

A Survey of “t-h” Joins in the Definite Article and Other Wording

J. R. Welch, B.Sc., MCSFS.

8 Egdean Walk, Sevenoaks, Kent, TN13 3UQ, United Kingdom

A simple classification for the methods of joining letter “t” to letter “h” in lower-case handwriting was developed and used to gather data on “th” joins in the definite article and in other wording. The frequency of occurrence of each category of join was obtained in the two environments. The results suggest that the two environments are largely the same in respect of this classification of this feature.

The frequency of occurrence of handwriting characteristics in a population of writers is of considerable importance in the comparison of handwriting. The gathering of such data is a continuing area of research in document examination and important work has been undertaken in the UK by Nicholson with regard to block capital writing [1], and by Eldridge et al concerning cursive writing [2], and more recently and more extensively in America by Johnson, et al with regard to both sorts of writing [6].

A pre-requisite for the collection of data is the defining of characteristics which are to be looked for and counted. The process can give rise to confusing terminology: in this paper the aspect of handwriting observed is termed the feature and the differing variants of it are termed categories. Both the process of placing into categories, and the resulting simplified description of the handwriting, are termed classification. One of the greatest difficulties lies in defining the categories of each feature such that they are readily distinguishable. This is undoubtedly achieved with, for example, the feature “o” if the categories chosen are “written clockwise” and “written anti-clockwise” but the nature of handwriting means that such clear definitions are not always possible.

Previous work by Dr. Hilary Pritchard (Personal Communication) suggested that several variants of the letter “t” were likely to be present in any individual’s writing. This within person variability is possibly due to the form of the “t” depending on its position in the word and/or the letters before and after it. The number of variables which may affect the form of the “t” can be reduced by considering only one aspect of the letter “t” when it is in a particular environment. That is the approach taken here.

In the English language “th” is the most frequently occurring two letter group, [3] a fact which contributed to Muehlberger et al choosing that letter group for their research. [4]. The features chosen by Muehlberger and colleagues to classify required fine judgements to be made on, for example, relative heights, lengths and curvatures, when assigning features of the writing to different categories. In the work reported here the feature chosen to classify was the method of joining “t” to “h” in the hope that the observations would be more unequivocal. The most frequently occurring three letter group in English is “the” [3]. It has been suggested that the writing of “the” as the definite article is especially characteristic (Dr. A. Wing, Personal Communication) and a survey of how children make “t-h” joins in the definite article has been reported by Sassoon et al [5]. In the work reported by the current author two surveys were made of the manner of the “t” joining the “h” in “the” as the definite article and the “th” group in other wording. The surveys were intended to determine

1. whether or not “t-h” joins in each environment are constant for one writer,
2. the frequency of occurrence of each category in the two environments, and
3. whether or not the “t-h” joins of each writer in the two environments are similar enough for the two environments to be treated as one.

The defined categories of “t-h” join were “Via Cross-Stroke”, “Via Foot”, “Not Joined” and “Other”. (Please note that the terms “cross-stroke” and “cross-bar” are used interchangeably in this article.)

FIGURE 1

A simple classification such as this obviously makes use of only a small fraction of the information contained within the letters under study. This is illustrated by the widely differing "t-h" formations which share the same category with regard to this feature. Inevitably, some formations could not be assigned to any of the three

main categories and were classified as "other". Two examples are shown below. In example 1 the loop of the "h" may or may not incorporate the "t" cross-stroke. In example 2 the join could be regarded as either via foot or via cross-stroke, a situation complicated by the writer occasionally adding a separate cross-stroke.

VIA CROSS-STROKE		VIA FOOT	NOT JOINED
Th ₃₁	Th ₁₃₅	Th ₁₂₀	th ₁₅₆
Th ₃₇	Th ₁₃₉	th ₁₅₉	th ₁₈₅
Th ₄₆ NOT CAPITAL 't'	Th ₂₂₄	th ₂₀₇	Th ₂₀₁
Th ₅₀	Th ₂₂₉	th ₂₀₈	th ₂₀₅
Th ₇₂	th ₂₄₀	th ₂₃₂	Th ₂₁₁
Th ₇₂	Th ₂₄₁		th ₂₁₆
Th ₈₉	th ₂₄₁		Th ₂₄₃
Th ₁₁₁			th ₂₅₅

Figure 1.

FIGURE 2

The classification requires some subjective decisions to be made as to whether the two letters are "joined" by whichever route or "not joined". Instances where there is a clear gap between the two letters or there is a solid ink line joining them present no problem. The difficulty arises where there is an appreciable reduction in pen-pressure when the pen is moving from one letter to the next: if the pen lifts from the paper is the resulting gap intended or accidental? Similarly, where a writer habitually makes a pen-lift, should he be resting on a springy surface the paper may move forward sufficiently as pen-pressure is reduced to stay in contact with the pen. In these surveys

the apparent intention of the writer to leave a gap or not was estimated by considering the variation in pen-pressure displayed elsewhere in the writing. The difficulty of this task has probably contributed to the number of writers who appear in both the "not Joined" and "Joined via Foot" categories.

The writings surveyed were those of 257 people attending courses at the Metropolitan Police Training Centre, Hendon, London. Each participant was provided with a typewritten passage of 117 words and asked to copy it out twice, once "in your normal style of writing" and once "in block capital writing". All participants interpreted "your normal style of writing" as some form of lower case style; they were probably en-

couraged to do so by the contrast with “block capital writing”. They were asked to write with a ball-point pen and most but not all did so. The writings were made in a classroom environment.

The writers were not a cross-section of the population: females were grossly under-represented at around 3% of the total (11 in 257). The ages of the great majority of the writers were in the twenties and thirties. Left-handed writers formed approximately 12% of the total (30 in 257). Approximately 90% of the writers were educated in the

U.K., the remainder were educated abroad mainly in countries with historic, cultural or trading links with the U.K.: all the participants were fluent in reading and writing English.



EXAMPLE 1.

EXAMPLE 2.

Figure 2.

TABLE 1

From Table 1 it can be seen that in the survey of the definite article the “t-h” joins of 217 writers (84%) fall into one or other of the three mutually exclusive categories “Cross Stroke Only”, “Foot Only” and “Not Joined Only”: for the survey of “t-h” joins excluding the definite article the figure is 206 (80%). It is therefore clear that in each survey a large proportion of writers are consistent in the way they write “t-h” joins. In each situation very few writers, 4 and 10 respectively, have examples of joins via the cross-stroke and via the foot. Where two categories of “t-h” join occur in the writing of one person they are likely to be “Not Joined” and “Joined via Foot”.

In practice, when only a small amount of writing is available for comparison, a given feature may be present once only. Under these circumstances the document examiner cannot know whether or not the observed variant of the feature is the only one used by the writer. Consequently, a more useful frequency distribution is that showing the number of writers who al-

ways, or sometimes, write a particular variant. Such a frequency distribution can be obtained by summing data contained in Table 1 to produce Table 2.

Table 1.

	t-h JOINS IN THE DEFINITE ARTICLE	t-h JOINS ELSEWHERE
CROSS-BAR ONLY	33	32
FOOT ONLY	118	111
BOTH		10
NOT JOINED ONLY	66	63
NOT JOINED/JOINED CROSS BAR	7	5
NOT JOINED/JOINED FOOT	24	
NOT JOINED/JOINED CROSS BAR/ JOINED FOOT		2
OTHER	5	
OTHER/JOINED CROSS BAR		1

TABLE 2

In Table 1 the frequency of occurrence of the defined categories in the definite article is similar to that in “t-h” elsewhere. The pooled data for the two situations in Table 2 are also similar. The data from the two surveys was inspected to determine the number of writers who wrote “t-h” joins in the same way, and in differing ways, in the two environments. The number writing “t-h” joins of the same category (or categories where the writer uses more than one variant) in both the definite article and elsewhere was 207, approximately 81% of the writers. For the remaining 50 writers the differences were all caused by the use of an extra category in one or other of the environments; there were no writers who totally changed categories from one environment to the other. Therefore, there is appreciable evidence that, for the purpose of classifying “t-h” joins the two environments, “the definite article” and “th elsewhere” may be treated as one. However, it bears repetition that the classification is concerned with only one feature and makes use of only a small fraction of the information contained with the letters. This is illustrated by the observation that some writers who used the same category of “t-h” join in the two environments used a “t” of more flamboy-

Table 2.

	IN THE DEFINITE ARTICLE	IN "th" ELSEWHERE
TOTAL NUMBER OF WRITERS WRITING SOME OR ALL "th"s WITH CROSS BAR JOIN		50
TOTAL NUMBER OF WRITERS WRITING SOME OR ALL "th"s WITH FOOT JOIN	146	152
TOTAL NUMBER OF WRITERS WRITING SOME OR ALL "th"s WITHOUT JOIN	97	99
TOTAL NUMBER OF WRITERS WRITING "th" JOINS OF MORE THAN ONE CATEGORY	35	

ant appearance in the definite article. The different appearance was produced by the use of, for example, longer lead-in strokes, larger loops and longer retraces.

DISCUSSION

The writings surveyed were produced in the 1980s by a relatively small group of people in which there were few females and whose ages fell within a fairly small range. Consequently, the results should be considered as indicative rather than conclusive. In general, the results of the survey show that "t-h" joins are a usefully classifiable feature of lower-case handwriting. Whilst "th" and "the" occur with great frequency in most extended writing in English the author considers that other small words occurring frequently in written English such as "to", "of", "for", "by", "at", "it", and "and" may also be suitable subjects for specific classification schemes.

SUMMARY

The two surveys conducted have provided frequency distributions for the defined categories of "t-h" joins in the two environments "definite article" and "th elsewhere". The results indicate that in the definite article and in other

wording some 84% and 80% of writers, respectively, use only one of the defined categories of "t-h" join: that is, in each environment people are reasonably consistent in their writing of this feature. The frequency distributions in the two environments are sufficiently similar to suggest that the two environments may be regarded as the same for the purposes of classifying "t-h" joins in this way and, indeed, when comparing passages of handwriting in cases. Direct inspection of the survey data shows that some 81% of the writers used the same variant or variants in the two environments; indicating again that to a reasonable degree the two environments may be treated as one.

REFERENCES

1. Nicholson P.J. A system for the Classification of Block Capital Handwriting. Proceedings of the Tenth Meeting of the International Association of Forensic Sciences, Oxford, 1984, Paper G.426 (Microfiche).
2. Eldridge M.A., Nimmo-Smith I., Wing A.M. and Totty R.N. The Variability of Selected Features in Cursive Handwriting: Categorical Measures. *Journal of the Forensic Science Society*. 1984; 24: 179-219.
3. Higenbottom F. Codes and Ciphers. London: Hodder and Stoughton.
4. Muehlberger R.J., Newman K.W., Regent James, Wichman J.G. A Statistical Examination of Selected Handwriting Characteristics. *J Forensic Sci.*, Jan. 1977, Vol 22, No.1
5. Sassoon R., Nimmo-Smith I. and Wing A.M. Developing Efficiency in Cursive Handwriting: An Analysis of "t" Crossing Behaviour in Children. Computer Recognition and Human Production of Handwriting. Eds: Plamondon R., Suen C.Y., and Simmer M.K. World Scientific Publ.Co. 1989. pp 287-297.
6. Johnson M. E., Vastrick T. W., Boulanger M., and Schuetzner E.. Measuring the Frequency Occurrence of Handwriting and Hand-Printing Characteristics. *Journal of Forensic Sciences*, Vol. 62, No.1, January, 2017.
7. Pritchard, H. Personal Communication.
8. Wing, A. Personal Communication.