

# 남북한 방송언어의 차이에 대한 기초 분석

이창환\*, 김정일\*\*, 박종민\*\*\*

\*서강대학교 심리학과, \*\*아주대학교 심리학과 \*\*\*경희대학교 언론정보학과  
김정일(kyungil@ajoug.ac.kr)

## Preliminary analysis about the differences between South and North Korean Broadcasting Languages

Chang H. Lee\*, Kyungil Kim\*\*, Jongmin Park\*\*\*

\*Department of Psychology, Sogang University

\*\*Department of Psychology, Ajou University

\*\*\*The School of Journalism and Communications

### 요 약

본 연구는 장기간의 남북한 분단으로 인한 언어적 이질성의 정도를 가늠해보기 위하여 남북한 방송언어 비교하였다. 연구의 주 목적은 남북한 언어간 차이에 대한 데이터가 부족한 상황에서 언어사용 실태에 대한 토대 데이터를 제공하는 것이었다. 남북한의 주요 방송사 뉴스 동영상에서 추출한 텍스트를 대상으로 한국어분석프로그램 KLIWC (Korean Linguistic Inquiry and Word Count)으로 분석하였다. 분석 결과, 북한 방송언어는 KLIWC의 각 차원에서 남한 언어와 유의미한 차이를 나타냈으며, 특히 정서적 단어, 인지적 단어, 사회적 단어 등에서 유의미한 차이가 발견되었다. 또한 북한 방송에서는 인칭대명사나 품사와 같은 언어학적 기능어에 있어서도 남한방송보다 사용빈도가 높았다. 이러한 차이에 대한 예비적인 심리학적 신문방송학적 해석을 제공하였다.

### 1. 서론

남, 북한이 분단된지도 60년이 넘는 시점에서 그동안 남, 북한 국민들의 사고방식이나 심리에서는 많은 차이가 나고 있다. 특히 심리적 기제와 현상을 추론할 수 있는 언어적 차이에 대한 정량적인 연구는 많이 이루어지지 않았다. 즉, 남북한 국민이 어떠한 언어적 유형을 많이 쓰고, 어떠한 내용에서 상대적인 차이가 나는지에 대한 연구가 없는 실정이다.

본 연구에서는 남북한의 언어적 차이를 알기 위한 정량적 토대 자료를 구축하기 위하여 남북한 뉴스방송언어차이를 분석하였다.

### 2. 선행이론 연구

기존의 남북 언어 비교는 남북한의 언어가 역사적으로 어떤 방식으로 변화해 왔는가, 그리고 그 결과 어휘 수준에서 단어들이 어떤 차이를 보이며 언어학적인 차이는 무엇인지를 주로 연구해왔다. 또한 띄어쓰기, 맞춤법 등과 같은 문법수준에서의 차이와 동음이의어 분석과 같은 언어의 형

태적 차이에 관한 연구에 치중해왔다. 하지만, 남북의 사람들 간의 본질적인 언어적 차이를 알아보기 위해서는 언어의 내용이 인지적, 정서적, 사회적으로 어떻게 차이가 나고 언어의 논리성과 복잡성에서 얼마나 차이가 나는지에 초점을 두어야 한다. 보다 구체적으로 사람들의 심리사회적인 차이로 인하여 유사한 상황에서 내용상 그리고 언어 구조상 어떤 차이가 나타나는지를 밝혀야 하며 이러한 차이를 사회과학적 방법론에 입각하여 수치로 정량화 시킬 필요가 있다. 사회과학에서는 종속측정치를 정량화 시키는 것이 다른 연구자에 의한 반복적인 검증을 보장하고 보편타당한 객관적인 정보를 제공한다는 측면에서 필수적인 연구 방법인 것이다.

기존의 주요 연구를 살펴보면, 육훈(1996)은 몇 가지의 관점으로 남북 언어의 차이를 설명했다. 우선 언어 정책상의 차이이다. 북한은 언어를 ‘혁명과 건설에 힘 있는 무기’라고 주장하고, 남한과 달리 한자 사용을 금지하고, 외래어를 모두 한글로 바꾸어 사용하고 있다. 둘째, 자모배열 순서와 명칭상의 차이를 이야기한다. 이는 기본적으로 남북한 언어가 자모의 배열과 명칭들이 다르다는 것인데, 남한의 자모 수는 24개인 반면, 북한의 자모 수는 40개이다. 이러한

자모수의 차이는 사전과 컴퓨터 사용에 있어 큰 차이를 유발하게 된다. 세 번째는 맞춤법의 차이를 이야기 한다. 이러한 맞춤법의 차이는 맞춤법의 개념을 살펴보면 명확히 알 수 있다. 남한은 ‘한글 맞춤법은 표준어를 소리대로 적되, 어법에 맞도록 함을 원칙으로 한다(육훈, 1996, 재인용)’는 것이 기본 방침이 반면 북한의 경우는 ‘조선말 맞춤법은 단어에서 뜻을 가지는 매개 부분을 언제나 같이 적는 원칙을 기본으로 하면서 일부 경우 소리나는대로 적거나 습관을 따르는 것을 허용(육훈, 1996 재인용)’하는 것을 기본 원칙으로 하고 있다. 다음 네 번째로 육훈(1996)이 제시하는 남북한 언어의 차이는 이질적인 문법들인데, 예를 들면, 북한의 ‘토’와 남한의 ‘조사’가 가지는 차이, 사동형과 피동형에서의 남북한 언어 차이, 띄어쓰기의 차이점, 그리고 무엇보다도 중요한 것은 어휘 어미의 본질적 의미 차이<sup>1)</sup> 등을 이야기 할 수 있다. 마지막으로 호칭(예, 안해와 아내)의 차이와 오해 받기 쉬운 “괜찮습니다”와 “일없습니다”의 표현을 들어 실생활에 있어 남북한 언어 사용의 차이를 육훈은 설명하고 있다.

박시균, 권병로(2003)는 남북한 언어를 발음과 어휘를 중심으로 비교 분석하였다. 우선 발음의 경우 발음원칙, 모음, 어두의 비음 ‘ㄴ’과 유음 ‘ㄹ’, 겹받침, 음의 동화, 음의 첨가, 호격 조사 ‘아’, 한글 자모 받침의 연음 시 발음 등을 기준으로 분석하였다. 간단히 살펴보면, 먼저 발음에 관하여 남한은 ‘표준어의 실제 발음을 따르되, 국어의 전통성과 합리성을 고려하여 정함(1988년 표준어 규정)’으로 되어 있는 반면, 북한은 1988년 <조선말 규범집>에 따라 ‘조선말 발음법은 혁명의 수도 평양을 중심지로 하고 평양말을 토대로 하여 이룩된 문화어의 발음에 기준한다’고 되어 있다. 한편, 모음에서 발생하는 남북한 발음차이의 경우 ‘ㄷ’의 경우 북한어는 ‘ㄱ, ㄹ, ㅎ’ 뒤에서만 <에>로 발음되며, 남한은 ‘예, 레’ 말고는 ‘예’가 <에>로 발음되는 차이가 있다. 또한 남한에는 어두에 ‘ㄷ’이나 ‘ㄴ’ 선행모음을 동반하는 ‘ㄴ’의 소리가 오면 안 되는 두음법칙이 존재하지만 북한은 두음법칙을 인정하지 않고, 음의 동화에 있어서도 북한은 ‘ㄷ’이 근접한 비음(ㄴ, ㄹ, ㅇ)의 영향으로 비자음화를 허용하지 않는 특징(예, 식량의 발음 - 남한<식량>, 북한<싱량>)이 있다(박시균, 권병로, 2003). 이외에도 발음에서 남북한 언어의 차이를 박시균, 권병로(2003)은 호격조사 ‘아’를 설명하면서, 일반적으로 남북한 공히 모음으로 시작되는 어미나 조사, 접미사의 앞에 놓인 받침은 원래대로 발음하도록 하였으나, 예외적으로 북한말은 부름의 조사 ‘아’앞에 놓인 ‘ㅅ’은

<ㄷ>으로 발음하도록(예, ‘우리의 영원한 벗아 <남한 - 버사> <북한 - 버다>)되어 있다고 설명한다.

김상준(2002)은 6.15 남북정상회담의 남북한 TV보도를 중심으로 남북한간 방송보도의 비언어적 메시지 특성에 관한 비교를 하였다. 그에 따르면 북한 언론의 목적은 “인민을 교양, 개조하고 조직하는 유력한 선전, 선동의 수단이며, 사상적 무기이자 공식적인 국가기관의 하나”(p.176)이므로 방송 또한 체제 수호, 자주성의 강조, 그리고 당의 정책을 비판 없이 반영하는데 그 목적이 있다고 한다. 그러므로 북한의 방송언어는 남한에 비해 보다 구어체적이며 단순하고 감정적이고 장황한 수사적 표현을 즐겨 사용한다고 전제하고 있다. 남한 방송의 비언어적 특성으로는 시청자에게 호감을 주는 부드러운 뉴스인 반면 북한 방송은 목적성을 띠며 사실적 내용 전달보다는 인민의 사상과 감정을 표현할 것을 요구하며 이를 위해 방송원의 중립적 자세보다는 특정 사안에 대한 의사표시(예, 지지, 반대, 기쁨, 슬픔 등)가 분명하게 반영한다. 또한 남한의 뉴스 전달 속도는 생동감을 더욱 부각시키기 위해 북한(1분당 평균 330음절) 보다 분당 약 20음절 빠른 평균 350음절을 발음하는 것으로 나타났다. 이외에도 성별에 따른 뉴스분류, 의상과 표정도 분석하였다.

최근에는 사람들의 언어 사용방식과 유형을 보다 정량적이고 체계적으로 밝히고자 하는 방향으로 연구가 진행되는 것이 세계적 추세인데, 현재 가장 의미 있는 연구물들을 제공하는 프로그램 중 하나가 바로 최근에 개발된 LIWC(Linguistic Inquiry and Word Count)이다. 이 언어분석 프로그램은 최초로 James Pennebaker 에 의하여 1990년대 후반 이후 개발된 것으로서, 영어권 심리학계는 물론 관련 사회과학의 다양한 분야에서 이 프로그램을 사용하여 언어에서부터 추출 가능한 다양한 변인들과 이에 관련된 심리적 현상들을 밝혀내는 연구들이 활발하게 진행되어 오고 있다(Pennebaker, Mehl, & Niederhoffer, 2003). 그 결과 언어에 반영되는 신체/정신적 건강 증진 여부(Pennebaker, 1997; Pennebaker, Mayne, & Francis, 1997), 우울증 등 정신병리 현상과 관련된 언어 변인들의 도출(Rude, Gortner, & Pennebaker, 2004), 성격 및 성취 욕구(Pennebaker & King, 1999), 사회적 위기 상황 시 대중의 평균적인 언어적 변화(Lee & Park, 2004; Mehl & Pennebaker, 2003), 자살에 대한 노출 가능성(Stirman & Pennebaker, 2001), 면역지수 CD-4 (t-lymphocyte) 세포의 양 (Booth, Petrie, & Pennebaker, 1997), Hepatitis B 항체 수준(Petrie, Booth, Pennebaker, 1999), 언어의 양식(글쓰기 대 말하기(Lee, Nam, & Pennebaker, 2004), 허위(Deception)진술 여부(Newman, Pennebaker, Berry, Richards, 2002), 나이에 따른 변화 (Pennebaker & Stone, 2003), 다문화권 비교

1) 어휘 어미의 차이란 단어 자체가 가지고 있는 본질적 개념의 차이로 육훈(1996)은 훈민정음, 교수, 자본주의, 지도자, 학교 등의 개념 차이를 예를 들어 설명하고 있다.

(Ramirez-Esparza, Gosling, Benet-Martinez, Potter, & Pennebaker, in press), 실직자 언어분석 (Spera, Buhrfeind, & Pennebaker, 1994) 등에 이르기까지 실로 광범위한 영역에서 다양한 심리적 특성과 그 언어적 반영을 확인하는 연구들이 진행되어 왔다.

이러한 영어권의 획기적인 연구 성과에 근거하여 LIWC의 한국어 분석프로그램인 K(Korean)-LIWC가 개발되었다 (프로그램 웹사이트: <http://k-liwc.ajou.ac.kr>). K-LIWC도 주요 심리학분야외에도 언어를 정보전달 수단으로 사용하는 신문방송학 분야(박종민, 박재진, 이창환, 2007; 박재진, 박종민, 이창환, 2008-> 참고문헌추가됨)와 정치언어(최윤선, 이창환, 2008)를 분석하는 분야에서 사용되었다. 본 연구에서는 KLIWC와 같이 널리 쓰이고 신뢰도와 타당도가 검증된 (이창환, 심정미, 윤애선, 2005) 언어분석 프로그램을 사용하여 남북한 언어를 비교하고자한다. 이에 대한 첫출발로 일상생활의 표준어를 반영하는 방송뉴스언어를 분석하고자한다. 방송언어가 일반 국민들의 언어사용과 흡사하며 언어사용에 지대한 영향을 미친다는 것에 대해서는 많은 부분 이론의 여지가 없다. 특히 오늘날과 같이 TV 등 영상매체의 성장이 급진전하는 상황에서 그 영향력은 더욱 커질 것이다. 남북한 간 정치 분야를 넘어서 경제, 사회, 문화 등 다양한 분야에서 교류가 확산, 증진되고 있는 현 시점에서 남북한간 방송언어 사용에 대한 차이를 체계적으로 분석해보는 것은 학술적 가치를 넘어서 다양한 분야에서 남북한간의 이해의 격차를 좁혀 상호간 보다 발전적인 관계형성을 이루는데 실질적 도움을 줄 수 있는 중요한 작업이 될 것으로 여겨진다.

KLIWC를 통하여 분석되게 될 언어변인은 일반 내용단어들을 포함하여 정서, 인지, 사회 등을 반영하는 심리적 단어들과 기능적 변인의 비율이다. 기능적 변인이라함은 전체 단어수를 비롯하여 문장당 단어수, 조사, 형태소, 의문문의 비율 등 실제 의미와는 관계없지만 전체 텍스트의 구조와 복잡성을 결정짓는 변인들이다. 따라서 남북한 방송언어간 내용적 차이 뿐 만 아니라 글의 위계적 체계적 차이에 관한 토대 데이터를 제공할 것이다.

### 3. 연구 방법 및 자료수집

샘플 수집 절차. 우선 북한의 뉴스 언어를 분석대상은 2009년 8월에서 2010년 1월에 걸쳐 방송된 KBS 남북의 창에 북한 방송 동영상이었다. 각 방송 언어는 프로그램의 스크립트를 복사하여 엑셀 스프레드시트에 붙여넣기를 하여 수집되었으며, 방송 내용을 보면서 스크립트가 방송 내용과 다를 경우에 수정을 하였다. 추후 K-LIWC 프로그램으로

분석하기 위해 날짜별로 Text 문서 파일 형태로 변경하였다. 수집된 스크립트 내용은 수집된 날짜, 세부 코너, 발화자의 성별, 발화의 주제와 함께 기록되었다.

이결과 총 166회분의 방송 분량의 스크립트가 수집되었다. '남북의 창'에서 주로 다뤄지는 북한 방송의 내용은 프로그램 내의 각 섹션에 따라 달랐다. '기획리포트' 코너에서는 주로 북한 중앙TV의 정치, 외교, 사회 관련 뉴스들이 주를 이루었다. '북한 상식 백과'와 '클로즈업 북한'코너는 주로 북한 사회의 현실을 보여주는 내용으로 문화, 서민 생활 관련뉴스, 정보 프로그램, 드라마나 영화 등등 내용이 다양했다.

분석절차. 수집된 방송 스크립트 샘플 중 뉴스만을 대상으로 하였다. 동영상의 내용이 텍스트로 전사된 것들중에서 남북한 각각 21개를 무선적으로 뽑았다.

최종적으로 각각의 전사된 뉴스 동영상 내용을 KLIWC로 분석하였다. KLIWC 분석방식은 앞서 상술했 바와 같이 특정 텍스트내의 단어들을 유형별로 사용비율을 자동적으로 계산해내는 방식이다. 따라서 각각의 전사된 뉴스 텍스트에 대하여 KLIWC는 명사, 대명사 등 9개 기본 품사별 사용빈도, 감정적 단어와 인지적 단어 등 심리적 단어의 사용빈도, 28개의 각종 내용단어의 사용빈도를 산출하였다. 하지만, 텍스트가 KLIWC에 투입되면 즉시적으로 계산되는 방식은 아니고 단어하나하나에 대하여 중의성을 분석자에게 되묻는 단계가 있었다. 일례로 "가는"이라는 단어는 가늘다, 가다, 가+는 등 다양한 뜻이 있기에 KLIWC는 여러 가지 가능한 옵션을 산출하며 분석자가 이를 결정하게 된다. 한국어 단어는 대부분의 단어가 중의적이기에 A4용지 반장만 되더라도 텍스트에 대한 중의성 판단은 30-40분 가량의 작업이 필요하였다.

### 4. 연구 결과 및 논의

표1은 남북한 방송에 있어서 언어학적 변인들의 차이이다. 먼저 주목할 만한 부분은 어절/문장, 즉 문장 당 어절의 비율이다. 남한 언어 사용자들이 북한 언어 사용자들에 비해 유의미하게 높은 비율을 보였다. 이 비율은 언어의 복잡성을 나타내는 지수로 기존 연구에 의하면 일반적으로 언어가 단순할수록 걱정이나 근심이 적은 경향을 보인다 (Pennebaker, Mehl, & Niederhoffer, 2003).

그리고 북한 언어 사용자들은 거의 모든 종류의 인칭 대명사를 더 빈번하게 사용하는 것으로 관찰되었다. 개별 인

칭 대명사와 관련된 심리적 상태의 추정은 다양한 연구에서 찾아볼 수 있다. 예를 들어, 자기중심적이고 우울한 사

람일수록 1인칭 단수 대명사의 사용을 많이 하며, 다른 기능어의 종류와 수도 단순해지는 반면, 정상인의 경우에는 글이 복잡하고 3인칭 대명사를 자주 쓰는 경향이 있는 것으로 관찰되고 있다(Pennebaker, et al., 2003; Pennebaker & Stone, 2003). 그런데 북한 언어 사용자들은 모든 인칭 대명사의 사용에 있어서 더 높은 빈도를 보인 것으로 보아 심리적 차원의 문제로 보기 보다는 언어 사용에 있어서 현대 우리 국어 사용 경향성에서 쉽게 관찰할 수 있는 '인칭 대명사의 생략'이 덜 일어나고 있는 것으로 보인다. 이는 언어학적 관점에서 볼 때도 주목할 만한 차이이다.

그 외에 북한 방송 언어에서는 지시대명사, 동사, 형용사, 관형사, 부사, 감탄사, 어미 등의 높은 사용빈도가 관찰되었고, 남한 방송 언어에서는 일반명사, 고유명사, 양수사, 조사, 수관형사, 접미사, 영어/외래어 표현 등이 더 빈번하게 사용되는 것으로 나타났다.

다음으로 표 2는 남북한 방송에 있어서 심리학적 단어들과 내용단어들의 차이를 나타내고 있다. 가장 주목할 만한 차이점은 북한 방송에서 유의미하게 많은 감정적 정서적 단어가 있었다는 것이다. 이는 평소 일반인들이 느꼈던 북한 뉴스의 많은 정서적 반응과 감정적 호소가 데이터로도 입증됨을 의미한다. 매우 흥미로운 사실은 북한방송에서는 긍정적인 정서라든가 낙천성과 같이 정적인 감정 단어가 남한 방송보다 많이 사용되며 부정적인 정서단어는 적게 사용된다는 것이다. 이는 북한 뉴스가 항상 호전적이고 부정적이지 않으며 많은 내용에 있어서 긍정적인 주제를 담고 있을 수 있으며 이에 대한 사례연구가 필요하겠다. 아울러 슬픔이나 불안과 같은 부적인 단어에 있어서는 북한방송의 사용량이 남한보다 많았다. 이는 북한 뉴스 방송의 감정적 정서적 변화의 스펙트럼이 다양함을 의미한다.

다음으로 친구, 가족, 학교, 직장, 타인참조, 체면등 사회적 과정과 관련된 단어의 사용빈도가 북한방송에서 많았으나 성취, 운동, 재정적 이슈와 관련된 단어의 사용빈도는 남한 방송에서 많았다. 이는 도시화와 산업화가 진행되지 않을수록 가족과 사회관련 가치를 더욱 중시하는 집단주의적 문화가 발달함을 보여주는 데이터이다.

인지적 과정 범주와 관련하여서 흥미로운 사실은 원인, 기대, 확신관련 변인에 걸쳐서 북한방송의 사용빈도가 남한보다 많았다. 이러한 연유에 대한 연구는 별도로 필요하겠지만, 북한방송에서는 뉴스에 대한 객관적인 사실 보도 외에도 뉴스와 관련된 원인과 결과에 대한 확신, 향후기대를 언급하는 경우가 많을 수 있음을 추론할 수 있다.

그 외에 수면, 성 등 신체와 관련된 단어의 사용빈도가 북한방송에서 많았으며, 최근의 식량문제를 반영하듯이 식사와 관련된 단어의 사용빈도가 북한방송에서 월등히 많았

다. 그리고 형이상학적 이슈와 관련된 단어의 사용빈도는 남한 방송에서 많았으며 의외로 속어나 맹세관련 단어의 사용빈도가 남한방송에서 많았다. 하지만 속어나 맹세관련 단어의 경우는 전체텍스트내에서 그 사용 비율이 0.3%인 관계로 집단간 비교에 큰 의미를 둘 수는 없다.

## 참고문헌

- [1] 김상준, "남북한 방송보도의 비언어 커뮤니케이션에 관한 비교 연구: 6.15 남북정상회담의 남북한 TV 보도를 중심으로", 화법연구, 4권, 147-181, 2002.
- [2] 박시균, "남북한 언어에 대한 비교 연구 2: 화술과 화법, 방송화법을 중심으로", 국어문학, 39집, 117-141, 2004.
- [3] 박시균, 권병로, "남북한 언어에 대한 비교 연구: 발음과 어휘를 중심으로", 국어문학, 38집, 29-52, 2003.
- [4] 육훈, "남북한 언어 차이", 새국어교육, 53호, 255-267, 1996.