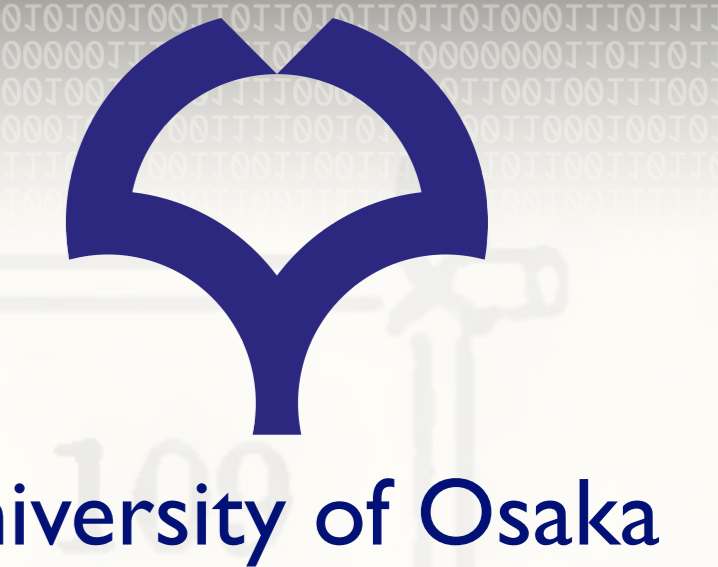


Collaborative Texts under a Stylometric Microscope: Investigating Cases of Mixed Authorship

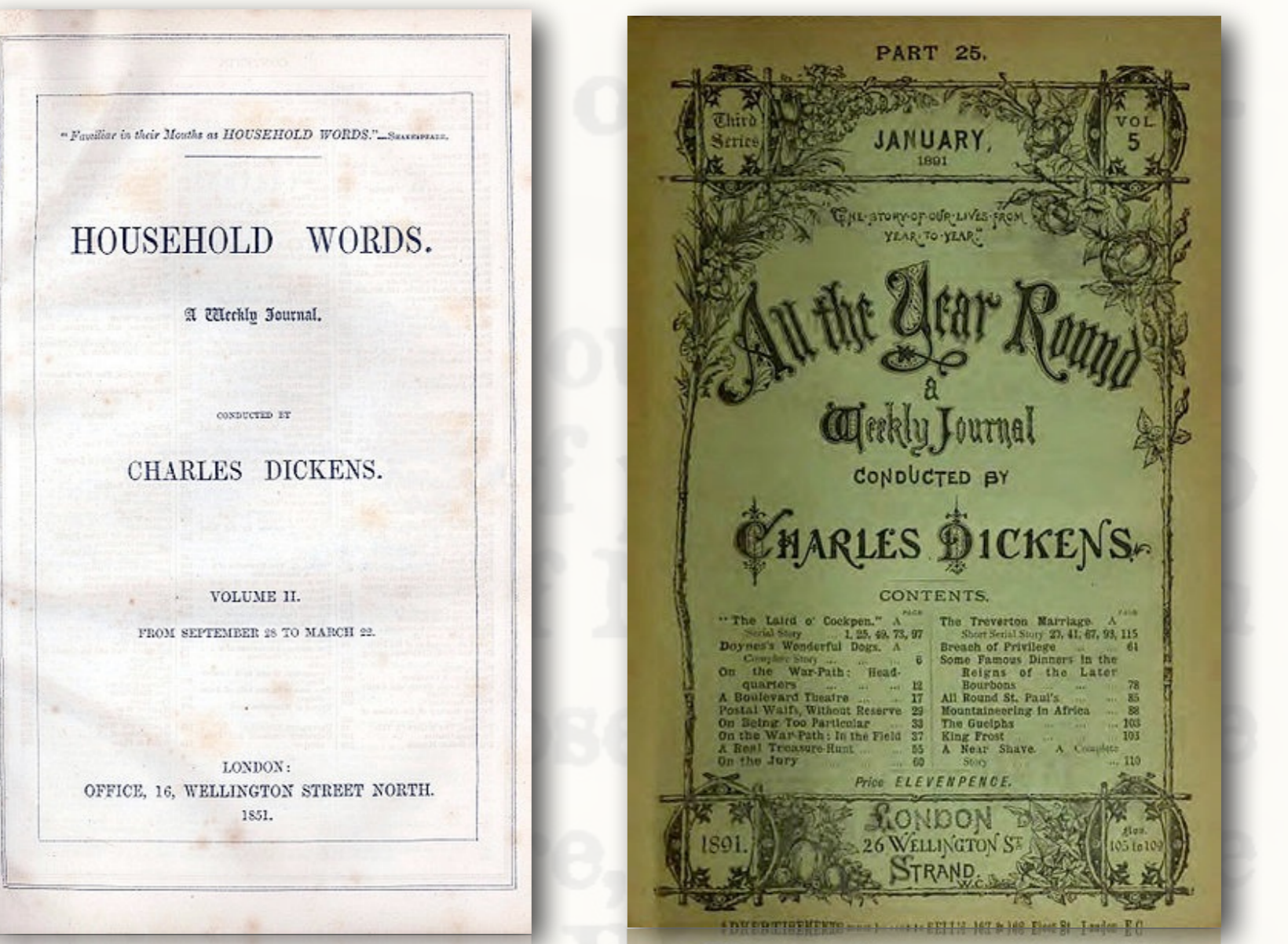
Digital Humanities

Tomoji Tabata
Dept. Linguistics and Language Informatics
Graduate School of Language and Culture



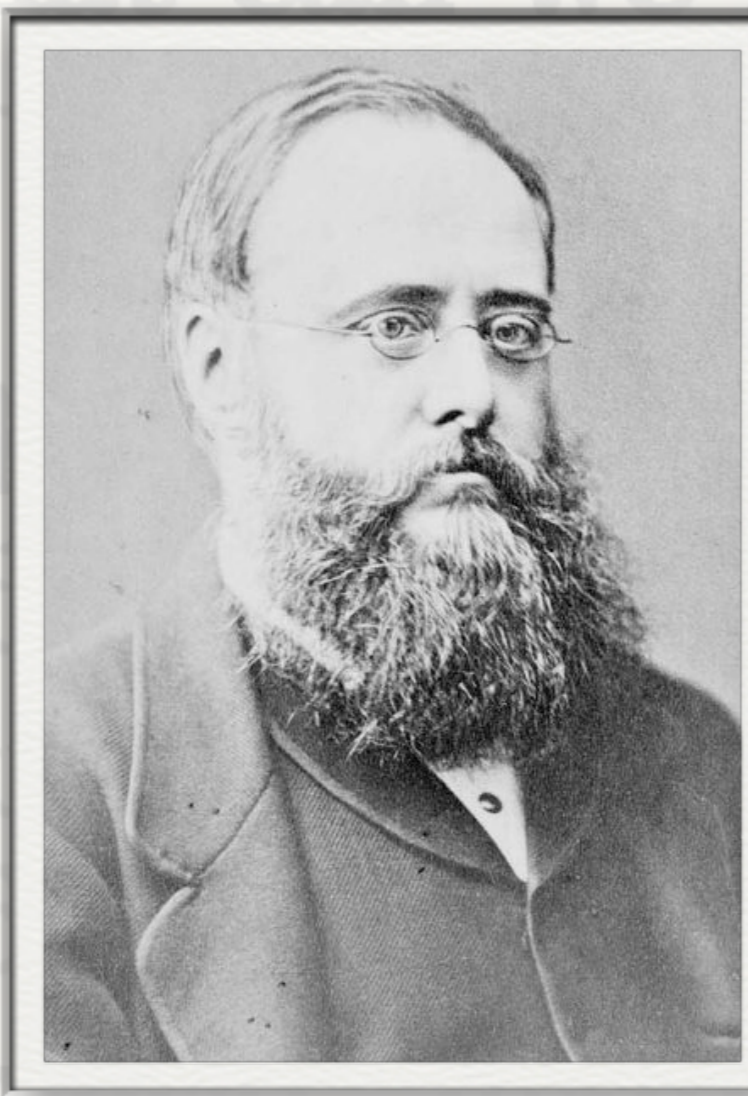
The Victorian author **Charles Dickens** was among the first publishing entrepreneurs to run mass-produced weekly magazines on a successful commercial basis. He employed many 'salaried staff writers' (Nayder, 2002), who had to write under anonymity, including Elizabeth Gaskell, Adelaide Anne Proctor among others, in *Household Words* and *All the Year Round*, the journals 'conducted by' Dickens (Stone, 1968; Thomas, 1982; Allingham, 2011).

On the other hand, **Dickens collaborated with his younger contemporary Wilkie Collins** on a number of stories, typically for the Christmas Numbers of his journals. While some of their collaborative pieces were written with the assistance of other staff writers, four works are known to have been co-authored by Dickens and Collins alone (Nayder, 2002). The four collaborations can be seen as betokening what appears to be a firm presence of Collins, a foothold he had gained, in the Dickens circle by the time he and Dickens launched into the joint works beginning in 1857.



Journals 'conducted' by Dickens:

- *Household Words* (27 March 1850–28 May 1859)
- *All the Year Round* (1859–1895)



Dickens's collaboration with Collins

- *The Frozen Deep* (1857, 1866, 1874)
- *The Lazy Tour of Two Idle Apprentices* (1857)
- *The Perils of Certain English Prisoners* (1857)
- *No Thoroughfare* (1867)



Bibliographic details about the collaborations

No.	Date	Title	Part	Authorship
1	1857, 1866, 1874	<i>The Frozen Deep</i>		Collins & Dickens ⇒ Collins
2	1857	<i>The Lazy Tour of Two Idle Apprentices</i>	Chapter I	Dickens & Collins
			Chapter II	Dickens & Collins
			Chapter III	Dickens & Collins
			Chapter IV	Dickens
			Chapter V	Dickens & Collins
3	1857	<i>The Perils of Certain English Prisoners</i>	Chapter I	Dickens
			Chapter II	Collins
			Chapter III	Dickens
			Overture	Dickens
4	1867	<i>No Thoroughfare</i>	Act I	Dickens & Collins
			Act II	Collins
			Act III	Dickens
			Act IV	Dickens & Collins

Rolling SVM (Support Vector Machine)

Figures 2 & 3 show a result of Rolling SVM run with a window size set to 3,000 words, a step size of 300 words, using 100 features. The whole texts of collaborations are sliced into consecutive 3,000 word-segments, with each segment compared with the centroids of Dickens and Collins, respectively, to determine the likelihood of it being written by one author rather than the other.

Fig. 1: Dickens, Collins, and the collaborations

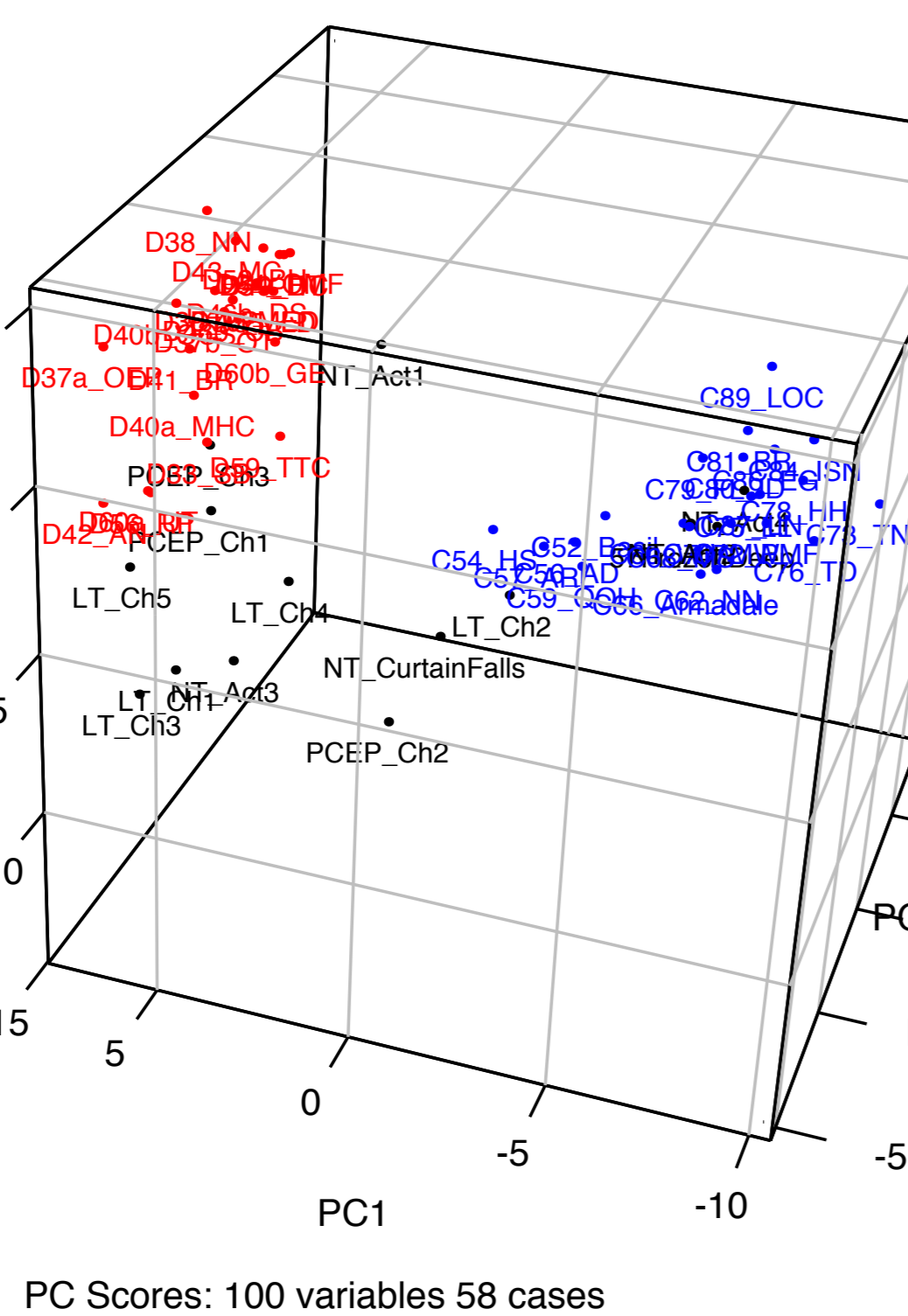
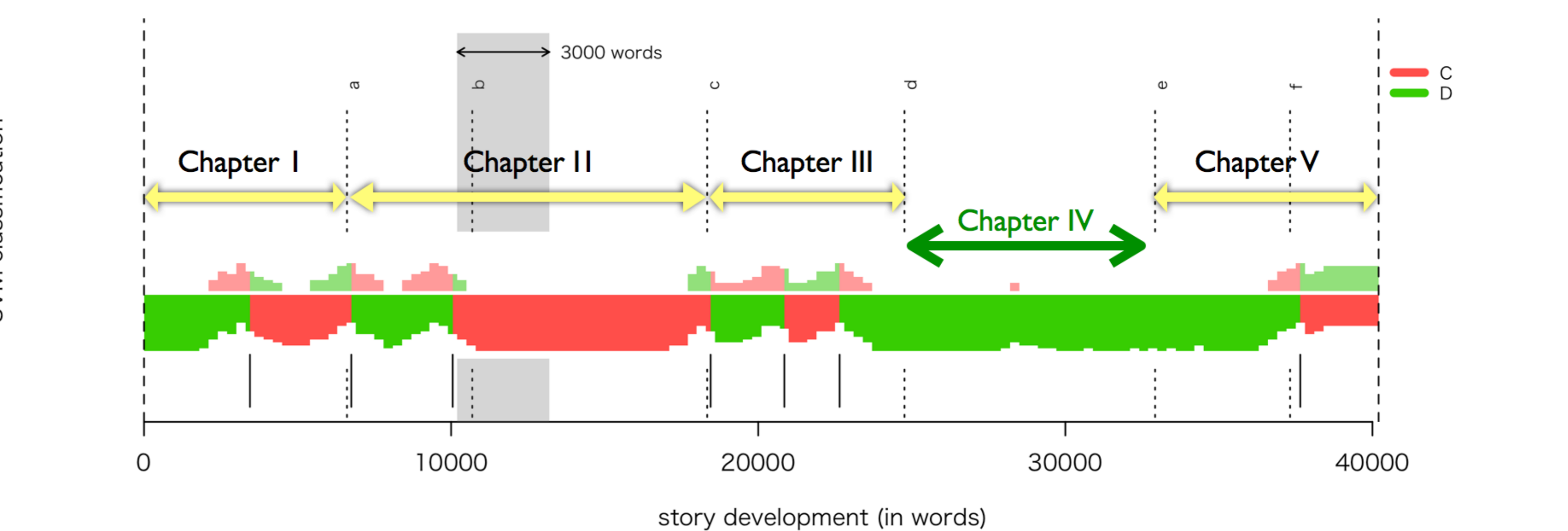
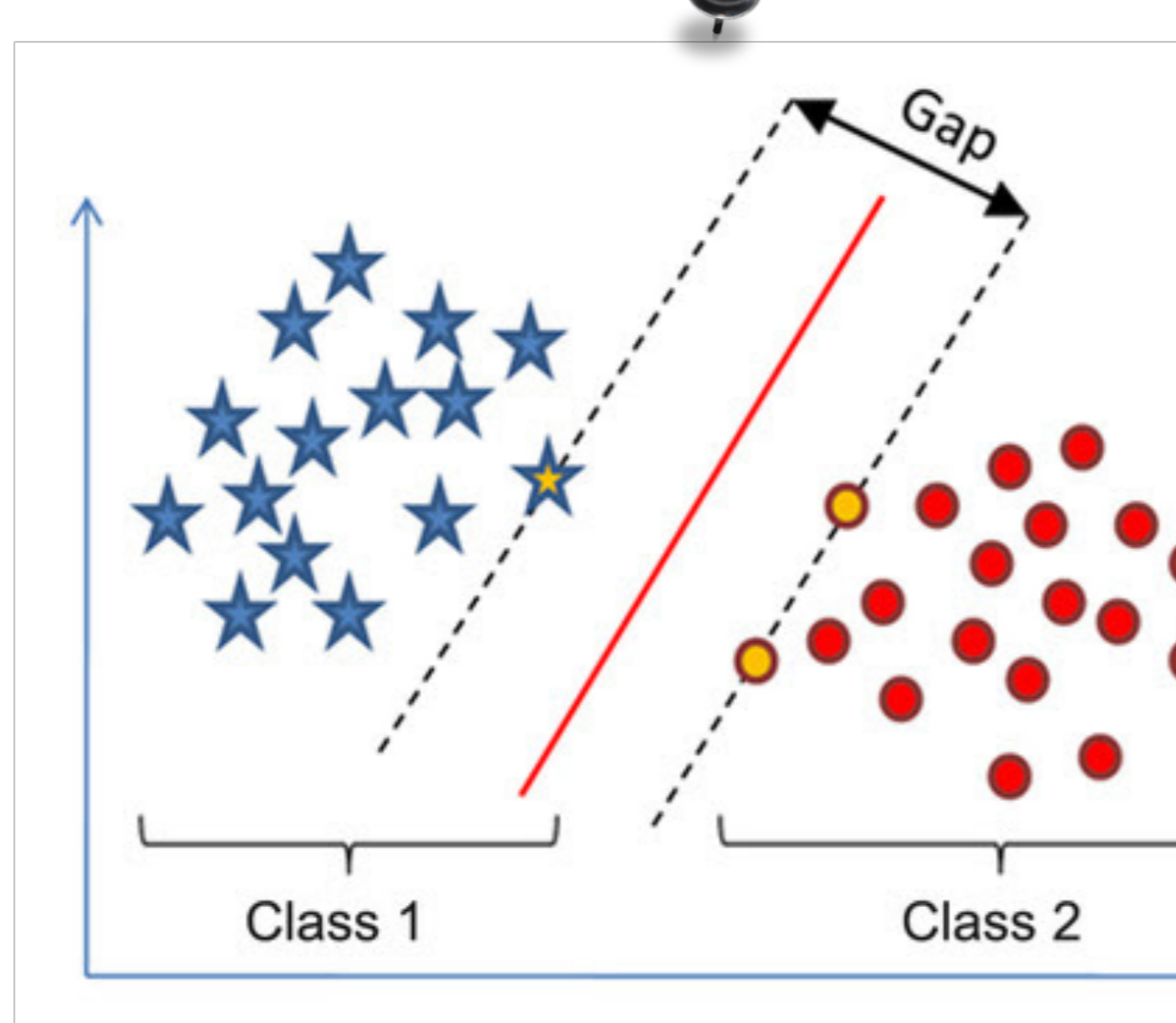
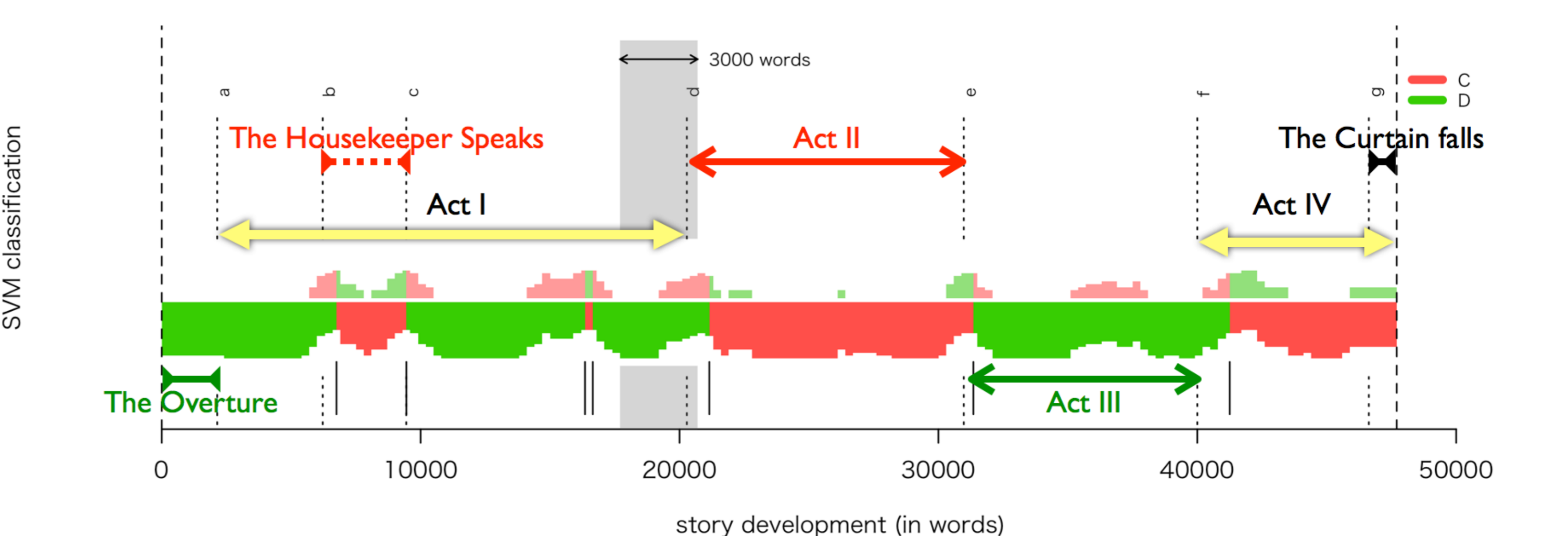


Fig. 2: Rolling SVM tested on *The Lazy Tour of Two Idle Apprentices*



Rolling SVM makes it possible to pinpoint exactly where one author takes over from the other with a high precision. Of particular interest is that a remarkably similar pattern holds throughout the five collaborated chapters: it is always Dickens who takes the lead at the outset of a chapter. He runs about one-third to halfway into each chapter before passing over to Collins. The diagrams seem to reflect an interesting nature of collaboration as well as the unequal partnership (Nayder, 2002) between Dickens and Collins: Dickens always takes initiative, sets a keynote for the whole chapter, which Collins takes over and continues the rest of the chapter, a typical relationship between a master and his disciples.

Fig. 3: Rolling SVM tested on *No Thoroughfare*



What Support Vector Machine does ...