Dictionary methods

Because hand-coding methods are highly labour intensive, a more current approach replaces the hand-coding of texts with computerized coding techniques (Laver and Garry 2000; Kleinnijenhuis and Pennings 2001). Most of these techniques reproduce the hand-coding of texts, using computer programs instead of people to do the coding based on dictionaries linking specific words to predefined policy positions.

Three distinct dictionary techniques are discussed below. The so-called "Positional Word Count" technique builds a dictionary based on the most frequently used words that are illustrative of the policy position of a given political party. A content analysis of party platforms in the 1992 British election by Laver and Garry (2000) is a good illustration. The analysis matches the words in predefined Conservative, neutral, and Labour dictionaries with those found in the Labour and the Conservative platforms, and gives each word a value of -1 if it is also found in the Conservative dictionary, 0 if it is found in the neutral dictionary, and +1 if it is found in the Labour dictionary. The same dictionary is then used to analyze the Labour and Conservative platforms in the 1997 British election.

Kleinnijenhuis and Pennings (2001) developed a method that makes computer-coding compatible with the CMP handcoding technique. Instead of giving each word in a party platform a value of -1, 0 or +1 depending on whether it is found in a Conservative, neutral, or Labour dictionary, like Laver and Garry did in their study of British party platforms, Kleinnijenhuis and Pennings associate a list of words with each of 54 issue categories found in the initial CMP scheme. Their method consists of counting how frequently each word in the dictionary is found in the party platforms they analyze.

The natural sentence approach developed by Ray (2001) also uses CMP-compatible coding. However, unlike Kleinnijenhuis and Pennings' technique which uses the word as its unit of analysis, Ray's technique uses the sentence (or quasi-sentence) as its unit of analysis. Ray's dictionary also records the policy direction ("in favour off, opposed to, support") of a party platform statement, in addition to the issue. Computerized coding schemes such as the ones just described are able to produce much more accurate and valid

estimates of policy positions of the same texts when compared with the manual coding method used by the CMP.

However, there are several limits associated with the type of dictionary used by these methods:

1. One needs to construct a new dictionary for every new scale on which to position political texts. As a result, no matter how heavily computerized it is, the method still requires a lot of human effort in developing and testing coding dictionaries that fit particular times and political cultures.
2. As Laver, Benoit and Garry point out (2003: 312) “heavy human involvement in the generation of coding dictionaries imports some of the methodological disadvantages of traditional techniques based on potentially biased human codes.”
3. Another limit is the small amount of data on which the typical dictionary is constructed, so that adding or subtracting one word from the dictionary may have a very large impact on the final result.
4. The method does not permit to precisely calculate the error associated with the measures generated by the dictionary. The construct validity of the method is therefore undermined.

References

Kleinnijenhuis, J. and P. Pennings (1999). A Probabilistic Keyword Approach to Textual Analysis. Mannheim Joint Sessions of the ECPR. Mannheim.

Kleinnijenhuis, J. and P. Pennings (2001). Measurements of Party Positions on the Basis of Party Programmes, Media Coverage, and Voter Perceptions. Estimating the Party Position of Political Actors. M. Laver. New York, Routledge: 162-182.

Laver, M., K. Benoit, and J. Garry (2003). "Extracting Policy Positions from Political Texts Using Words as Data." The American Political Science Review **97**(2): 311-331.

Laver, M. and J. Garry (2000). "Estimating Policy Positions from Political Texts." American Journal of Political Science **44**(3): 619-634.

Ray, L. (2001). A Natural Sentence Approach to the Computer Coding of Party Manifestos. Estimating the Positions of Political Actors. M. Laver. New York, Routledge: 149-161.