Horace Walpole and his correspondents : social network analysis in a historical context

Henstra, F.H.

Chapter 6. Social network analysis and the problem of small numbers[1]

6.1. Introduction

In this chapter I will deal with the problem of small numbers of instances in sociohistorical linguistic analysis on the basis of data from the Horace Walpole correspondence. This, as I have demonstrated in chapter 5 above, is a particular challenge for linguists, and it has been argued that it can be approached from the viewpoint of what Labov calls "making the best use of bad data" (Labov 1994: 11). In focusing on this issue, I will discuss the kind of problems which seem to be inherent in the type of *bad data* that sociohistorical linguists use and in the models they have available for analysis, and propose possible strategies for dealing with these difficulties. My approach will draw on and combine two models that have been in use within social network analysis: what I have called in section 4.4 the classical model of network strength analysis, which is based largely on biographical and contextual information on the authors analysed, and a linguistic model for measuring linguistic involvement used as an indicator of network strength. These models will be applied to a closed network cluster in Horace Walpole's social network, for which I will analyse the language of his correspondence with Thomas Gray (1716–1771), Richard West (1716–1742) and Thomas Ashton (1715–1775). The linguistic feature I will analyse is the variation in the usage of *be* and *have* in constructions of the perfect with mutative intransitive verbs in order to test the suitability of the different models for explanatory purposes.

[1] This chapter is based on an earlier version of Henstra (2009). I am grateful to Anni Sairio for her helpful comments and suggestions during the writing of this chapter, as well as for the comments of anonymous readers of the original article.

6.2. The network cluster and the corpus

6.2.1. The correspondents

Within the greater network of the correspondents of Horace Walpole that is the object of this study I will focus in this chapter on the cluster in the network consisting of Walpole and his Eton school friends Ashton, Gray and West. With some of them Walpole maintained a lifelong friendship and correspondence. The group was very close knit and the young men even identified themselves by means of a special name: the "Quadruple Alliance" (HWC 13: xxiii). They also gave each other nicknames: according to the *ODNB*, "Ashton was Almanzor from John Dryden's *Conquest of Granada*; Walpole was Celadon from D'Urfé's *Astrée*; West was Favonius or Zephyrus from Latin names for west winds; and Gray was Orosmades from Nathaniel Lee's *The Rival Queens*" (*ODNB*, s.v. Thomas Gray). Walpole is claimed by the editors of his correspondence to have been somewhat of a leader to the group: "[h]e assumed that rôle inevitably, not because he was the Prime Minister's son … but because he was gay and gregarious and had a gift for friendship" (HWC 13: xxiii). Perhaps this gift was an important factor in the development of the large network of friends and acquaintances with whom Walpole corresponded throughout his life.

Walpole had been at Eton from 1727 until 1734, when he went to the University of Cambridge to continue his education. At Cambridge, he joined his friend Thomas Ashton at King's College. Ashton was to make a career in the clergy, and is said to have been much furthered in this by Walpole. The editors of HWC note that Ashton is often "put down as a time-server who attached himself to the Prime Minister's son at Eton with a view to securing future preferments" (HWC 13: xxvii). Ashton and Walpole eventually fell out over a religious pamphlet written by Ashton in July 1750 as an attack on Walpole's

friend the reverend Conyers Middleton (1683-1750), and the break was never mended: after that time Thomas Ashton was no longer a part of Walpole's life and the correspondence between the two men ceased. The poet and literary scholar Thomas Gray is perhaps the best known of Walpole's correspondents dealt with in this chapter. After studying at Eton and Cambridge, Gray went on a European tour together with Walpole in 1739. It was not long into the trip, however, before the men experienced difficulties in travelling as friends. Gray and Walpole eventually each went their own way at Reggio in 1741 after an incident of which the nature has never become quite clear. They resolved their differences in 1745, however, and not so much rekindled, as recreated their friendship. The friendship in later life centred around the publishing of Gray's poetry by Robert Dodsley, in which Walpole was closely involved, while Gray also assisted Walpole in his historical research. The fourth member of the quadruple alliance is Richard West. He was with the others at Eton, but went to study at Oxford in 1735 instead of Cambridge. In September 1741, upon his return from Europe, Gray found their friend ill and declining. West died in London in 1742 at the age of only 26. Poetry was an important part of the lives of all four men, and their correspondence played an instrumental part in their poetical enterprises. As the *ONDB* puts it, "[m]ost of West's small output of poetry", for example, "was enclosed in letters to Gray" (*ODNB*, s.v. Richard West). West was considered by the others as "the truest poet among them" (HWC 13: xxviii), but because of his untimely death he would never mature to his full potential as a poet.

6.2.2. The letters

Throughout the correspondence, the four men are linked together by their shared love of poetry, music, art, classical culture and literature, which is also

reflected in the nicknames they gave each other (6.2.1). Their earliest letters abound in literary allusions and parodies, and they consequently often have a jocular tone, as the following letter from Gray to Walpole demonstrates:

> *To Mie Nuss att London*
>
> *Honner'd Nurse,*
>
> *This comes to let you know that I am in good health, but that I should not have been so if it had not been for your kind promise of coming to tend m e yourself and see the effect of your own prescription.*
> *And I should desire of you, so please you, as how that you would be so good as to be so kind as to do me the favour of bringing down with you a quantity of it, prepared as your grandmother's aunt, poor Mrs Hawthorn (God rest her soul, for she was as well a natured a good gentlewoman as ever broke bread or trod upon shoe-leather, though I say it that should not say it, for you know she was related to me, and marry, not a jot the worse, I trow!) used to make it. Now I would not put you to this trouble if I could provide myself of the ingredients here, but truly, when I went to the poticaries for a drachm of spirit of ridicule, the saucy jackanapes of a prentice-boy fleered at me, I warrant ye, as who should say, you don't know your errand. So by my troth, away ambles me I (like a fool as I came) home again, and when I came to look of your receipt, to be sure, there was spirit of RIDICULE in great letters, as plain as the nose in one's face. And so, back hurries I, in a making-water-while, as one may say; and when I came there, says I, you stripling, up-start, worsted-stocking, white-livered, lath-backed,s impudent princox, says I, abuse me that am your betters every day in the week, says I, you ill-begotten, pocky, rascally, damned son of a bitch, says I—for you know, when he put me in such a perilous passion how could one help telling him his own—why, 'twould have provoked any Christian? in the world, though 'twere a dog, to speak. And so if you'll be so kind, I'll take care you shall be satisfied for your trouble.*
> *So, this is all at present from Your ever-dutiful and most obedient and most affectionate loving god-daughter,*
> *PRU. OROSMADES*
>
> (Thomas Gray to Walpole, 17 November 1734, HWC 13:61-62)

Walpole's correspondence in later life is also often funny and full of allusions to art, political events and history, but none of those letters can match the plain fun that is emanating from the early letters written by the members of this

network cluster. In the letters the men often expressed how displeased they were at not seeing or hearing enough of each other, sometimes in a more serious tone, but often in a light-hearted and witty manner. In one of his letters Gray labours the point, for instance, that Walpole must think he is dead, or else he would not have neglected him so badly: "Dear *Dimidium animae meae,* As you take a great deal of pleasure in concluding that I am dead, and resolve not to let me live any longer; methinks you ought to be good to my ashes" (Gray to Walpole, ca. 29 December 1734, HWC 13: 69).

As the relationships matured and circumstances changed, so did the tone and content of the letters. The boyish jokes and wittiness in the early letters gave way (though never completely) to a more serious attempt at maintaining a friendship while physical distances increased. In the letters dating from after 1739, when Gray and Walpole were travelling the continent together, they tried to share as much of what they saw in Europe as they could with West and Ashton back in England:

> Ever since Wednesday, the day we were [at Versailles], we have done nothing but dispute about it. They say. We did not see it to advantage, that we ran through the apartments, saw the garden *en passant,* and slubbered over Trianon. I say, we saw nothing. However, we had time to see that the great front is a lumber of littleness, composed of black brick, stuck full of bad old busts, and fringed with gold rails (Walpole to West, 15 May 1739, HWC 13: 168).

It may be noted that the arts and culture were still important subjects in the lives and correspondence of the men, as they had been the case since their earliest days at Eton. Social gossip is a second recurring theme in their letters, a subject on which Walpole was able to employ his sarcastic but humorous tone to its full potential, jokingly, for instance, calling Lady Mary Wortley Montagu a

"she-meteor" and her conversations with two friends a "rhapsody of mystic nonsense" (Walpole to West, 31 July 1740, HWC 13: 227–228).

The correspondence between Gray and Walpole after their reconciliation in 1745 is also mainly concerned with art and poetry, though more in a practical way than in the allusive manner of their youthful letters:

> I am very glad my objections serve only to strengthen your first opinion about the subject of your picture; if I casually meet with anything more, I shall send it you. The reason I trouble you at present is to tell you that I have got in my hands the Dugdale Mr Chute inquired after (Gray to Walpole, 11 April 1754, HWC 14: 81).

After West's premature death in 1742 and Walpole's definitive break with Ashton in 1750, Gray and Walpole were the only two members left of the quadruple alliance. The later letters reflect the lasting closeness between Walpole and Gray, a closeness which was maintained perhaps because of their shared history, certainly because of their shared interests in art, history and culture. Ironically enough, the men also shared a medical problem, which occupied both to a great extent in the latter part of their lives: "You are very kind to inquire so particularly after my gout: I wish I may not be too circumstantial in my answer; but you have tapped a dangerous topic; I can talk gout by the hour" (Walpole to Gray, 19 November 1765, HWC 14: 142).

6.2.3. The corpus for analysis

The first of the problems concerning models and data that needs to be addressed here has to do with the nature of the material which is available to me, which was referred to in chapter 5 as the problem of *bad data*. The problem in question may be best illustrated by a description of the corpus which I have compiled for the current linguistic analysis. The corpus is divided

into three discrete periods, based on the background information about the correspondents and their letters discussed above. Period I, from April 1734 to March 1739, is the time at which the correspondents were first in transition from Eton to university and later when they were all at their respective colleges. However, being at university did not necessarily imply a permanent physical presence there: the men went on trips abroad and were sometimes otherwise away from university during lengthy periods of time. It is noted in the *ODNB*, for example, that Walpole "left Cambridge … after increasingly erratic appearances there" (*ODNB*, s.v. Horace Walpole). This is the period of the early letters in which the correspondents can be seen to mature from boys to gentlemen. The sub-corpus of Period II, from March 1739 to July 1741, consists of the letters written by Walpole and Gray during their tour on the continent, to West and Ashton at home, and of the letters from England to the continent. Period III, ranging from July 1741 to March 1771, consists of the letters by Walpole and Gray in their adult life, when they were the only two members of the quadruple alliance left, after West's death and the break with Ashton. The three periods represent different phases in the lives of the correspondents, and also different phases in the network: as discussed in chapter 3, the make-up of a social network is dynamic, networks change with time and circumstances as relationships do, and these three periods will therefore represent three radically different network structures for one and the same group of people. This should reflect possible changes in the linguistic reality of the network as well.

The overview of the corpus per period, presented in Table 6.1., instantly reveals a number of weaknesses and gaps in the corpus:

	Correspondent	Number of letters	Number of words
Period I: April 1734 to March 1739	Gray to Walpole	38	13,761
	West to Walpole	10	4553
	Ashton to Walpole *	1*	77*
	Walpole to Gray	1	1003
	Walpole to West	3	1053
	Total	***53***	***20,447***
Period II: March 1739 to July 1741	Gray to West	3	1062
	West to Walpole	7	4014
	Ashton to Walpole *	1*	549*
	Walpole to West	17	13,298
	Walpole to Ashton	3	1926
	Total	***31***	***20,849***
Period III: July 1741 to March 1771	Gray to Walpole	88	28,229
	Walpole to Gray	11	8859
	Walpole to West *	1*	481*
	Total	***100***	***37,569***

Table 6.1. Overview of the corpus per period (correspondences consisting of less than 1000 words have been marked with an asterisk)

In Period I there is hardly any material from Ashton. There are also only a few letters from Walpole, though the word count for those letters is higher than that for Ashton and should be sufficient for some cautious linguistic analyses. Since Period II is the time during which Walpole and Gray travelled through Europe together, there are no letters from Walpole and Gray to each other for this period. Gray did finish three of Walpole's letters to West, though, so we get some glimpse of Gray's language use in that period through these letters. Finally, for Period III there is only the correspondence between Gray and Walpole, since West had died in the meantime (only one letter remains, sent to West by Walpole after his return from the continent and shortly before West's death). Very small corpora cannot realistically be used for linguistic analysis,

and I have here marked correspondences of less then 1000 words (which is the minimum number of words on which, for practical purposes, I have decided in this specific analysis) with an asterisk. These parts of the corpus have not been used in the linguistic analysis which follows. What furthermore may be noted is the great unevenness in the number of letters and amount of text available for each of the correspondents. Overall, the writings of Gray to Walpole and of Walpole to West seem to be overrepresented in this corpus.

Focused historical corpora such as this one are likely to be (problematically) unbalanced and to contain gaps, since, as I have already discussed in chapters 2 and 5, as a researcher one is completely dependent on the historical material which is available for the selected informants in compiling a corpus. When performing linguistic analysis, the potential problem involved with underrepresentation or overrepresentation of certain correspondents in the corpus needs to be kept in mind. Even if the linguistic data are normalised to occurrences per 1000 words, the corpus itself remains unbalanced and therefore generalisations about the language of a network in relation to, for example, general eighteenth-century usage are to be made cautiously. It is true that historical linguists often have no choice but to work with what they have at their disposal, which is therefore often enough called *bad data.* However, I would also like to point out that in the kind of sociohistorical research which is described in this paper the use of focused rather than representative corpora for a micro or idiolectal analysis, rather than for making general statements on the state of the language, is a common practice.

6.3. Applying two models for sociolinguistic analysis

6.3.1. The classical network strength scale

For my analysis of the data from the Walpole correspondence Eton Network Cluster, I will draw on what I have called the classical network strength scale following Milroy (1987) (see chapter 4). One of the main propositions of that model is that closed or dense network clusters are able to maintain a norm of their own and that someone's relative integration into a network is one of the best predictors of linguistic behaviour. As the "Quadruple Alliance" in its full form qualifies as a closed network cluster, the model seems particularly appropriate to use. I will use the model as set out in chapter 4 but in the adapted form as proposed in section 4.3.3.; including notions on coalition formation taken from Fitzmaurice (2000b) as well as the comments made by Sairio (2005) that a scale from immediacy to distance, rather than Bax'ss scale from friend to enemy (Bax 2000), would be better suited for quantifying emotional relationships in historical data. The model adopted here thus consists of a functional and an emotional element, and I will consider all respective relationships within the Eton network cluster, assigning points accordingly. For full details, see Table 4.1. in chapter 4 above. The results of the analysis are presented in Table 6.2 below.

What can be seen from the results of the application of the network strength scale to the Walpole Eton Network in Table 6.2. is that network scores indeed differ greatly between the correspondents in the three discrete periods of time. It should be taken into account, though, that these periods were defined on the basis of the available biographical and contextual information about the different phases in the lives of the men and their friendships. The changes in network strength scores and relationships over time are in

accordance with the idea that relationships are in essence dynamic. In 5.5.1 I proposed a view of the network strength scale as a 'snapshot' of a social network at a particular time, while this particular time can also be a longer period which in case of any relative stability may still be characterised as a discrete one within the relationship.

			Correspondent B			
	Correspondent and period		*Gray*	*West*	*Ashton*	*Walpole*
Correspondent A	***Gray***	I		3	5	5
		II		2	2	6
		III		–	–2	4
	West	I	3		3	3
		II	2		3	2
		III	–		–	–
	Ashton	I	5	5		6
		II	2	3		2
		III	–2	–		–2
	Walpole	I	5	3	6	
		II	6	2	2	
		III	4	–	–2	
	Total network integration score of B	***I***	***13***	***11***	***14***	***14***
		II	***10***	***7***	***7***	***10***
		III	***2***	***–***	***–4***	***2***

Table 6.2. The network strength scale for the Walpole Family Network per correspondent and period analysed.

In this case the three periods represent three such *snapshots*, and the changes between the periods may be seen as representing the dynamic nature of the relationships between the four men. I have already cited Fitzmaurice's observation that in asymmetrical relationships "an individual may change network strength score with a shift from being the recipient of a non-reciprocal tie to gaining recognition as a reciprocal actor" (Fitzmaurice 2000b: 271) (4.5). However, no traces of shifting asymmetry within the network relationships can

be found in the classical network strength analysis presented in Table 6.2. This is due to the nature of the information upon which the analysis is based: this network analysis is wholly symmetrical, which is perhaps indicative of its rather generalised nature.

A basic premise of social network analysis is that if a network tie is stronger (represented in the current model by a higher network strength score between two members), influence is more likely to travel from one person to the other. If a network consists of many strong ties (and the total network integration scores are therefore high), the network may be considered dense, with, consequently, a likelihood of a strong internal norm being in operation. Networks which consist of weaker ties may be considered more open and are therefore more susceptible to the introduction of change (see Milroy 1987: 185–191). Looking at networks from the perspective of the individual one may say that a stronger tie between correspondents promotes the upkeep of the relationship-internal and network-internal linguistic norms by means of the travel of linguistic influence between the correspondents; on the other hand, a weaker, or in this model sometimes negative, tie leaves the correspondent more open to change from the outside and at the same time makes it less likely for him – in the case of the Eton Network Cluster – to be affected by the operation of the internal norm through linguistic influence within the network, since a weak tie represents a less responsive attitude towards any internal norm-enforcing influences. Given these general principles, the question arises as to what the network strength scale in Table 6.2. implies for the possibilities of linguistic influence between Walpole and the other members of the network cluster.

In Period I, the time during which Walpole and his friends were at university, all correspondents were very close to each other, although West

was slightly less connected. I have visualised the differences in connection strength as found in Table 6.2 by using thinner lines in Figure 6.1.

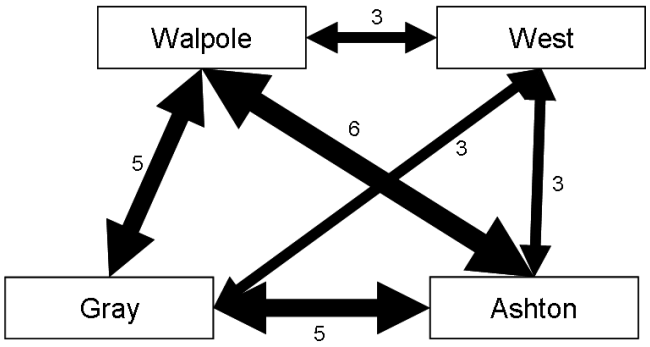

Figure 6.1. The Eton Network Cluster during Period I: April 1734 – March 1739

West was living away in Oxford at the time, and he therefore fulfils fewer functional relationship conditions than the other members of the network cluster who were living and studying in Cambridge together. In this network structure West, as the least integrated network member, is "relatively more *exposed* to the influence" (Milroy 1987: 196) of some outside or different norm, and he is therefore someone who would be able to introduce change into the network cluster. The scores of Ashton and Walpole are both slightly higher than those of the others, because they also lived at King's College together, and therefore fulfil an extra functional criterion. The possibility of linguistic influence is expected to be substantial between all network members, but also to be relatively equal between them.

During Period II, for which see Figure 6.2. below, there are literally two fronts: West and Ashton are in England on the one side of the diagram, and Walpole and Gray are taking the grand tour of Europe together on the other. Looking at the total network integration, the network integration scores for Walpole and Gray are much higher than those for West and Ashton, which

makes them more central to the network. Therefore the language of West and Ashton may be expected to approximate that of Gray and Walpole. Possible linguistic influence is also expected within the two groups, in a symmetrical way, with the network members adapting their language to one another.

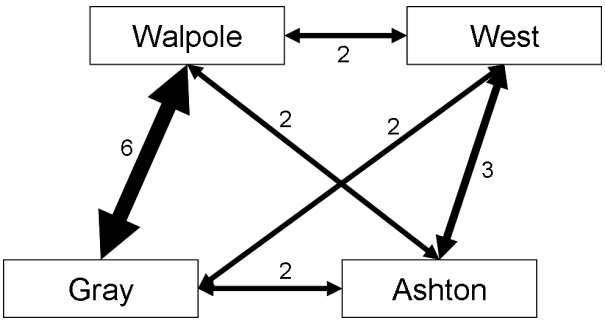

Figure 6.2. The Eton Network Cluster during Period II: March 1739 – July 1741

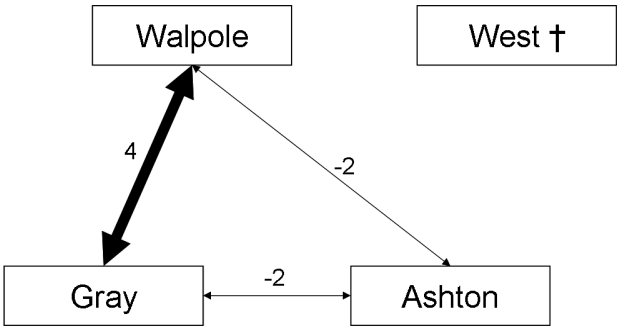

Figure 6.3. The Eton Network Cluster during Period III: July 1741 – March 1771

During Period III, represented in Figure 6.3., there is no longer a network score for West, as he had died at its very beginning, in 1742. Ashton is now an outsider to the network, due to his disagreement and subsequent break with Walpole. Gray and Walpole had reconciled in 1745: “In November 1745 …

Walpole wrote offering a meeting, and Gray went up to London, to be kissed on both cheeks … Ashton was no longer a person of significance in Gray's eyes" (*ODNB*, s.v. Thomas Gray). Gray sided with Walpole in his disagreement with Ashton. They would therefore be expected to dissociate linguistically from Ashton and *vice versa*, whereas the two friends are expected to have influenced each other reciprocally because of their close contact during the latter part of their lives (see Bax 2002 for more on CAT).

The closeness between Gray and Walpole during this period stems, among other things, from their shared enterprise of publishing Gray's poetry. However, due to the nature of the classical network strength model, the total integration score of the network is much lower here. There are fewer people and there are fewer functional relationships between them. One might wonder, though, whether this means that the network is actually to be considered weaker and thus more open to change from outside, or not. Gray and Walpole are clearly working together on a shared enterprise, a manner of closeness which perhaps cannot be expressed specifically enough in the current model (see also the application of Fitzmaurice's ideas on coalition formation (2000b) as discussed in chapters 4 and 5).

A second problem that needs to be raised here as being inherent in working with sociohistorical linguistic data and models, therefore, has to do with the risk of interpretation and generalisation in an analysis such as that of the Walpole Eton Network Cluster. It was noted that the basis of the classical network strength scale lies in the interpretation of biographical and contextual background information, as well as of the content of the correspondence. In this case the data consist of a network of informants who are no longer alive and therefore cannot be asked directly about their social situations. The best use of these *bad data* is made by interpreting the information we have from

first, sometimes second and often even third-hand parties. In chapter 5 I posited that many gaps in existing data can be filled that way, but that in the interpretation from contextual sources one is limited by the reliability of the second-hand information and the type of sources: was a personal letter written with possible other readers in mind? What is the influence of the subjective filter through which an observer describes a situation on the reliability of that background information? As discussed in chapter 4 (see table 4.2.), the nature of the sources used determines the reliability of the network analysis, and the most trustworthy source for background information is an author's private diary. Personal letters range somewhere in the middle of the continuum of reliability. Moreover, in the interpretation of the information there is also a risk that one only sees what is expected or what is a desired result, rather than a possibly more complicated or perhaps unclear truth.

A third problem to be considered in the context is that, in the light of the dynamic nature of relationships, the periods which I presented as relatively coherent periods of time in the analysis above are perhaps much too long and unequal (ranging from five to thirty years) to be seen as discrete, stable and comparable units. However, there are not enough data for analysing the relationships over much shorter time-spans. In analyses like the present one, a choice has to be made between discerning trends over longer periods of time, or taking snapshots of shorter periods, which has as a major drawback that there are often not enough data available to devise a reliable network analysis, or, indeed, to perform a linguistic analysis. The interpretation of network strength as discussed above for Walpole and his Eton correspondents must be seen as very tentative. This does not mean, however, that a classical network strength scale cannot be a useful tool for research, though one needs to be

cautious in applying it in view of the possibility of misinterpreting the available information.

6.3.2. Linguistic involvement

Whereas the classical network strength model of social network analysis is, for use in historical periods, very much dependent on the interpretation of background information, the model of analysis used by Sairio (2005), based on the work on involvement by Chafe (1985) and Palander-Collin (1999a, 1999b), hinges purely on linguistic elements, namely features of involvement (see chapter 3). According to this model, a higher degree in linguistic involvement is expected to coincide with a closer relationship in network terms (comparable to a higher network strength score in the classical model). In the following section I will apply a version of a linguistic involvement analysis to Walpole, Gray, Ashton and West, creating what I will call an *involvement* strength scale, in order to compare the implications of the results with those of the classical network strength scale discussed above.

As set out in chapter 4, the elements which make up an involvement score (calculated as a number of tokens per 1000 words) belong to three types. Firstly, there is self-involvement, including the use of first person pronouns, evidential constructions such as the use of *I think* and other references to the writer's mental processes. For the purpose of the present analysis I have adopted the list of evidential verbs given in Sairio (2005: 26) which consists of the verbs *think, know, believe, suppose, find, be sure* and *doubt*. Examples (1) and (2) illustrate this type of usage in the language of Walpole and his friends:

1. I am obliged on the sudden to come hither to see my poor mother, who is in a condition between life and death, though **(I think)** much nearer the latter (Gray to Walpole, 27 February 1753, HWC 14: 66).

2. The uncertainty of my silly health might have made me the duller companion, as you know very well; for which reason, fate took care to remove me out of your way; but my letters, **I am sure**, at least carry sincerity enough in them to recommend me to any one that has a curiosity (West to Walpole, 27 February 1737, HWC 13: 127).

Secondly, there is hearer-involvement, or in this case, rather adressee-involvement, which includes the use of what Chafe (1985: 117) calls "ubiquitous" *you know*, an example of which is given in (3), and the use of second person pronouns, for which see (4):

3. for we must give the Spaniards another drubbing, **you know** (Walpole to West, 20 July 1739, HWC 13: 180).

4. We have miserable weather for the season; could **you** think I was writing to **you** by my fireside at Rome in the middle of May (Walpole to Ashton, 14 May 1740 N.S., HWC 13: 214–215)?

By nature of the current concordance-based analysis of the texts, instances of *you know* function doubly as tokens of adressee-involvement (counted both as second person pronoun *you* and an instance of "ubiquitous" *you know)*. This is a problematical factor which is to be reconsidered carefully in future use of the model.

And finally there is subject-involvement, which is here represented by the use of intensifying degree adverbs such as *very* and *so*. For my analysis I have followed the features analysed by Sairio (2005: 26–27), including the intensifier adverbs *very, so, quite, pretty* and *really*, in which she applies Chafe's notion that subject-involvement, as "a speaker's lively interest in the subject matter being communicated", may be found in the use of "vivid particles" (Chafe 1985: 117). This is illustrated by examples (5) and (6):

5. I'm **pretty** sure, if I were divided into two persons one half would forget t'other very quickly (West to Walpole, 15 October 1739 O.S., HWC 13: 185).

6. About a month ago he was three votes of being Pope. He did not apply to any party, but went gleaning privately from all, and of a sudden burst out with a number; but too soon, and that threw him **quite** out (Walpole to Ashton, 15 May 1740 N.S., HWC 13: 214–215).

In Table 6.3., the results of the involvement analysis for the Walpole Eton Network Cluster can be found:

Period I	**Gray to HW**	**West to HW**	**HW to Gray**	**HW to West**
1st pers. pro.	56.1	66.11	30.9	46.53
2nd pers. pro.	31.03	25.92	2.99	31.34
you know	0.58	1.1	0	0
evidential	3.63	5.05	1.99	3.8
degree adverb	9.23	9.44	7.98	4.75
Involvement	***100.57***	***107.62***	***43.86***	***86.42***

Period II	Gray to West	West to HW	HW to West	HW to Ashton
1st pers. pro.	38.61	64.77	44.44	50.36
2nd pers. pro.	25.42	36.12	16.91	12.46
you know	0	1.99	0.53	1.04
evidential	0.94	5.48	3.53	5.71
degree adverb	8.47	9.97	7.82	5.71
Involvement	***73.44***	***118.33***	***72.23***	***75.28***

Period III	Gray to HW	HW to Gray
1st pers. pro.	46.76	55.31
2nd pers. pro.	25.36	17.83
you know	0.38	0.45
evidential	4.14	4.97
degree adv.	6.09	9.37
Involvement	***82.73***	***87.93***

Table 6.3. Network involvement scores per correspondent and per period, expressed in number of tokens per 1000 words (HW = Horace Walpole)

Once again, a few possibly problematical issues for this model present themselves quite clearly. Firstly, because of the small size of some of the sub-corpora, some involvement features do not occur in every period for every correspondent. This makes the total involvement score, which is an aggregate of all of the involvement features' frequencies per 1000 words, less reliable: if for instance an intensifying degree adverb does not occur in a sample which is only slightly larger than 1000 words, that does not necessarily mean that its mean frequency per 1000 words in a larger sample would also be zero. This is a clear drawback of working with very small sub-corpora.

Secondly, as noted in chapter 4, because the data for the network model are taken from the same textual sources as will be studied for linguistic properties in light of the network make up, there is a risk of circular reasoning. Also, some of the features are subject to linguistic change in the period or thereafter, and text-type may also influence the occurrence and frequency of some of the linguistic parameters. For all these reasons, I feel a linguistic analysis like this should always be combined with other data, as I will argue in section 6.5.2. as well.

Thirdly, because the involvement model draws wholly on linguistic data, it is impossible to devise an involvement score for all network members relating to each of the other network members for all three periods of time, which was indeed possible in the network strength scale using biographical background information for periods from which no letters exist (see 6.3.1). The gaps which are left by the lack of linguistic data for certain network members during particular periods of time become more clearly visible when the total involvement scores are ordered in the same way as for a classical network strength analysis into what I have called an involvement strength scale, in Table 6.4.

Correspondent A	Correspondent B					
	Correspondent and period		*Gray*	*West*	*Ashton*	*Walpole*
	Gray	I II III		73.44		100.57 82.73
	West	I II III				107.62 118.33
	Ashton	I II III				
	Walpole	I II III	43.86 87.93	86.42 72.23	75.28	

Table 6.4. Involvement strength scale (B is scored for the total involvement as expressed in A's letters to B in that specific period)

We see, for instance, that it is impossible to devise a reliable score for Ashton's language due to the very limited data, less than 1000 words, we have of his language and the same holds for the single last letter from Walpole to West from period III. However, in leaving so many gaps, the model using involvement is in a way more suited to the linguistic variation it is supposed to help explain, since both analyses are purely linguistic: a network analysis based on involvement in this fashion will show the same gaps in the data as an analysis of a linguistic feature within the same network.This may be considered a positive side to the circularity of reasoning which checking linguistic data against a linguistically based model entails. Comparable to classical network strength scores, a higher involvement score signals the possibility of greater linguistic influence. I will therefore now take a closer look at the total involvement scores for this model and what their implications are for predicted

linguistic influence in the network, as was done for the classical network strength scale.

Firstly, the asymmetry which was absent from the network strength scale, perhaps due to interpretational difficulties of the relatively sparse background information, is clearly present in the results from the application of the involvement model. If we are to interpret a high involvement score as a high network strength score, we see that the current model suggests that Gray, whose involvement score towards Walpole is almost twice as high as vice versa, is possibly influenced by Walpole in Period I. This is in agreement with Fitzmaurice's remark that "[i]t may be rare for an interpersonal tie to be perceived in the same way by both of its actors", a contrast which is "captured in the notions of asymmetry and reciprocity" (Fitzmaurice 2000b: 271). As discussed in chapter 4, recipients of non-reciprocal ties very likely present the route by which social influence is transmitted. The convergence of linguistic choices between two correspondents can in that way be said to reflect the desire of one correspondent to be approved of by the other. Translating this into the involvement model leads us to expect that in the case of asymmetrical relationships, influence is likely to travel from the person who is less involved to the person who is more involved in the relationship. However, more extensive linguistic analysis is needed to provide more evidence for ideas on asymmetrical relationships and linguistic influence, and discussion of influence is therefore still necessarily highly theoretical and hypothetical.

In Figure 6.4. we see that in Period I Walpole is less involved with, and less connected to Gray than Gray is to Walpole, and possible linguistic influence is therefore expected to travel from Walpole to Gray: Gray is expected to accommodate (either consciously or subconsciously, see Bax 2002)

to Walpole's language. The relationship with West seems to be more equal. As noted above, no score is available for Ashton due to lack of linguistic material.

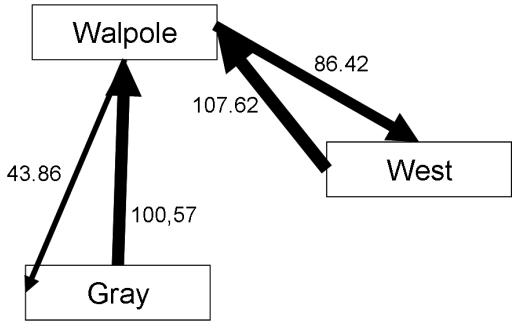

Figure 6.4. The Eton Network Cluster during Period I: April 1734 – March 1739.

For Period II, represented in Figure 6.5., we see asymmetry between West and Walpole. This asymmetry leads us to expect an influence of Walpole on West. There is also asymmetry between Walpole and Ashton and Gray and West but this is caused by the non-existence of letters from Ashton to Walpole and West to Gray, and therefore we cannot presuppose any direction of influence.

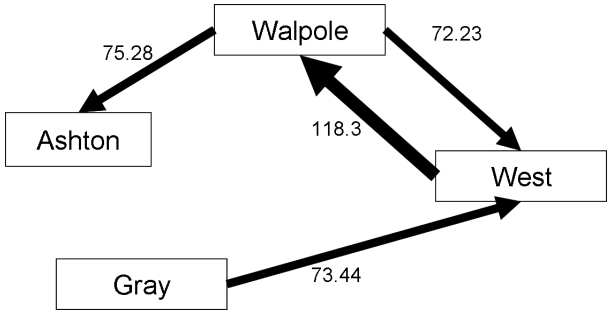

Figure 6.5. The Eton Network Cluster during Period II: March 1739 – July 1741.

The possibility should be considered though that the lower involvement scores may not be entirely due to social network related causes. We see that

Walpole's involvement with West in Period II, during the European tour, decreases. However, both Walpole and Gray's writing in Period II are less involved than that of West, as one would perhaps expect with the greater physical distance between the correspondents and the lack of day-to-day contact. It may also be due to influence from other external factors, such as the fact that the content of the letters from Period II consists mostly of travel descriptions, that the writing of the two travellers shows less personal and interpersonal involvement during Period II. As I noted before, the linguistic make up of a text is also influenced by its genre, and travel writing might have to be considered a different kind of text type than personal correspondence. It is not unikely that the letters from abroad therefore have different linguistic properties, which influence the results of the involvement analysis. Notwithstanding these considerations, West's asymmetrically high involvement is in line with the idea that he wants to belong to a group to which he does not belong: the travellers. This would make linguistic accommodation by West to Walpole more likely.

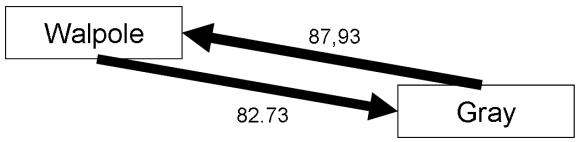

Figure 6.6. The Eton Network Cluster during Period III: July 1741– March 1771

Period III, visualised in Figure 6.6., shows an almost symmetrical relationship between Gray and Walpole with a relative closeness which suggests some possible reciprocal influence. The one remaining letter between Walpole and West does not contain enough linguistic data to devise a reliable involvement score for Walpole towards West in Period III.

In the previous discussion I have illustrated a third problem which is linked to the data and models that sociohistorical linguists use: the influence of external factors other than social network integration on linguistic reality. The analysis of the network above, which is rather general, shows that the linguistic basis of the involvement model makes the model more suited to the data and less dependent on interpretation, but at the same time also more open to the influence of other external factors on linguistic reality. Extra-linguistic factors other than social network integration may have an influence on the occurrence of involvement markers in the language of the correspondents; linguistic changes (such as the subjectification of evidential verbs, see Brinton 2006) and the influence of text-type are only two possibilities. Speech and writing are two radically different media. Chafe (1985) focuses on the differences between spoken and written language on the basis of features like involvement, and states that in spoken discourse "[t]here is involvement with the speaker's own ego, with the process of interaction with the hearer, and with subject matter", whereas "[w]ritten language lacks these manifestations of involvement" (Chafe 1985: 122). He also notes, however, that "[t]hese generalizations apply best to the extremes of spoken and written language" (Chafe 1985: 122), which is supported by Palander-Collin's remark that "Biber and Finegan (1989, 1997) … addressed register variation and identified linguistic features characteristic of different written and speech-based genres … showing that personal letters contain a high number of so-called involved features" (Palander-Collin 1999b: 129). Sairio argues that "[i]n a later study by Chafe and Danielewicz (1987: 107, 111) personal letters were seen to show the highest amount of ego involvement when compared with conversations, lectures and academic papers" (Sairio 2005: 24).

However, although personal letters are more *speechlike* than other genres of writing, letters are still a fundamentally written medium. As Tieken-Boon van Ostade (2009:122) notes, "Letters [...] are *not* speech, and should be treated accordingly". Furthermore, it would be unwise to treat personal letters as a single text type, especially because of their varied forms and subject matter (see for instance Görlach 2004: 23-88). In the involvement analysis for the Walpole network during Period II, I argued that the differences in the levels of involvement between Gray and Walpole on the one hand and West on the other may have to do with the content and perhaps the genre of their letters. Therefore, even while using a purely linguistic model, for historical linguistic analysis one always needs to take notice of contextual information and subject matter. Involvement analysis can be a very useful tool, I believe, in predicting or explaining linguistic influence, but perhaps not in a completely straightforward way. It is important to watch out for overgeneralising influences on language such as text type and underestimating the influence of other extra-linguistic factors when devising a model of linguistic influence which is itself linguistically based.

6.4. Linguistic analysis and evaluation of results

After having looked at the material and the two different models, and having noted some of the problems and possibilities of their application, I will now turn to the language in the present corpus in order to see whether the predictions about possible linguistic influence in the Walpole Eton Network Cluster, based on the two models, are in line with the linguistic data. To this end, I have carried out an analysis of variation in the use of *be* and *have* in the perfect with mutative intransitive verbs such as *I am come* and *he was gone to town*. As already discussed in chapter 5, during the eighteenth century a shift

occurred from the predominant use of *be* to *have* as in modern usage and I expect to find evidence of this development in the language of the Walpole Eton Network Cluster as well. Rydén and Brorström describe a levelling development in the *be/have* paradigm for mutative intransitives during the LModE period (1700–1900), which led to "an almost complete *have* dominance" in the nineteenth century (1987: 197). For this analysis I have adopted the list of verbs from Rydén / Brorström (1987: 234–265). Table 6.5. below shows the results of the analysis of *be/have* variation with these verbs. Once again, the results show a number of gaps, and a number of zero and very low token counts. Similar results were obtained for the Walpole Family Network in chapter 6, which probably has to do with the relatively small corpus of letters which is available for the analysis carried out there. For now, I will only discuss the results of the analysis in the context of the focus of the present study, asking the following question: can the social network model (in two different manifestations) be successfully used to explain linguistic variation in a network context and at the level of the individual?

The results of my analysis of *be/have* variation which were obtained for the Walpole Eton Network cluster can now be compared to the expectations raised by the two different models as discussed above. For Period I, the classical network strength scale predicted possible linguistic influence between all correspondents. The involvement model led to expectations of possible influence from Walpole upon Gray as well as linguistic convergence for all correspondents. Unfortunately, the analysis of *be* and *have* seems insufficient for a reliable comparison and test of the models, for this period at least, since there are no data from Walpole. Therefore, any influence which includes Walpole cannot be tested, and since there are no letters from, for

example, Ashton to West or West to Gray in Period I, the prediction of all-round linguistic convergence cannot be tested either.

	Period I: 1734–1739			
	Number of tokens		Percentage of use	
	be	*have*	*be*	*have*
Gray to West				
Gray to HW	9	3	75	25
West to HW	1*	2*	33.33*	66.67*
HW to Gray	0*	0*	0*	0*
HW to West	0*	0*	0*	0*
Total	*10*	*5*	*66.67*	*33.33*

	Period II: 1739–1741			
	Number of tokens		Percentage of use	
	be	*have*	*be*	*have*
Gray to West	1*	1*	50*	50*
Gray to HW				
West to HW	2*	2*	50*	50*
HW to Gray				
HW to West	15	4	78.95	21.05
Total	*18*	*7*	*72*	*28*

	Period III: 1741–1771			
	Number of tokens		Percentage of use	
	be	*have*	*be*	*have*
Gray to West				
Gray to HW	11	8	57,89	42.11
West to HW				
HW to Gray	5	4	55.56	44.44
HW to West				
Total	*16*	*12*	*57,14*	*42.86*

Table 6.5. Variation in the use of *be* and *have* in perfective constructions with mutative intransitives (numbers of tokens which are too small to draw reliable conclusions about usage have here been marked with an asterisk)

In the second period the network strength analysis predicted influence from Walpole and Gray on West and Ashton, as well as linguistic convergence within the two separate groups. The involvement model also predicted an influence of Walpole on West. Because the linguistic analysis yields no results

for Walpole and Gray (since there are no letters between them in the CHWC), we cannot test any of these claims. The Thomas Gray Archive[2] lists a number of unlocated letters and letters which are not extant, as well as one French letter from West to Gray and/or Walpole, as well as seven letters from Gray to Ashton which I have not been able to study and some of which are in private collections.

West's usage in Period II shows an equal preference for *be* and *have*, while there was a tentative *have* preference in Period I (two instances of *have* versus one of *be*). Walpole shows a clear preference for *be*, so this could be seen as West adapting to Walpole, which is what is expected in both models. However, West's results are based on only three tokens in the first period and another four in the second period, so no significant claims can be made here.

For Period III we see a convergence in the usage of Walpole and Gray (their usage percentages are virtually the same), which confirms the predictions from both models. Moreover, both men show a decreasing preference for *be*, towards the modern usage of *have* in these constructions. Gray's use in period II seems somewhat more modern than in period III, albeit a tentative conclusion, regarding the very low number of instances. Periods I and III are the only periods for which a slightly larger number of tokens is available, I have therefore disregarded the asterisked data in Table 6.5. There could be several reasons for this besides influence on each other. Walpole and Gray may, for instance, also have been influenced by the publication of normative grammars (though one may wonder whether they belonged to the

[2] The Thomas Gray Archive is, according to the website "a fully browseable, searchable and annotated digital archive of the life and works of Thomas Gray (1716-1771)" , which is currently housed at the Bodleian Library at Oxford University <http://www.thomasgray.org/>.

intended audience of those books) which condemned the use of *be* in these types of constructions.

The seemingly disappointing results of this analysis may still provide some new questions and directions that will be relevant for continuing research along these lines. It would, for instance, be interesting to make a further separation in Period III, somewhere in the middle of the eighteenth century, representing the time before and after the rise in the publication of normative grammars such as Lowth's (1762), in order to look at whether the normative linguistic environment which is represented by the grammars led to a decrease in the usage of the construction with *be*. Conversely, it might be asked whether the language of the upper class in the middle of the eighteenth century may actually have been a model for the norm as it was written down in the grammars. McFadden (2007) for instance, who mainly concentrated in his paper on the linguistic context of the variation in usage of *be* and *have* with mutative intransitives, mentions the decrease of the use of *be* in mutative intransitive perfects in the second half of the eighteenth century as a puzzle that cannot be solved by purely linguistically driven change. Could it be more than a coincidence that this puzzling change coincides with the time in which the publication of so many grammars allowed a normative influence to make itself felt?

6.5. Suggestion for further research: the combination model

It has been demonstrated in the discussion of the methods and the several case studies above that the models for network strength analysis that have been used hitherto have all had their own challenges: concerning their fit with the data, the reliability of the information that was needed to be able to analyse the network, the influence of external factors on the results and the

reliance of the models on the subjective interpretation of background information. As a suggestion for further research I would like to make the case for a combination model. With the suggested model I aim to avoid these issues as much as possible, for instance by combining sociometric with cognitive data, as suggested by Fitzmaurice (2000a: 205). The basis of the NSS goes back to Milroy (1987), but comes more directly from Sairio (2009b). Sairio's model was based on Fitzmaurice (2007) and also took some suggestions from Henstra (2008, see chapter 5), which in turn was based on Bax (2000). Following Fitzmaurice (2007) I use a 5-point scale ranging from 0 to 4, for two reasons. Firstly, unlike was attempted in Henstra (2008, see chapter 5) there is no need to use negative scores for the more distant relationships, which complicated the statistical analysis of network strength a great deal. Secondly, the greater differences in scores between network members, created by a broader bandwidth of scoring the ties, simplifies interpretation by enlarging the differences between the network members, it also makes it possible to compare the classical NSS with involvement scores which have been recalculated to a 5-point scale. What I shall call the *first layer* of analysis can be found in Table 6.6. below.

I recall here Milroy (1987) saying that the chosen indicators must "must reflect … conditions which have repeatedly been found important in a wide range of network studies, in predicting the extent to which normative pressures are applied by the local community" and that "[t]hey must be recoverable from data collected in the field and easily verifiable" (Milroy 1987: 141). The indicators in the model in Table 6.6. are based on previous research by Bax (2000), Fitzmaurice (2007), Sairio (2005, 2008, 2009a) and myself (2008, 2009, the current study) and have been shown by Sairio (2009b) to be at least in some sense effective measures of network strength. The interpretation of

these parameters will differ for each network: in this case I have adapted some terms for use with Walpole's network, and in turn they may be adapted again for other networks. These adaptations will be discussed below. For the other terms I follow Sairio (2009b: 149-152).

1. Same domicile		
	yes often (e.g., during the season) rarely (e.g., abroad) no	4 points 3 points 1 point 0 points
2. Type of relationship		
	intimates kin acquaintances not acquainted	4 points 3 points 2 points 0 points
3. Same social circle		
	yes: primary yes: secondary no	4 points 2 points 0 points
4. Professional collaboration		
	yes: balanced/"giver" yes: "receiver" no	4 points 2 points 0 points
5. Social status		
	equals superior inferior	4 points 2 points 0 points
6. Age		
	same generation older generation younger generation	4 points 2 points 0 points
7. Gender		
	same other	4 points 0 points
8. Previous network connection		
	yes no	2 point 0 points

Table 6.6. The proposed NSS. Layer 1: functional analysis (based on Sairio 2009b: 149-152)

An option that I have added for the Walpole network in the parameter "relationship type" is kinship. Friendship and kinship are somewhat problematical notions for the eighteenth century, as was noted in chapter 5 and also in Sairio (2009b: 149). However, since some of Walpole's correspondents are family members and others are not, this is a distinctive paramterer which could not be ignored in the analysis. The parameter "same social circle" I have defined on the level of the network clusters as identified in the Walpole correspondence and "professional collaboration" may be seen as incorporating all types of collaboration in coalition-like associations, political alliances as well as collaborative writing projects.

Self involvement	
(a)	first person pronoun use
(b)	evidential constructions with *think, know, believe, suppose, find, be sure* and *doubt*
Hearer involvement/addressee inclusion	
(a)	second person pronoun use
(b)	nominal third person reference to addressee: ubiquitous "you know"
Subject involvement	
(a)	intensifying degree adverbs

Table 6.7. The proposed NSS. Layer 2: Linguistic involvement

This classical NSS is then superposed with a *second layer* of network strength analysis, in which an analysis of involvement features in the language of the correspondents is carried out, following Sairio (2005), Palander-Collin (1999a, 1999b) and Henstra (2009, see also this chapter). The NSS is quite simply a reflection of the normalised frequencies of linguistic tokens of involvement, and their occurrence in the language of the correspondents per 10,000 words. The analysed features have been shown to reflect involvement

strategies in previous research such as Chafe (1985), Palander-Collin (1999a, 1999b and 2009) and Sairio (2005). The background of these features was discussed in more detail above. A summary can be found in Table 6.7. above.

Tokens for all these involvement features are added up for each network tie and then normalised per 10,000 words. In Table 6.8. this is done with the data from the correspondence between Walpole and his Eton friends as found in Tables 6.3. and 6.4. above:

Period I	*Gray to HW*	*West to HW*	*HW to Gray*	*HW to West*
Involvement	**10.06**	**10.76**	**4.34**	**8.64**
Period II	*Gray to West*	*West to HW*	*HW to West*	*HW to Ashton*
Involvement	**7.34**	**11.83**	**7.22**	**7.52**
Period III	*Gray to HW*	*HW to Gray*		
Involvement	**8.27**	**8.79**		

Table 6.8. Involvement scores for Gray, West, Ashton and Walpole, expressed in number of tokens per 10,000 words (HW = Horace Walpole), based on Table 6.3.

The scores are then re-calculated to fit a 5-point scale (from 0 or – to 4), to make them easier to compare to the scores given in the first layer: the classical NSS. I have chosen this method to create a universal fit for the very different ranges of involvement that different text types, times, and authors will have.

- No involvement (i.e. no extant letters) equals a – or 0 (not shown in Table 6.9.)
- The maximum involvement score (*M*) in Table 6.8. is 11.83
- The lowest involvement score (*m*, noting that *m*>0) in Table 6.8. is 4.34

The formula for calculation of the involvement scores (*I*) is then the following:

$$\frac{N-m}{(M-m)/3}+1=I$$

The involvement score for Gray to Walpole in Period I, for instance, is calculated as follows:

$$\frac{10.06 - 4.34}{(11.83 - 4.34)/3} + 1 = 3.29$$

When all scores in the network are recalculated this way, and rounded off to the nearest whole number, this leads to the following involvement scores, based on Table 6.8.:

Period I	*Gray to HW*	*West to HW*	*HW to Gray*	*HW to West*
Involvement	**3**	**4**	**1**	**3**
Period II	*Gray to West*	*West to HW*	*HW to West*	*HW to Ashton*
Involvement	**2**	**4**	**3**	**2**
Period III	*Gray to HW*	*HW to Gray*		
Involvement	**3**	**3**		

Table 6.9. Layer 2. Involvement scores for Gray, West, Ashton and Walpole, expressed on a 5-point scale (scores between 0 / – and 4)

The highest involvement score is used as the basis for the calculation of the top of the scale so that scores are never higher than four points. The lowest involvement score is used as the basis for calculating the score of one, so that scores are never lower than one (if no involvement tokens are found whatsoever, or if there is no correspondence between two network members a score of zero or a – is noted, to mark a clear difference between low involvement and no involvement). In this way the scores within a network can also always be related to scores calculated for another network in the same way, because they are relative scores expressing *degree* of involvement as compared to the other network members, rather than as an *absolute* score. This also provides some relief for the influence of changes in the language over time, such as the subjectification of the evidential verbs used to calculate

involvedness (see Brinton 2006 and Biber *et al.* 1999). Rounding off the scores creates another margin for error of course, and one may choose to round off the numbers to half points, for instance, for greater precision. However, as may be seen from chapters 5 and 6 above, it is difficult to back a very fine-grained network analysis with statistically significant results.

As a final step we can now combine the classic NSS and the Involvement network scores into the following Combination Network Strength Scale:

Correspondent A		**Correspondent B**							
Correspondent and period		***Gray***		***West***		***Ashton***		***Walpole***	
		NSS	***Inv.***	***NSS***	***Inv.***	***NSS***	***Inv.***	***NSS***	***Inv.***
Gray	I	n/a	n/a	3	–	5	–	**5**	**3**
	II			2	2	2	–	6	–
	III			–	–	–2	–	4	3
West	I	3	–	n/a	n/a	3	–	3	4
	II	2	–			3	–	**2**	**4**
	III	–	–			–	–	–	–
Ashton	I	5	–	5	–	n/a	n/a	6	–
	II	2	–	3	–			2	–
	III	–2	–	–	–			–2	–
Walpole	I	**5**	**1**	3	3	6	–	n/a	n/a
	II	6	–	2	3	2	2		
	III	4	3	–	–	–2	–		

Table 6.10. The Combination Network Strength Scale for Walpole, Gray, West and Ashton (based on Tables 6.2. and 6.9.)

We see that only a few of the Involvement scores correspond exactly with the classic NSS scores, these cells have been highlighted in the darker shade of grey. Such a similarity may be seen as a confirmation of both the NSS and Involvement analysis. More scores, however, are relatively close to each other (the difference is 1 point), these have been highlighted in a lighter shade of

grey and may be seen as a tentative confirmation of NSS and Involvement scores. It may be helpful to look at the Involvement scores at a higher livel of significance, for instance rounded off at two digits to see if the difference is then smaller or greater. Finally, there are a number of cases in which the difference is rather large (<1), in one case even 4 points. I suggest further research is needed to see if there is perhaps something interesting going on, or if the method of collecting data for the involvement scores needs to be finetuned. However, considering that there are ten instances in which a combination of NSS and Involvement data for the same correspondent and period exist, a tally of 30% identical scores and another 40% similar scores (difference of no more than 1 point) seems like a good start, and I think the combination model shows promise for use as an objectifying tool within historical social network analysis when developed and tested more fully.

6.6. Concluding remarks

In this chapter I have shown that the two versions of a social network model which I drew upon for the analysis of my data for the Walpole Eton Network Cluster here both have their problems, but also that they have distinctive advantages and value for historical sociolinguistic research. The problems which were addressed along the way may be summarised as follows: firstly, there is the underrepresentation and at the same time overrepresentation of certain authors in focused corpora. These corpora are more suited to social network analysis with its dependence on background knowledge (which is more easily gathered for smaller network clusters). They are, however, often unbalanced, and cannot easily be used for more generalised research. Furthermore, focused corpora run a greater risk of not containing enough linguistic data to find statistically significant results in a linguistic analysis.

Secondly, a classical network strength scale (as well as any other model which depends on contextual information for filling gaps in the linguistic information) is open to interpretational difficulties. As a researcher one always runs the risk of overinterpreting information and of applying a tunnel vision towards the desired result when dealing with this kind of analysis. Thirdly, linguistically based analyses such as those drawing on the involvement model are very much open to the influence of other extra-linguistic factors on linguistic reality and to linguistic change. A genre such as personal letters cannot be seen as stylistically and linguistically homogeneous and there is a risk of overgeneralising the specific linguistic characteristics of these different text types when they are put together in a linguistically based analysis (see for example Biber 1999: 133; 146; 148 on the importance of register variation). Furthermore, circular reasoning is a serious consequence of using linguistic data to predict linguistic change or usage. I believe that all these problems can be taken under one heading: there seems to be a mismatch between the type of research sociolinguistic models make possible, namely a very specific, micro-level network analysis, and the type of research for which our historical linguistic data allow.

I have provided a suggestion for further research in the form of a combination model, in which a layered model provides us with a more complex representation of the truly complex reality than the classical NSS or the linguistic involvement model can, as was found in a wide variety of historical network analyses presented in chapters 3, 4, 5 and 6. It provides the necessary *double check* for the researcher, who cannot just trust a single-tier analysis. The social network model has proved to be an enticing model for explaining linguistic variation on a micro-level. However, I have shown that a successful and statistically sound application of it using historical data is difficult. The

combination model provides a much needed objectified view of the subjective and flawed measuring methods available to us, which invite anecdotal use of the model, rather than theoretically sound applications.