Intercoder Reliability Report

Election News 2011

30 newspapers were coded for this project. 6 editions for each of the following:

1. The Korean Times Daily publishes six days per week in Korean.

2. The Russian Express is a weekly that publishes every Friday in Russian.

3. The Canadian Punjabi Post publishes six days per week in Punjabi.

4. Ming Pao publishes seven days a week in Chinese.

5. The Punjabi Daily publishes six days per week in Punjabi.

There were four coders: One coder fluent in Korean, a second coder fluent in Punjabi, a third who was fluent in Russian and a fourth who was fluent in Chinese. The coders coded 6 issues of the newspapers collected from March 25, 2011, to May 4, 2011.

To ensure intercoder reliability amongst all coders in four different languages, all coders were trained to code election stories from the Toronto Star, National Post and Globe and Mail. Once they reached acceptable levels of intercoder agreement, we felt confident that their coding decisions for the respective ethnocultural newspapers would be consistent and accurate.

Coders received 20 hours of coder training for the Star/Globe/Post content. After random testing indicated they were coding in a similar/consistent manner with each other, they then coded 40 randomly selected federal election stories from the three English-language dailies. The 40 stories represented about 10 percent of 431, the total number of election stories coded in all the ethnic papers. Upon completing the coding of the 40 stories, the coders met to discuss discrepancies revealed by the intercoder test. Disagreements were resolved through discussion.

We measured intercoder reliability for nominal variables using Cohen's kappa, with the aid of PRAM (Program for Reliability Assessment with Multiple Coders) software. Cohen's kappa is a relatively conservative index that measures the extent to which coders make identical coding decisions, and takes into account the agreement expected by chance. Research suggests that for multiple coders, values of between .400 and .75 may be taken to represent fair to good agreement beyond chance (Banerjee, Capozzoli, McSweeney & Sinha, 1999; Fleiss, 1981; Landis & Koch, 1977).

Intercoder reliability was at or above 0.4 for all variables, with the exception of Stand-key phrases. We kept Stand-key phrases because coders did pass the test requiring them to identify stories that contained a specific slant. Having done that with a satisfactory degree of agreement, they were free to choose from many key phrases in the stories to illustrate the slant. The fact that in some cases they chose different phrases or copied longer or shorter passages from stories would not affect the analysis.

Variable	Cohen's Kappa Test Result
Publication ID	0.665
Issue date	0.913
UID	1
Page #	0.495
Headline	0.466
Item Form	0.975
Exclusive Interview	1.000
Horserace	0.869
Ad Size	1.000
Story Origin	0.684
Photo Status	1.000
CandidateEthnicity	0.576
Stand	0.81
Stand Key Phrase	0.25
Issue1	0.829
Party1	0.921
Party2	0.808
Party3	0.78
CandidateMention1	0.936
CandidateMention2	0.819
CandidateMention3	0.841
Party/Candidate1	0.919
Party/Candidate2	0.795
Party/Candidate3	0.768
Party/Candidate4	0.802

REFERENCES

Banerjee, Wayne, Laura McSweeney and Sinha Debajyoti. 1999. Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics, 27*(1), 3-23.

Fleiss, J.L. 1981. *Statistical methods for rates and proportions* (2nd edition). New York: John Wiley.

Landis, J.R. and Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.