




---

```

name: <unnamed>
log: \\wsl.localhost\Debian\home\lsys\neutrality\analysis\logs/summ.smcl
log type: smcl
opened on: 2 Dec 2025, 12:56:29

1 .
2 . qui do init.do

3 .
4 . qui reg ssl_quote_to_speech i.opposition $length_s $base_controls $min,
5 . keep if e(sample)
   (2 observations deleted)

6 .
7 . *=====
8 . * Unique MPs (mp): 204
9 . * - Opposition MPs:    16
10. * - Non-partisan MPs: 44
11. * - Ruling-party MPs: 144
12. count_unique mp
   (0 observations deleted)
204

13. count_unique mp if opposition
   (13796 observations deleted)
16

14. count_unique mp if non_partisan
   (14426 observations deleted)
44

15. count_unique mp if ruling
   (1580 observations deleted)
144

16.
17. *=====
18. * Total quotes: 14,901
19. * - Opposition MPs: 1,105 quotes
20. * - Non-partisan MPs: 475 quotes
21. * - Ruling-party MPs: 13,321 quotes
22. count
14901

23. count if opposition
1105

24. count if non_partisan
475

25. count if ruling
13321

26.
27. *=====
28. * Unique articles (article_id): 3,421
29. * - Articles with opposition quotes:    456
30. * - Articles with non-partisan quotes: 261

```

31. \* - Articles with ruling-party quotes: 3,354  
 32. count\_unique article\_id  
 (0 observations deleted)  
**3421**

33. count\_unique article\_id if opposition  
 (13796 observations deleted)  
**456**

34. count\_unique article\_id if non\_partisan  
 (14426 observations deleted)  
**261**

35. count\_unique article\_id if ruling  
 (1580 observations deleted)  
**3354**

36.  
 37. \*=====

38. \* Quote length (words, variable: quote)  
 39. \* - All quotes: mean 20.9, sd 38.5  
 40. \* - Opposition quotes: mean 20.1, sd 18.2  
 41. \* - Non-partisan quotes: mean 23.7, sd 19.3  
 42. \* - Ruling-party quotes: mean 20.9, sd 40.2  
 43. su quote

Variable	Obs	Mean	Std. Dev.	Min	Max
quote	<b>14901</b>	<b>20.94316</b>	<b>38.45591</b>	<b>1</b>	<b>4255</b>

44. su quote if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
quote	<b>1105</b>	<b>20.06516</b>	<b>18.22448</b>	<b>1</b>	<b>275</b>

45. su quote if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
quote	<b>475</b>	<b>23.68211</b>	<b>19.28725</b>	<b>1</b>	<b>116</b>

46. su quote if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
quote	<b>13321</b>	<b>20.91832</b>	<b>40.16441</b>	<b>1</b>	<b>4255</b>

47.  
 48. \*=====  
 49. \* Article length (words)  
 50. \* - All articles: mean 625.4, sd 263.4  
 51. \* - Opposition articles: mean 696.9, sd 339.7  
 52. \* - Non-partisan articles: mean 670.2, sd 272.0  
 53. \* - Ruling-party articles: mean 617.9, sd 254.7  
 54. su article

Variable	Obs	Mean	Std. Dev.	Min	Max
article	<b>14901</b>	<b>625.391</b>	<b>263.4284</b>	<b>115</b>	<b>3955</b>

55. su article if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
article	1105	696.9158	339.6973	115	3579

56. su article if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
article	475	670.1642	271.9951	192	2338

57. su article if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
article	13321	617.8614	254.7088	115	3955

58.

59.

60. =====

61. \* Article length (words)

62. \* - All speeches: mean 2,095.9, sd 1,808.8

63. \* - Opposition speeches: mean 1,156.0, sd 901.0

64. \* - Non-partisan: mean 1,808.6, sd 1,177.3

65. \* - Ruling-party: mean 2,184.2, sd 1,859.9

66. su speech

Variable	Obs	Mean	Std. Dev.	Min	Max
speech	14901	2095.959	1808.833	4	6100

67. su speech if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
speech	1105	1156.037	900.9681	10	5099

68. su speech if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
speech	475	1808.577	1177.276	14	5142

69. su speech if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
speech	13321	2184.175	1859.937	4	6100

70.

71. =====

72. \* Speech length (words)

73. \* - All speeches: mean 2,095.9, sd 1,808.8

74. \* - Opposition: mean 1,156.0, sd 901.0

75. \* - Non-partisan: mean 1,808.6, sd 1,177.3

76. \* - Ruling party: mean 2,184.2, sd 1,859.9

77. su speech

Variable	Obs	Mean	Std. Dev.	Min	Max
speech	14901	2095.959	1808.833	4	6100

78. su speech if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
speech	1105	1156.037	900.9681	10	5099

79. su speech if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
speech	475	1808.577	1177.276	14	5142

80. su speech if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
speech	13321	2184.175	1859.937	4	6100

81.

82. \*=====

83. \* QUotes per article

84. \*=====

85. \* - All articles: mean 4.36, sd 3.11

86. \* - Opposition articles: mean 6.34, sd 3.71

87. \* - Non-partisan articles:mean 4.84, sd 3.15

88. \* - Ruling-party articles:mean 4.22, sd 3.02

89. bysort article\_id: gen quotes\_per\_article = \_N

90. egen art\_tag = tag(article\_id)

91.

92. su quotes\_per\_article if art\_tag

Variable	Obs	Mean	Std. Dev.	Min	Max
quotes_per~e	3421	4.355744	3.109881	1	29

93. su quotes\_per\_article if art\_tag & opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
quotes_per~e	174	6.028736	3.63577	1	19

94. su quotes\_per\_article if art\_tag & non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
quotes_per~e	94	4.755319	2.738362	1	12

95. su quotes\_per\_article if art\_tag & ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
quotes_per~e	3153	4.251507	3.06215	1	29

96.

97. drop quotes\_per\_article art\_tag

98.

99. \*=====

```

100 * QUotes per MP
101 *=====
102 * - All MPs: mean 73.0, sd 130.8
103 * - Opposition MPs: mean 69.1, sd 92.1
104 * - Non-partisan MPs: mean 10.8, sd 9.1
105 * - Ruling-party MPs: mean 92.5, sd 147.6
106 bysort mp: gen quotes_per_mp = _N

107 egen mp_tag = tag(mp)

108
109 su quotes_per_mp if mp_tag

    Variable |   Obs      Mean   Std. Dev.     Min     Max
  quotes_per~p |  204  73.04412  130.7727      1     808

110 su quotes_per_mp if mp_tag & opposition

    Variable |   Obs      Mean   Std. Dev.     Min     Max
  quotes_per~p |   16  69.0625  92.10029      6     330

111 su quotes_per_mp if mp_tag & non_partisan

    Variable |   Obs      Mean   Std. Dev.     Min     Max
  quotes_per~p |   44  10.79545  9.090193      1      36

112 su quotes_per_mp if mp_tag & ruling

    Variable |   Obs      Mean   Std. Dev.     Min     Max
  quotes_per~p |  144  92.50694  147.6037      1     808

113
114 drop quotes_per_mp mp_tag

115
116 *=====
117 * QUotes per Speech
118 *=====
119 * - All speeches: mean 2.91, sd 3.86
120 * - Opposition speeches: mean 2.41, sd 2.40
121 * - Non-partisan: mean 1.97, sd 1.51
122 * - Ruling speeches: mean 3.01, sd 4.06
123 bysort speech_id: gen quotes_per_speech = _N

124 egen sp_tag = tag(speech_id)

125
126 su quotes_per_speech if sp_tag

    Variable |   Obs      Mean   Std. Dev.     Min     Max
  quotes_per~h | 5129  2.905245  3.864413      1      78

127 su quotes_per_speech if sp_tag & opposition

    Variable |   Obs      Mean   Std. Dev.     Min     Max
  quotes_per~h |  460  2.408696  2.379636      1      23

```

128 su quotes\_per\_speech if sp\_tag & non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
quotes_per~h	239	2.033473	1.746223	1	15

129 su quotes\_per\_speech if sp\_tag & ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
quotes_per~h	4430	3.003837	4.057526	1	78

130

131 drop quotes\_per\_speech sp\_tag

132

133 \*-----

134 \* QUotes per year

135 \*-----

136 \* - All quotes: mean 1241.8, sd 402.3

137 \* - Opposition: mean 92.1, sd 50.6

138 \* - Non-partisan: mean 39.6, sd 14.7

139 \* - Ruling party: mean 1110.1, sd 349.6

140 bysort year: gen quotes\_per\_year = \_N

141 egen yr\_tag = tag(year)

142

143 bysort year opposition: gen qpy\_oppo = \_N if opposition  
(13796 missing values generated)

144 bysort year non\_partisan: gen qpy\_nonp = \_N if non\_partisan  
(14426 missing values generated)

145 bysort year ruling: gen qpy\_ruling = \_N if ruling  
(1580 missing values generated)

146

147 egen yr\_tag\_oppo = tag(year) if opposition

148 egen yr\_tag\_nonp = tag(year) if non\_partisan

149 egen yr\_tag\_ruling = tag(year) if ruling

150

151 su quotes\_per\_year if yr\_tag

Variable	Obs	Mean	Std. Dev.	Min	Max
quotes_per~r	12	1241.75	402.2797	782	2104

152 su qpy\_oppo if yr\_tag\_oppo

Variable	Obs	Mean	Std. Dev.	Min	Max
qpy_oppo	12	92.08333	50.59727	26	160

153 su qpy\_nonp if yr\_tag\_nonp

Variable	Obs	Mean	Std. Dev.	Min	Max
qpy_nonp	12	39.58333	14.66882	18	64

154 su qpy\_ruling if yr\_tag\_ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
qpy_ruling	12	1110.083	349.5513	701	1883

155

156

157 \*=====

158 \* QUotes per parl

159 \*=====

160 \* - All MPs: mean 3,725.3, sd 2,205.8

161 \* - Opposition: mean 276.3, sd 237.6

162 \* - Non-partisan: mean 118.8, sd 67.9

163 \* - Ruling: mean 3,330.3, sd 1,944.4

164 bysort parl: gen quotes\_per\_parl = \_N

165 egen parl\_tag = tag(parl)

166

167 \* Opposition quotes per parliament

168 bysort parl: egen qpp\_oppo = total(opposition)

169 egen parl\_tag\_oppo = tag(parl) if !missing(qpp\_oppo)

170

171 \* Non-partisan quotes per parliament

172 bysort parl: egen qpp\_nonp = total(non\_partisan)

173 egen parl\_tag\_nonp = tag(parl) if !missing(qpp\_nonp)

174

175 \* Ruling quotes per parliament

176 bysort parl: egen qpp\_ruling = total(ruling)

177 egen parl\_tag\_ruling = tag(parl) if !missing(qpp\_ruling)

178

179 \* Summaries

180 su quotes\_per\_parl if parl\_tag

Variable	Obs	Mean	Std. Dev.	Min	Max
quotes_per~l	4	3725.25	2205.813	1641	6194

181 su qpp\_oppo if parl\_tag\_oppo

Variable	Obs	Mean	Std. Dev.	Min	Max
qpp_oppo	4	276.25	237.5547	112	628

182 su qpp\_nonp if parl\_tag\_nonp

Variable	Obs	Mean	Std. Dev.	Min	Max
qpp_nonp	4	118.75	67.85954	60	188

183 su qpp\_ruling if parl\_tag\_ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
qpp_ruling	4	3330.25	1944.355	1469	5378

184  
 185 \*=====

186 \* Acc.  
 187 \*=====

188 \* - All quotes: mean 91.5, sd 12.8  
 189 \* - Opposition: mean 90.9, sd 12.7  
 190 \* - Non-partisan: mean 91.7, sd 11.5  
 191 \* - Ruling: mean 91.6, sd 12.9  
 192 su ss1\_quote\_to\_speech

Variable	Obs	Mean	Std. Dev.	Min	Max
ss1_quote~ch	<b>14901</b>	<b>91.53473</b>	<b>12.82502</b>	<b>45</b>	<b>100</b>

193 su ss1\_quote\_to\_speech if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
ss1_quote~ch	<b>1105</b>	<b>90.86878</b>	<b>12.69393</b>	<b>46</b>	<b>100</b>

194 su ss1\_quote\_to\_speech if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
ss1_quote~ch	<b>475</b>	<b>91.67368</b>	<b>11.4568</b>	<b>48</b>	<b>100</b>

195 su ss1\_quote\_to\_speech if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
ss1_quote~ch	<b>13321</b>	<b>91.58502</b>	<b>12.8812</b>	<b>45</b>	<b>100</b>

196  
 197 \* - All quotes: mean 96.7, sd 10.2  
 198 \* - Opposition: mean 95.3, sd 13.2  
 199 \* - Non-partisan: mean 97.5, sd 7.3  
 200 \* - Ruling: mean 96.7, sd 10.0  
 201 su ss2\_quote\_to\_speech

Variable	Obs	Mean	Std. Dev.	Min	Max
ss2_quote~ch	<b>14901</b>	<b>96.66177</b>	<b>10.22995</b>	<b>0</b>	<b>100</b>

202 su ss2\_quote\_to\_speech if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
ss2_quote~ch	<b>1105</b>	<b>95.31855</b>	<b>13.18953</b>	<b>0</b>	<b>100</b>

203 su ss2\_quote\_to\_speech if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
ss2_quote~ch	<b>475</b>	<b>97.50105</b>	<b>7.289486</b>	<b>29</b>	<b>100</b>

204 su ss2\_quote\_to\_speech if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
ss2_quote~ch	<b>13321</b>	<b>96.74326</b>	<b>10.02869</b>	<b>0</b>	<b>100</b>

```

205
206 * - All quotes:      mean 89.9, sd 26.8
207 * - Opposition:     mean 89.3, sd 26.7
208 * - Non-partisan:   mean 93.8, sd 20.1
209 * - Ruling:         mean 89.8, sd 27.1
210 su ce_max_quote2speech

    Variable | Obs      Mean   Std. Dev.      Min      Max
ce_max_quo~h | 14901  89.93332  26.84338  .001209  99.99845

211 su ce_max_quote2speech if opposition

    Variable | Obs      Mean   Std. Dev.      Min      Max
ce_max_quo~h | 1105   89.32841  26.71708  .0013391  99.99763

212 su ce_max_quote2speech if non_partisan

    Variable | Obs      Mean   Std. Dev.      Min      Max
ce_max_quo~h | 475    93.83273  20.06376  .0012998  99.99446

213 su ce_max_quote2speech if ruling

    Variable | Obs      Mean   Std. Dev.      Min      Max
ce_max_quo~h | 13321  89.84445  27.05498  .001209  99.99845

214
215
216 *=====
217 * 'Perfect' quotes
218 *=====
219 * Perfect match indicators (1 = perfect, 0 = not perfect)
220 gen ssl1_perfect = ssl1_quote_to_speech == 100 if !missing(ssl1_quote_to_speech)

221 label var ssl1_perfect "Substring-1: perfect (=100)"

222
223 gen ss2_perfect = ss2_quote_to_speech == 100 if !missing(ss2_quote_to_speech)

224 label var ss2_perfect "Substring-2: perfect (=100)"

225
226 * For CE, this will probably all be 0 because max < 100, but for completeness:
227 gen ce_perfect = ce_max_quote2speech >= 99 if !missing(ce_max_quote2speech)

228 label var ce_perfect "CE max: perfect (=100)"

229
230 * Substring
231 * - All quotes:      42.2%
232 * - Opposition:     38.8%
233 * - Non-partisan:   35.8%
234 * - Ruling:         42.7%
235 su ssl1_perfect

    Variable | Obs      Mean   Std. Dev.      Min      Max
ssl1_perfect | 14901  .4217838  .4938609      0       1

```

236 su ss1\_perfect if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
ss1_perfect	1105	.3882353	.4875693	0	1

237 su ss1\_perfect if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
ss1_perfect	475	.3578947	.4798864	0	1

238 su ss1\_perfect if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
ss1_perfect	13321	.4268448	.4946379	0	1

239

240 \* Bow  
 241 \* - All quotes: 70.6%  
 242 \* - Opposition: 63.0%  
 243 \* - Non-partisan: 70.1%  
 244 \* - Ruling: 71.3%

245 su ss2\_perfect

Variable	Obs	Mean	Std. Dev.	Min	Max
ss2_perfect	14901	.7064627	.4553977	0	1

246 su ss2\_perfect if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
ss2_perfect	1105	.6298643	.4830595	0	1

247 su ss2\_perfect if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
ss2_perfect	475	.7010526	.4582794	0	1

248 su ss2\_perfect if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
ss2_perfect	13321	.7130095	.4523741	0	1

249

250 \* Semantic  
 251 \* - All quotes: 77.0%  
 252 \* - Opposition: 74.0%  
 253 \* - Non-partisan: 78.1%  
 254 \* - Ruling: 77.2%

255 su ce\_perfect

Variable	Obs	Mean	Std. Dev.	Min	Max
ce_perfect	14901	.7696799	.4210518	0	1

256 su ce\_perfect if opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
ce_perfect	1105	.7402715	.4386841	0	1

257 su ce\_perfect if non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
ce_perfect	475	.7810526	.4139688	0	1

258 su ce\_perfect if ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
ce_perfect	13321	.7717138	.4197438	0	1

259

260 \*=====

261 \* Age

262 \*=====

263

264 \* Per-MP age (in case age is stored at quote level)

265 bysort mp: egen age\_mp = mean(age)

266

267 \* Tag unique MPs

268 egen mp\_tag = tag(mp)

269

270 \* Overall and by party

271 su age\_mp if mp\_tag

Variable	Obs	Mean	Std. Dev.	Min	Max
age_mp	204	49.11675	8.311532	29.2372	76.9319

272 su age\_mp if mp\_tag &amp; opposition

Variable	Obs	Mean	Std. Dev.	Min	Max
age_mp	16	46.45687	10.61591	35.45626	70.89134

273 su age\_mp if mp\_tag &amp; non\_partisan

Variable	Obs	Mean	Std. Dev.	Min	Max
age_mp	44	47.99285	8.721769	29.2372	63.75

274 su age\_mp if mp\_tag &amp; ruling

Variable	Obs	Mean	Std. Dev.	Min	Max
age_mp	144	49.75571	7.860029	30.65339	76.9319

275

276 \*=====

277 \* Tenure

278 \*=====

```

279
280 * Per-MP tenure (if stored at quote level)
281 bysort mp: egen tenure_mp = mean(tenure)

282
283 su tenure_mp if mp_tag

    Variable |   Obs      Mean   Std. Dev.      Min      Max
  tenure_mp | 204  7.465426  8.005543 -.0383562 45.52368

284 su tenure_mp if mp_tag & opposition

    Variable |   Obs      Mean   Std. Dev.      Min      Max
  tenure_mp | 16  4.584745  6.610105 .218929 21.80367

285 su tenure_mp if mp_tag & non_partisan

    Variable |   Obs      Mean   Std. Dev.      Min      Max
  tenure_mp | 44  1.033389  .6683909 -.0383562 3.551541

286 su tenure_mp if mp_tag & ruling

    Variable |   Obs      Mean   Std. Dev.      Min      Max
  tenure_mp | 144  9.750846  8.206327 .2109589 45.52368

287
288 *=====
289 * Gender of MPs (share female)
290 *=====
291 bysort mp (gender): gen gender_mp = gender if _n == 1
  (14697 missing values generated)

292
293 * - All MPs: 47 / 204 (23.0% female)
294 * - Opposition MPs: 3 / 16 (18.8% female)
295 * - Non-partisan MPs: 18 / 44 (40.9% female)
296 * - Ruling-party MPs: 26 / 144 (18.1% female)
297 tab gender_mp if mp_tag

    gender_mp |   Freq.      Percent      Cum.
  0 | 157  76.96  76.96
  1 | 47  23.04  100.00
  Total | 204  100.00

298 tab gender_mp if mp_tag & opposition

    gender_mp |   Freq.      Percent      Cum.
  0 | 13  81.25  81.25
  1 | 3  18.75  100.00
  Total | 16  100.00

299 tab gender_mp if mp_tag & non_partisan

    gender_mp |   Freq.      Percent      Cum.
  0 | 26  59.09  59.09
  1 | 18  40.91  100.00
  Total | 44  100.00

```

300 tab gender\_mp if mp\_tag &amp; ruling

gender_mp	Freq.	Percent	Cum.
0	118	81.94	81.94
1	26	18.06	100.00
Total	144	100.00	

301

302 \*=====

303 \* Gender of quotes

304 \*=====

305 \* - All quotes: 83.5% male, 16.5% female

306 \* - Opposition: 72.0% male, 28.0% female

307 \* - Non-partisan: 59.6% male, 40.4% female

308 \* - Ruling-party: 85.3% male, 14.7% female

309 tab gender

Gender of politician	Freq.	Percent	Cum.
male	12,441	83.49	83.49
female	2,460	16.51	100.00
Total	14,901	100.00	

310 tab gender if opposition

Gender of politician	Freq.	Percent	Cum.
male	795	71.95	71.95
female	310	28.05	100.00
Total	1,105	100.00	

311 tab gender if non\_partisan

Gender of politician	Freq.	Percent	Cum.
male	283	59.58	59.58
female	192	40.42	100.00
Total	475	100.00	

312 tab gender if ruling

Gender of politician	Freq.	Percent	Cum.
male	11,363	85.30	85.30
female	1,958	14.70	100.00
Total	13,321	100.00	

313

314

315 su ss1\_quote\_to\_speech ss2\_quote\_to\_speech ce\_max\_quote2speech

Variable	Obs	Mean	Std. Dev.	Min	Max
ss1_quote~ch	14901	91.53473	12.82502	45	100
ss2_quote~ch	14901	96.66177	10.22995	0	100
ce_max_quo~h	14901	89.93332	26.84338	.001209	99.99845

```
316
317
318 log close
    name: <unnamed>
    log: \\wsl.localhost\Debian\home\lsys\neutrality\analysis\logs\summ.smcl
  log type: smcl
closed on: 2 Dec 2025, 12:56:37
```

---