Appendix 2 Answers to exercises

Note: In numerical questions, where more than one example of a particular method is given, the later examples are not necessarily worked through in full detail.

Chapter 1

2 (i) Ratio: time scale has 'absolute zero'.
  (ii) Ordinal: pleasantness scale cannot be assumed to have precisely equal intervals.
  (iii) Nominal: either there is, or there is not, a finite verb.
  (iv) Ordinal: we cannot claim equality of intervals between points on the grammaticality scale.
  (v) Ratio: there is an 'absolute zero' (no sentences remembered correctly) and equality of intervals on the scale.

Chapter 2

1 It will usually be found that word length distributions are positively skewed. The precise location of the peak of the curve may vary according to the nature of the text: for instance, certain types of written language have a high proportion of two-letter prepositions such as *in*, *on*, *of*, *at*, *to*, *by*, or a high proportion of two- and three-letter pronouns such as *we*, *me*, *us*, *she*, *him*, *her*, *you*.

2

	Frequency			
Intensity	*Stressed*		*Unstressed*	
7–9			/	1
10–12			/	1
13–15			卌 /	6
16–18	///	3	卌 ///	8
19–21	卌 卌 卌 ////	19	卌 卌 卌 卌	20
22–24	卌 卌 卌 卌 卌 ///	28	卌 卌 卌 卌 卌 //	27
25–27	卌 卌 卌 卌 卌 卌 /	31	卌 卌 卌 卌 ///	23
28–30	卌 卌 ////	14	卌 卌 ////	14
31–33	卌	5		
		100		100

The polygons corresponding to these values show that the differences between the distributions are quite small, and this corroborates the view that stress may not be a quantitatively important phenomenon in French. Both curves are fairly symmetrical about their highest point. The peak for stressed syllables is at a slightly higher intensity than that for unstressed syllables. The range of values is greater for the unstressed syllables.

3

Pause length (1/50 sec units)	*Frequency*
0–5	34
6–10	23
11–15	11
16–20	12
21–25	31
26–30	20
31–35	14
36–40	4
41–45	1

The histogram will show two peaks, one at very low pause lengths, the other at 21–25 time units, suggesting two functionally different types of pause. The next stage of the investigation would be to examine the environments of the pauses in each category, to see whether any generalisations could be made.

Chapter 3

1 (iii) Since the data are of the ratio type, and do not appear to contain any highly untypical values, the mean is an appropriate measure of central tendency.

$$\bar{x} = \frac{\Sigma x}{N} = \frac{602}{30} = 20.07 \text{ msec.}$$

$$s = \sqrt{\frac{\Sigma x^2 - (\Sigma x)^2/N}{N-1}} = \sqrt{\frac{12\,482 - 602^2/30}{29}}$$

$$= 3.72 \text{ msec.}$$

2 Since word length distributions are usually positively skewed, the median would be the most appropriate measure of central tendency in most cases.

3 The exact values calculated will depend on whether the raw data or grouped data have been used. The following calculations are based on the class intervals suggested in the answer to question 2 of chapter 2.

Stressed syllables

$$\bar{x} = \frac{\Sigma fx}{N}$$

where f = frequency in a given class
x = mid-point of class interval

$$\bar{x} = \{(3 \times 17) + (19 \times 20) + (28 \times 23) + (31 \times 26) + (14 \times 29) + (5 \times 32)\}/100$$

$$= \frac{2\,447}{100} = 24.47 \text{ units}$$

($\bar{x}$ for raw data = 24.31 units)

Since there are 50 observations in the range up to 24 units, and 50 in the range from 25 units upwards, the median is the mid-point of the combined classes 22–24 and 25–27, i.e. 24.50 units. The mode is the 25–27 class, whose mid-point is at 26 units. The range is from the 16–18 class to the 31–33 class (from the raw data the range is from 16 to 31, i.e. 15 units). The standard deviation is given by:

$$s = \sqrt{\frac{\Sigma fx^2 - (\Sigma fx)^2/N}{N-1}} = \sqrt{\frac{61\,129 - 2\,447^2/100}{99}}$$

$$= 3.55 \text{ units.}$$

(for raw data also, $s = 3.55$ units).

Unstressed syllables

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{2\,264}{100} = 22.64 \text{ units for grouped data}$$

($\bar{x}$ = 22.43 units for raw data).

$$\text{median} = L + \left(\frac{N/2 - F}{f_m}\right)h = 21.5 + \left(\frac{50-36}{27}\right) \times 3$$

$$= 23.06 \text{ units.}$$

mode = 22–24 class, mid-point 23 units

range = 7–9 to 28–30 (or 23 units from the raw data)

$$\text{standard deviation} = \sqrt{\frac{\Sigma fx^2 - (\Sigma fx)^2/N}{N-1}}$$

$$= \sqrt{\frac{53\,278 - 2\,264^2/100}{99}}$$

$$= 4.52 \text{ units, for grouped data}$$

(s = 4.39 units for raw data).

These results confirm our conclusions from frequency polygons (question 2 in chapter 2): the measures of central tendency are slightly higher, but the variability is lower, for the stressed syllables. The mean, median and mode are close together, reflecting the symmetrical nature of the distribution.

4

	Experimental group	*Control group*
Mean	38.60	38.60
Standard deviation	21.96	8.14

Although the average performance of the two groups is identical, the variability is much greater for the treated group than for the control group, suggesting that the treatment considerably improves the performance of some aphasics, but is deleterious to others.

5 The median is the most appropriate measure of central tendency, since the data are of the ordinal type. The frequency distributions are as follows:

	Frequency for	
Politeness rating	*Sentence 1*	*Sentence 2*
1	23	9
2	17	16
3	6	13
4	2	11
5	1	1
6	0	0
7	1	0

$$\text{median for sentence 1} = L + \left(\frac{N/2 - F}{f_m}\right)h$$

$$= 1.5 + \left(\frac{50/2 - 23}{17}\right) \times 1$$

$$= 1.62.$$

Median for sentence 2 = 2.50, since there are 25 ratings of 1 or 2, and 25 of 3 or above, so that the median clearly lies half way between 2 and 3.

Chapter 4

1 (ii) The histogram will be quite symmetrical.

$\bar{x} = 49.88$ marks

$s = 15.53$ marks

$\bar{x} \pm s = 49.88 \pm 15.53 = 34$ to 65 marks

$\bar{x} \pm 2s = 49.88 \pm 31.06 = 19$ to 81 marks

% of observations lying in range 34–65 marks

$= 35/50 \times 100$

$= 70\%$

% of observations lying in range 19–81 marks

$= 48/50 \times 100$

$= 96\%$.

Comparing these figures with those for the normal curve (68.2 and 95.4 per cent, respectively), we see that the data correspond closely to the predictions for a normal distribution.

(iii) (a) We have

$$z = \frac{x - \bar{x}}{s} = \frac{40 - 49.88}{15.53} = -0.64.$$

From table A2, 0.2611 of the area under the curve lies below this z value. Thus, the proportion of marks expected to fall below 40 is 26.1 per cent.

(b) For 60 marks, we have

$$z = \frac{60 - 49.88}{15.53} = 0.65.$$

From table A2, 0.2578 of area under curve lies above this value. For 70 marks,

$$z = \frac{70 - 49.88}{15.33} = 1.30.$$

0.0968 of area under curve lies above this. Thus, the proportion of marks expected to lie between 60 and 70 is

$$0.2578 - 0.0968 = 0.1610 = 16.1\%.$$

2 (i) For 10 dB,

$$z = \frac{x - \bar{x}}{s} = \frac{10 - 23.4}{5.8} = -2.31.$$

The proportion lying below this is 0.0104, or 1.0 per cent.

(ii) For 30 dB,

$$z = \frac{30 - 23.4}{5.8} = 1.14.$$

The proportion above this is 0.1271, or 12.7 per cent.

(iii) For 20 dB,

$$z = \frac{20 - 23.4}{5.8} = -0.59.$$

The proportion below this is 0.2776. For 25 dB,

$$z = \frac{25 - 23.4}{5.8} = 0.28.$$

The proportion above this is 0.3897. Therefore the proportion between 20 dB and 25 dB is

$$1 - 0.2776 - 0.3897 = 0.3327 = 33.3\%.$$

3 (i) The histogram is quite symmetrical.

(ii) $\bar{x} = 53.30$ marks.

$s = 16.62$ marks.

(iii) $\bar{x} \pm s = 53.30 \pm 16.62 = 37$ to 70 marks.

$\bar{x} \pm 2s = 53.30 \pm 33.24 = 20$ to 87 marks.

% in range 37–70 = 34/50 × 100 = 68%

% in range 20–87 = 47/50 × 100 = 94%.

The percentages agree very well with those for the normal distribution.

Chapter 5

1 $$\text{standard error} = \frac{s}{\sqrt{N}} = \frac{6.87}{\sqrt{100}} = 0.69 \text{ words.}$$

The 95 per cent confidence limits are

$$\bar{x} \pm 1.96 \times \text{standard error} = 14.21 \pm (1.96 \times 0.69)$$
$$= 14.21 \pm 1.35$$
$$= 12.86 \text{ to } 15.56 \text{ words.}$$

2 (i) $\bar{x} = 26.05$ msec.

(ii) $s = 6.26$ msec.

(iii) standard error = $6.26/\sqrt{20} = 1.40$ msec.

(iv) Since the sample is small, the t distribution must be used. From table A3, the critical value for the 5 per cent level and $(20-1) = 19$ df is 2.093. The 95 per cent confidence limits are thus given by

$$\bar{x} \pm 2.093 \times \text{standard error} = 26.05 \pm (2.093 \times 1.40)$$
$$= 26.05 \pm 2.93$$
$$= 23.12 \text{ to } 28.98 \text{ msec.}$$

The confidence limits are rather wide apart, because of the high standard deviation relative to the mean.

3 (i) $\bar{x} = 0.51$ sec.

(ii) $s = 0.16$ sec.

(iii) standard error = $0.16/\sqrt{30} = 0.03$ sec.

(iv) 99 per cent confidence limits are:

$$\bar{x} \pm 2.58 \times \text{standard error} = 0.51 \pm (2.58 \times 0.03)$$
$$= 0.51 \pm 0.08$$
$$= 0.43 \text{ to } 0.59 \text{ sec.}$$

4 proportion of auxiliaries $= p = 63/300 = 0.21$.

$$\text{standard error} = \sqrt{\frac{p(1-p)}{N}} = \sqrt{\frac{0.21(1-0.21)}{300}} = 0.02.$$

95 per cent confidence limits are:

$$p \pm 1.96 \times \text{standard error} = 0.21 \pm (1.96 \times 0.02)$$
$$= 0.21 \pm 0.04 = 0.17 \text{ to } 0.25$$
$$= 17\% \text{ to } 25\%$$

5 error tolerated $= 0.01 = 1.96 \times$ standard error.

Thus

standard error $= 0.01/1.96$.

But

standard error $= \sqrt{\frac{p(1-p)}{N}}$ and $p = 0.21$ (from question 4)

Therefore

$$\sqrt{\frac{0.21(1-0.21)}{N}} = \frac{0.01}{1.96}$$

$$\frac{0.21(1-0.21)}{N} = \left(\frac{0.01}{1.96}\right)^2$$

$$N = \frac{0.21(1-0.21)}{(0.01/1.96)^2} = 6\,373.$$

We thus need a sample of about 6 400 finite verbs to estimate the proportion of auxiliaries to within 1 per cent.

Chapter 7

1 (i) *Text A*

$\bar{x} = 311.00$ word types.

$s = 38.52$ word types.

standard error $= 38.52/\sqrt{50} = 5.45$ word types.

99 per cent confidence limits are:

$$\bar{x} \pm 2.58 \times \text{standard error} = 311.00 \pm (2.58 \times 5.45)$$
$$= 296.94 \text{ to } 325.06 \text{ word types.}$$

Text B

$\bar{x} = 346.64$ word types.

$s = 37.21$ word types.

standard error $= 37.21/\sqrt{50} = 5.26$ word types.

99 per cent confidence limits are:

$$\bar{x} \pm 2.58 \times \text{standard error} = 346.64 \pm (2.58 \times 5.26)$$
$$= 333.07 \text{ to } 360.21 \text{ word types.}$$

(ii) Since the samples are large, we can use the z-test.

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{N_1} + \dfrac{s_2^2}{N_2}}} = \frac{311.00 - 346.64}{\sqrt{\dfrac{38.52^2}{50} + \dfrac{37.21^2}{50}}} = -4.71.$$

The critical value for $p \leqslant 0.01$ in a non-directional test is 2.58. The calculated value of z (ignoring the sign) exceeds this. The means are therefore significantly different at the 1 per cent level.

2 Since the samples are small, a t-test must be used.

	Group A	*Group B*
$\bar{x}$	51.36	47.33
s	9.33	10.42

Note that the values of s, and hence the variances, are similar. The samples are really too small for any serious assessment of normality, but the following figures are at least consistent with a normal distribution:

% in range	*Group A*	*Group B*
$\bar{x} \pm s$	64	67
$\bar{x} \pm 2s$	100	100

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2(1/N_1 + 1/N_2)}}$$

where

$$s_p^2 = \frac{\Sigma x_1^2 - (\Sigma x_1)^2/N_1 + \Sigma x_2^2 - (\Sigma x_2)^2/N_2}{N_1 + N_2 - 2}.$$

$$s_p^2 = \frac{29\,891 - 565^2/11 + 21\,032 - 426^2/9}{11 + 9 - 2} = 96.59.$$

$$t = \frac{51.36 - 47.33}{\sqrt{96.59(\frac{1}{11} + \frac{1}{9})}} = 0.91.$$

The critical value of t for the 5 per cent level and $(11 + 9 - 2) = 18$ df in a non-directional test is 2.101. Since the calculated value is lower than this, the means do not differ significantly at the 5 per cent level.

3 The t-test for correlated samples is appropriate here.

Pitch levels (arbitrary units) for:			
Sentence 1	*Sentence 2*	d	d^2
30	27	3	9
41	36	5	25
34	35	−1	1
28	30	−2	4
35	38	−3	9
39	44	−5	25
40	46	−6	36
29	31	−2	4
27	33	−6	36
33	37	−4	16
		$\Sigma d = -21$	$\Sigma d^2 = 165$

$$t = \frac{\Sigma d}{\sqrt{\dfrac{N\Sigma d^2 - (\Sigma d)^2}{N-1}}} = \frac{-21}{\sqrt{\dfrac{(10 \times 165) - (-21)^2}{10-1}}}$$

$$= \frac{-21}{\sqrt{\dfrac{1\,650 - 441}{9}}} = -1.81.$$

The critical value for t for the 5 per cent level and 9 df in a directional test is 1.833. Since the calculated value for t is lower than this, the means are not significantly different at the 5 per cent level.

4 Since the samples are small and the groups independent, the appropriate test is the t-test for independent samples. We shall assume normality and homogeneity of variance in the populations from which the samples are derived.

$$\bar{x}_1 = 53.2.$$

$$\bar{x}_2 = 60.1.$$

$$s_p^2 = \frac{28\,668 - 532^2/10 + 36\,575 - 601^2/10}{10 + 10 - 2} = 45.58.$$

$$t = \frac{53.2 - 60.1}{\sqrt{45.58(\frac{1}{10} + \frac{1}{10})}} = -2.29.$$

The critical value of t for the 5 per cent level and 18 df in a directional test is 1.734. Since the calculated value of t exceeds this, the means are significantly different at the 5 per cent level. Since $\bar{x}_1 > \bar{x}_2$, the difference is in the predicted direction.

5

$$z = \frac{p_1 - p_2}{\sqrt{p_p(1-p_p)\left(\dfrac{1}{N_1} + \dfrac{1}{N_2}\right)}}$$

where

$$p_p = \frac{f_1 + f_2}{N_1 + N_2}.$$

$$p_p = \frac{32 + 24}{100 + 100} = 0.28.$$

$$z = \frac{32/100 - 24/100}{\sqrt{0.28(1-0.28)\left(\frac{1}{100} + \frac{1}{100}\right)}} = \frac{0.08}{\sqrt{0.0040}} = 1.26.$$

The critical value of z for the 1 per cent level in a non-directional test is 2.58. Since the calculated value of z is lower than this, we cannot claim that the proportions are significantly different at the 1 per cent level.

Chapter 8

1 Since the data are probably best regarded as ordinal, and the groups are independent, the Mann-Whitney U-test is appropriate. We rank all the data in the combined groups, and then find the rank sum for the smaller group.

Group 1	*Rank*	*Group 2*	*Rank*
8	27	4	5.5
6	17	6	17
3	2	3	2
5	10.5	3	2
8	27	7	23
7	23	7	23
7	23	5	10.5
6	17	5	10.5
5	10.5	4	5.5
6	17	4	5.5
6	17	6	17
7	23	5	10.5
8	27	6	17
		5	10.5
		4	5.5
	$R_1 = 241$		

$$U_1 = N_1 N_2 + \frac{N_1(N_1+1)}{2} - R_1$$

$$= (13 \times 15) + \frac{13 \times 14}{2} - 241$$

$= 195 + 91 - 241 = 45$

$U_2 = N_1 N_2 - U_1 = (13 \times 15) - 45 = 150$

$U = \text{smaller of } U_1 \text{ and } U_2 = 45.$

The critical value of U for the 2.5 per cent level in a directional test, with $N_1 = 13$ and $N_2 = 15$, is 54. The calculated value of U is below the critical value, so we can claim a significant difference at the 2.5 per cent level. Since group 1 has more of the higher ranks (this could be confirmed by calculating the mean rank for each group), the difference is in the predicted direction.

2 Since the data are of the ratio type, and the sets correlated, the Wilcoxon signed-ranks test is appropriate. We take the differences between pairs of scores; rank them, omitting any zero differences; assign to each rank the sign of the difference it represents; and find the sums of positive and negative ranks.

Set 1	*Set 2*	*Difference*	*Rank*
8	5	3	+18
7	7	0	–
4	9	−5	−23.5
7	3	4	+21
6	4	2	+11.5
5	5	0	–
2	0	2	+11.5
6	4	2	+11.5
9	7	2	+11.5
8	6	2	+11.5
10	5	5	+23.5
5	6	−1	−3.5
4	3	1	+3.5
10	9	1	+3.5
8	10	−2	−11.5
4	4	0	–
7	5	2	+11.5
8	6	2	+11.5
6	6	0	–
5	2	3	+18
8	7	1	+3.5

Set 1	*Set 2*	*Difference*	*Rank*
7	7	0	—
6	9	−3	−18
9	3	6	+25
4	6	−2	−11.5
3	1	2	+11.5
8	7	1	+3.5
7	8	−1	−3.5
5	1	4	+21
6	2	4	+21

sum of positive ranks = 253.5.

sum of negative ranks = 71.5.

W = smaller of the two = 71.5.

The critical value of W for the 5 per cent level and 25 pairs of non-tied scores in a non-directional test is 89. Since the calculated value of W lies below this, we can claim a significant difference between the two sets of scores at the 5 per cent level.

Since the number of pairs of scores is greater than 20, we can also use the z approximation:

$$z = \frac{W - \dfrac{N(N+1)}{4}}{\sqrt{\dfrac{N(N+1)(2N+1)}{24}}} = \frac{71.5 - \dfrac{(25 \times 26)}{4}}{\sqrt{\dfrac{25 \times 26 \times 51}{24}}} = -2.45.$$

The critical value of z for the 5 per cent level in a non-directional test is 1.96. Since the calculated value of z (ignoring the sign) exceeds the critical value, we again conclude that a significant difference can be claimed at the 5 per cent level.

3 The data are ordinal and a repeated measures design is used. The sign test is therefore appropriate. We find the sign of the difference between each pair of scores, subtracting consistently.

Sentence 1	*Sentence 2*	*Sign of sentence 1 – sentence 2*
1	3	–
2	2	0
1	4	–
2	3	–
3	1	+
2	4	–
1	1	0
2	3	–
3	5	–
1	3	–
2	3	–
1	4	–
2	1	+
2	4	–
1	3	–

There are 13 pairs of non-zero differences, 2 with the less frequent (positive) sign. Thus $x = 2$. The critical value of x for the 5 per cent level and for 13 pairs in a directional test is 3. Since the calculated value of x is smaller than the critical value, we may claim a significant difference at the 5 per cent level. Since the ratings for sentence 2 are higher than those for sentence 1 in 11 out of 15 cases, the differences are clearly in the predicted direction.

4 The data are ordinal and the groups independent. The Mann–Whitney U-test is therefore appropriate.

rank sum for method A (R_1) = 182.5

$U_1 = 162.5$

$U_2 = 62.5$

U = smaller of the two U values = 62.5

Critical value for U for the 5 per cent level and $N_1 = N_2 = 15$ in a non-directional test is 64. Since the calculated value of U is smaller than the critical value, we can claim significance at the 5 per cent level. The ranks for method B are higher overall.

Chapter 9

1 In the contingency table below, the calculated expected values are given in brackets.

Word length (letters)	*Frequency in* The Colossus	Winter Trees	*Total*
1–3	3 473 (3 701.3)	3 000 (2 771.7)	6 473
4–6	3 743 (3 669.3)	2 674 (2 747.1)	6 417
7–9	1 272 (1 157.9)	753 (867.1)	2 025
10–12	285 (257.3)	165 (192.7)	450
>12	54 (41.2)	18 (30.8)	72
	8 827	6 610	15 437

$$\chi^2 = \frac{(3\,473 - 3\,701.3)^2}{3\,701.3} + \frac{(3\,743 - 3\,669.3)^2}{3\,669.3} + \dots \text{(etc.)}$$

$$= 78.86$$

$$\text{df} = (5-1)(2-1) = 4.$$

The critical value of χ^2 for the 5 per cent level and 4 df is 9.49. There is therefore a significant difference in the word length distributions at the 5 per cent level (and indeed even at the 0.1 per cent level).

2

	The Colossus	Crossing the Water	Ariel	Winter Trees	*Total*
black	27 (34.3)	28 (30.8)	46 (36.2)	26 (25.7)	127
other words	8 800 (8 792.7)	7 921 (7 918.2)	9 294 (9 303.8)	6 584 (6 584.3)	32 599
	8 827	7 949	9 340	6 610	32 726

$\chi^2 = 4.48.$

The critical value of χ^2 for the 5 per cent level with 3 df is 7.82. We cannot reject the null hypothesis of even distribution of items containing *black*.

3

Punctuation mark	χ^2	*df*	*Critical value*	*Significant at 5% level*
Full stop	19.20	3	7.82	Yes
Comma	26.50	3	7.82	Yes
Colon	32.40	3	7.82	Yes
Semi-colon	7.90	3	7.82	Yes
Exclamation mark	39.09	3	7.82	Yes
Question mark	86.25	3	7.82	Yes
Dash	19.15	3	7.82	Yes

Examination of the appropriate contingency tables will reveal that the earliest volume of poetry (*The Colossus*) has a higher proportion of the more 'formal' punctuation marks (colon, semi-colon) than expected on the basis of even distribution, while the last two volumes (*Ariel* and *Winter Trees*) have fewer of these punctuation marks than expected. On the other hand, the later volumes have more of the less formal punctuation marks (exclamation mark, question mark, dash, comma) than expected, whereas the earliest volume has fewer than expected. *Crossing the Water* appears to be transitional in terms of its punctuation. These results support the author's general hypothesis that the language becomes less formal as we go from early to later works.

4 $\chi^2 = 134.63.$

The critical value of χ^2 for the 5 per cent level and $(9-1)(4-1)$ or 24 df is 36.42. We can thus claim a significant difference in the distribution of sentence length for the four texts. By considering the observed and expected frequencies in the contingency table, it is possible to show that *The Colossus* has fewer short sentences, and more long sentences, than expected for an even distribution, while *Winter Trees* has many more short sentences, and fewer long ones, than expected. *Crossing the Water* and *Ariel* have a low proportion of very short sentences (1–5 words) and very long sentences (>35 words), and appear to be transitional between the other two volumes. Overall, the results tend to confirm the conclusions drawn from punctuation studies: there seems to be a decrease in formal complexity from earlier to later works, reflecting a structurally simpler and less formal style.

5 To avoid cells with expected frequencies of <5, we group the data from question 2 of the chapter 2 exercises as shown in tables A2.1 and A2.2.

Stressed syllables

$\bar{x} = 24.47$ units, $s = 3.55$ units, from question 3 of the exercises to chapter 3.

There are 4 pairs of E and O values, but the distributions have been made to agree on 3 values (those of the sample size, mean and standard deviation), so that df $= 4 - 3 = 1$. The critical value at the 5 per cent level is 3.84. Since the calculated value of χ^2 is smaller than this, we cannot reject the null hypothesis that the data are normally distributed.

Unstressed syllables

$\bar{x} = 22.64$ units, $s = 4.52$ units, from question 3 of the exercises to chapter 3.

Number of df $= 5 - 3 = 2$. The critical value at the 5 per cent level is 5.99. Since the calculated value is smaller than this, we cannot reject the null hypothesis that the data are normally distributed.

6 *must*:

	British	*American*	*Total*
root	153	150	303
epistemic	74	47	121
	227	197	424

Table A2.1

Intensity	Frequency (O)	Upper limit	Deviation from $\bar{x}$	$z = \frac{\text{deviation}}{s}$	Proportion below boundary	Proportion within class	E	$\frac{(O-E)^2}{E}$
⩽21	22	21.50	−2.97	−0.84	0.201	0.201	20.1	0.18
22-24	28	24.50	0.03	0.01	0.504	0.303	30.3	0.17
25-27	31	27.50	3.03	0.85	0.802	0.298	29.8	0.05
⩾28	19				1.000	0.198	19.8	0.03
						1.000	100.0	$\chi^2 = 0.43$

Table A2.2

Intensity	Frequency (O)	Upper limit	Deviation from $\bar{x}$	$z = \frac{\text{deviation}}{s}$	Proportion below boundary	Proportion within class	E	$\frac{(O-E)^2}{E}$
⩽18	16	18.50	−4.14	−0.92	0.179	0.179	17.9	0.20
19-21	20	21.50	−1.14	−0.25	0.401	0.222	22.2	0.22
22-24	27	24.50	1.86	0.41	0.659	0.258	25.8	0.06
25-27	23	27.50	4.86	1.08	0.860	0.201	20.1	0.42
⩾28	14				1.000	0.140	14.0	0.00
						1.000	100.0	$\chi^2 = 0.90$

Using Yates's correction,

$$\chi^2 = \frac{424\{|(150 \times 74) - (153 \times 47)|\} - 424/2)^2}{227 \times 197 \times 121 \times 303} = 3.53.$$

Without Yates's correction,

$$\chi^2 = \frac{424\{(150 \times 74) - (153 \times 47)\}^2}{227 \times 197 \times 121 \times 303} = 3.95.$$

The critical value of χ^2 for 1 df at the 5 per cent level is 3.84. The χ^2 value obtained without Yates's correction is thus just significant at the 5 per cent level, while that obtained using Yates's correction just fails to achieve significance.

have to:

	British	*American*	*Total*
root	226	209	435
epistemic	2	9	11
	228	218	446

Using Yates's correction,

$$\chi^2 = \frac{446\{|(226 \times 9) - (209 \times 2)|\} - 446/2)^2}{228 \times 218 \times 11 \times 435} = 3.64.$$

Without Yates's correction,

$$\chi^2 = \frac{446\{(226 \times 9) - (209 \times 2)\}^2}{228 \times 218 \times 11 \times 435} = 4.90.$$

Again, the result from the uncorrected χ^2 is significant at the 5 per cent level, but that from the corrected calculation is not quite significant.

7 (i) χ^2 for variation within the Early Middle English texts $= 5.18$;

χ^2 for variation within the Late Middle English texts $= 3.30$.

The critical value of χ^2 for 2 df at the 5 per cent level is 5.99, so that neither result is significant. That is, there is no significant difference in the distribution of contiguity of the predicator and following major element, across texts *within* either group.

Pooling the frequencies gives the following contingency table:

	Early ME	*Late ME*	*Total*
Non-contiguous	75	66	141
Contiguous	455	534	989
	530	600	1 130

χ^2 (using Yates's correction) = 2.28.

The critical value for 1 df at the 5 per cent level is 3.84, so the difference in distribution *between* the two groups of texts is not significant. That is, no diachronic change in contiguity relations from Early to Late ME can be demonstrated.

(ii) χ^2 for Early ME texts = 7.38;
χ^2 for Late ME texts = 2.68.

The critical value for 2 df at the 5 per cent level is 5.99, so that the variation within the Early ME texts is significant, whereas that in the Late ME texts is not. There is thus a diachronic change towards greater homogeneity on this measure.

(iii) χ^2 for Early ME texts = 0.13;
χ^2 for Late ME texts = 11.87.

The critical value for 2 df at the 5 per cent level is 5.99, so that the variation within the Early ME texts is non-significant, but that in the Late ME texts is significant. We thus have the opposite situation to that in (ii): with respect to inversion of subject and verb after an introductory adverbial, there is a diachronic change towards greater heterogeneity.

Chapter 10

1 The standard deviations of the samples were calculated in question 3 of chapter 3, the results being as follows:

Stressed: $s = 3.55$ units $\quad \therefore$ variance $= s^2 = 12.60$
df = 99.

Unstressed: $s = 4.52$ units $\quad \therefore$ variance $= s^2 = 20.43$
df = 99.

$$F = \frac{\text{larger variance}}{\text{smaller variance}} = \frac{20.43}{12.60} = 1.62.$$

The critical value of F for the 10 per cent level in a non-directional test with 100 df for the smaller variance and 50 df for the larger variance is given in table A8 as 1.48. With ∞ df for the larger estimate it is 1.28. The critical value with 99 df for each variance estimate thus clearly lies between 1.28 and 1.48. The calculated value of F exceeds this, and we can therefore reject the null hypothesis of homogeneity of variance at the 10 per cent level.

2 In question 1 of chapter 7, we calculated the values of s:

For text A: $s = 38.52$ word types $\quad \therefore$ variance $= s^2 = 1\,483.8$
df = 49.

For text B: $s = 37.21$ word types $\quad \therefore$ variance $= s^2 = 1\,384.6$
df = 49.

$$F = \frac{1\,483.8}{1\,384.6} = 1.07.$$

The critical value of F for the 10 per cent level with 50 df for each estimate is 1.60 in a non-directional test. We therefore cannot reject the null hypothesis that the variances are equal.

3 Set A: $\Sigma x_A = 123 \quad \Sigma x_A^2 = 1\,979 \quad N_A = 8$

Set B: $\Sigma x_B = 394 \quad \Sigma x_B^2 = 5\,612 \quad N_A = 29$

Set C: $\Sigma x_C = 968 \quad \Sigma x_C^2 = 19\,534 \quad N_C = 50$

Set D: $\Sigma x_D = 274 \quad \Sigma x_D^2 = 4\,510 \quad N_D = 17$

For whole data set: $\Sigma x = 1\,759 \quad \Sigma x^2 = 31\,635 \quad N = 104.$

$$\text{total sum of squares, } SS_t = \Sigma x^2 - \frac{(\Sigma x)^2}{N}$$

$$= 31635 - \frac{(1\,759)^2}{104} = 1\,884.22.$$

between-groups sum of squares,

$$SS_b = \frac{(\Sigma x_A)^2}{N_A} + \frac{(\Sigma x_B)^2}{N_B} + \frac{(\Sigma x_C)^2}{N_C} + \frac{(\Sigma x_D)^2}{N_D} - \frac{(\Sigma x)^2}{N}$$

$$= 1\,891.13 + 5\,352.97 + 18\,740.48$$
$$+ 4\,416.24 - 29\,750.78$$
$$= 650.04.$$

within-groups sum of squares, $SS_w = SS_t - SS_b$

$$= 1\,884.22 - 650.04$$
$$= 1\,234.18.$$

between-groups mean square, $s_b^2 = \dfrac{SS_b}{k-1}$ (k = no. of groups)

$$= \frac{650.04}{3} = 216.68.$$

within-groups mean square, $s_w^2 = \dfrac{SS_w}{N-k} = \dfrac{1\,234.18}{104-4} = 12.34.$

$$F = \frac{s_b^2}{s_w^2} = \frac{216.68}{12.34} = 17.56.$$

The critical value of F at the 5 per cent level, with 3 df for s_b^2 and 100 df for s_w^2 is 2.70, in a directional test. Since the calculated value of F exceeds the critical value, the means of the groups differ significantly at the 5 per cent level.

4 Let German = group 1, Japanese = group 2, French = group 3, Russian = group 4.

$\Sigma x_1 = 38 \quad \Sigma x_1^2 = 184 \quad N_1 = 10 \quad \bar{x}_1 = 3.80$

$\Sigma x_2 = 60 \quad \Sigma x_2^2 = 466 \quad N_2 = 9 \quad \bar{x}_2 = 6.67$

Correction for page 206

Omit the heading *Pairwise t-tests (non-directional)* and the table which follows.

$\Sigma x_3 = 42$ $\Sigma x_3^2 = 216$ $N_3 = 11$ $\bar{x}_3 = 3.82$

$\Sigma x_4 = 58$ $\Sigma x_4^2 = 464$ $N_4 = 8$ $\bar{x}_4 = 7.25$

$\Sigma x = 198$ $\Sigma x^2 = 1\,330$ $N = 38$

$SS_t = 298.32$

$SS_b = 93.58$

$SS_w = 204.74$

$s_b^2 = 31.19$

$s_w^2 = 6.02$

$F = 5.18$

The critical value of F at the 5 per cent level, with 3 df for s_b^2 and 35 df for s_w^2 (the nearest, in table A8, to the actual 34 df), is 2.87. There is thus a significant difference in the means of the groups at the 5 per cent level.

Pairwise t-tests (non-directional)

Groups compared	*t value*	*df*	*Critical value at 5% level*	*Significant*
German and Japanese	2.51	17	2.110	Yes
German and French	0.02	19	2.093	No
German and Russian	3.19	16	2.120	Yes
Japanese and French	2.44	18	2.101	Yes
Japanese and Russian	0.44	15	2.131	No
French and Russian	3.06	17	2.110	Yes

Chapter 11

1 If x is the score for one learner on the French test, and y the score on the German test,

$\Sigma x = 1\,089$ $\Sigma y = 1\,065$ $\Sigma x^2 = 62\,795$
$\Sigma y^2 = 58\,957$ $\Sigma xy = 59\,918.$

$$r = \frac{N\Sigma xy - \Sigma x \Sigma y}{\sqrt{\{N\Sigma x^2 - (\Sigma x)^2\}\{N\Sigma y^2 - (\Sigma y)^2\}}}$$

$$= \frac{(20 \times 59\,918) - (1\,089 \times 1\,065)}{\sqrt{\{(20 \times 62\,795) - 1\,089^2\}\{20 \times 58\,957) - 1\,065^2\}}}$$

$$= 0.69.$$

The critical value of r for $N = 20$ at the 5 per cent level in a non-directional test is 0.444. There is thus a significant positive correlation between the two sets of scores.

2 Since the data are best regarded as ordinal, the Spearman rank correlation coefficient is the appropriate measure.

Reading score	*Rank*	*Writing score*	*Rank*	*d*	d^2
3	4.5	6	13.5	9	81
4	8	4	8.5	0.5	0.25
5	11	3	4.5	−6.5	42.25
6	13.5	7	15	1.5	2.25
3	4.5	4	8.5	4	16
4	8	2	1.5	−6.5	42.25
3	4.5	5	11.5	7	49
5	11	6	13.5	2.5	6.25
2	1.5	3	4.5	3	9
4	8	5	11.5	3.5	12.25
7	15	3	4.5	−10.5	110.25
5	11	4	8.5	−2.5	6.25
6	13.5	4	8.5	−5	25
2	1.5	3	4.5	3	9
3	4.5	2	1.5	−3	9
					$\Sigma d^2 = 420$

$$\rho = 1 - \frac{6\Sigma d^2}{N(N^2-1)} = 1 - \frac{6 \times 420}{15(15^2-1)} = 0.25.$$

The critical value of ρ for $N = 15$ at the 5 per cent level in a non-directional test is 0.521. No significant correlation can therefore be claimed at the 5 per cent level.

3 *Order*

The scattergram indicates negative correlation. If $x = \%$ order, $y =$ median politeness rating, then

$$\Sigma x = 1\,115 \qquad \Sigma y = 114.82 \qquad \Sigma x^2 = 70\,983$$

$$\Sigma y^2 = 477.67 \qquad \Sigma xy = 2\,186.6$$

$$N = 35$$

$$r = -0.78.$$

The critical value of r at the 5 per cent level in a directional test is 0.306 for $N = 30$ and 0.264 for $N = 40$, as given in table A9. The calculated value of r is thus clearly significant at the 5 per cent level.

Request

The scattergram indicates positive correlation. If $x = \%$ request, $y =$ median politeness rating, then

$$\Sigma x = 1\,428 \qquad \Sigma y = 114.82 \qquad \Sigma x^2 = 98\,074$$

$$\Sigma y^2 = 477.67 \qquad \Sigma xy = 6\,435.07$$

$$N = 35$$

$$r = 0.87$$

The critical value of r lies between 0.264 and 0.306, as shown above. The correlation is thus significant at the 5 per cent level.

Suggestion

The scattergram indicates no clear pattern of correlation. If $x = \%$ suggestion, $y =$ median politeness rating, then

$$\Sigma x = 722 \qquad \Sigma y = 114.82 \qquad \Sigma x^2 = 36\,066$$

$$\Sigma y^2 = 477.67 \qquad \Sigma xy = 2\,231.60$$

$$N = 35$$

$$r = -0.09$$

The critical value is as before, and the calculated r value is clearly non-significant.

4 *Language use and reading*

$$\Sigma d^2 = 804$$

$$\rho = 0.40$$

The critical value of ρ for $N = 20$ at the 5 per cent level in a non-directional test is 0.447. The correlation between language use and reading ability is thus not significant at the 5 per cent level.

Language use and social class

$$\Sigma d^2 = 728.5$$

$$\rho = 0.45$$

The critical value is 0.447 as before. The correlation between language use and social class is thus just significant at the 5 per cent level.

Reading and social class

$$\Sigma d^2 = 571.5$$

$$\rho = 0.57$$

The critical value is again 0.447. There is thus a significant correlation between reading ability and social class at the 5 per cent level.

5

		Predominantly use postvocalic /r/		
		—	+	*Total*
Male	+	36	24	60
	—	11	29	40
		47	53	100

$$\phi = \frac{(24 \times 11) - (36 \times 29)}{47 \times 53 \times 40 \times 60} = -0.32$$

$$\chi^2 = N\phi^2 = 10.18.$$

The critical value of χ^2 at the 5 per cent level for 1 df is 3.84. There is thus a significant correlation between sex and the use of postvocalic /r/. In the sample investigated, females use this feature more than males.