THE SURGICAL TREATISES OF THE CORPUS HIPPOCRATICUM: STATISTICAL LINGUISTICS AND AUTHORSHIP

Mikel Labiano

1. INTRODUCTION

The surgical treatises1 of the Corpus Hippocraticum (CH), Joints (Art.), Fractures (Fract.), Surgery (Off.), Instruments of reduction (Mochl.), Ulcers (Vlc.), Wounds in the head (VC), Fistulas (Fist.) and Haemorrhoids (Haem.) are the most attractive among the whole corpus of medical writings generally ascribed to Hippocrates in antiquity, mainly Art., Fract. and VC, due to their high level of technical composition and scientific and methodological care. However, we honestly know little or nothing about their author(s), date of composition, and place within the CH. Though it is true that throughout the centuries to present times there have been many different hypotheses attempting to solve these problems, they have met with little success and no certainty. See for example Lain (1970), Bourgey (1953: 15–105), Withington (1928: XXII–XXV, 2–5, 84–93), Jouanna (1992: 527–563), Duminil (1998: 25), and Rodríguez Alfaíme (2000: 174).

So, from linguistic, stylistic, historical, conceptual, medical, etc. perspectives, many hypotheses have been proposed, including a common author for Art., Fract., and VC, connected to the author of Prog., Acut., Aph., and Aër. Other possibilities include the following: Off. has connections with Fract.; Mochl. has connections with Art.; despite these connections Mochl. and Off. are from different authors than Fract. and Art., which share one common author as previously mentioned; Fist. and Haem. have one common author, probably the same as VC; Fist. and Haem. are from different authors than VC; Vlc. is probably from the author of VC; Fist. and Haem. share one common author with Vlc., etc. The list is endless, and, in short, chaotic. Some of these theories and hypotheses are evidently contradictory and experts can reach no agreement. So we are left with an important

1 This paper, based on research funded by project FFI2012–36050 of the Spanish Ministry of Economy, is a revision and update of a previous paper (Labiano 2004).
and crucial set of medical writings in the history of ancient medicine, but few clear ideas about their formation, composition, relation to each other, author(s), chronology, and adhesion to one medical school or another, among other concerns.

The purpose of this paper is therefore to attempt to clarify some of this ambiguity and contradiction, and, in this respect, a statistical study applied to linguistics – that is, statistical linguistics – is proposed as a useful method and a starting point for future studies. This procedure has been successfully applied to other medical writings in CH in previous studies by the research team of the Universidad Complutense de Madrid (UCM). See Ángel (2003) for Ep. V and VII, Sierra de Grado (2003) for VM, Aër., Carn., and Prorrh. II, and especially Rodríguez Alfageme for Gland. (1992), Vid. Ac. (1993), and Loc. Hom. (2010).

In short, this method involves a count of syntactic links (particles of coordination, links of coordination and subordination, absolute genitive, asyndeton), in groups of 25 sentences over a text, in order to compare the average and statistical deviations, using the Student’s t-test within two texts. This will determine whether the two compared texts may originate from the same writer, in terms of statistical probability. The first stage, therefore, is to “translate” the text into numbers, in order to make statistical analysis possible (Muller 1973: 15). Through comparisons of sets of two texts and their statistical deviations, several types of relations can be established between them based on statistical probabilities. These are relations of common/divergent authorship, proximity/distance in time, in style, etc.

Analysis of syntactic distribution of sentences and links throughout one text allows us to obtain a syntactic pattern – a syntactic style, unique to each writer, with a constant rate in every era. The use and distribution of these links, present or absent (asyndeton), in relation to the content of the text, reflects the logical and psychological methods each writer applies to his text, his attention to style, etc. In short, this syntactic pattern is a unique and personal way to represent what a writer wants to communicate. It is a personal brand; his personal brand.

According to these statistical probabilities – and it is important to emphasize that they are statistical – it can be shown statistically that two texts are from the same writer or, at least, that they belong to the same period. It is also possible, on the other hand, to deny all these connections. This is the main hypothesis, or null hypothesis, that two texts come from the same writer/author, whereas the alternative hypothesis is the opposite