnes (1) et (2). Mais on ne peut en rester là. Il faut procéder à une exploitation statistique de ces données "brutes". Le premier pas dans cette direction, c'est-à-dire le simple calcul de la moyenne des indices de densité, mettra en évidence une première difficulté. On s'apercevra que des valeurs comme 0,056 (ou 0,060, ou 0,068) se prêtent mal aux calculs, et on se souviendra que le traitement statistique des données de mesure peut s'effectuer sur des données multipliées ou divisées par une constante sans que le résultat des tests de signification les plus couramment utilisés, *t* de Student ou analyse de variance, en soit affecté. Bien entendu, si l'on tient à avoir la véritable moyenne, le véritable écart-type, et plus généralement les valeurs vraies des paramètres du premier degré, on multipliera les valeurs obtenues par la constante dans le cas où l'on a divisé au départ par cette constante, et on divisera dans le cas où l'on a multiplié au départ. Quant aux paramètres du second degré, et notamment la variance, c'est par le carré de la constante que l'on multipliera ou divisera selon le cas. Toutefois, il convient