

보완·대체의사소통도구 개발을 위한 학령기 아동 및 성인의 핵심어휘 조사¹

김영태* · 박현주** · 민홍기***

(*이화여자대학교 특수교육학과 · 언어병리학협동과정,

**이화여자대학교 언어병리학협동과정 박사과정

***인천대학교 정보통신공학과)

김영태 · 박현주 · 민홍기. 보완·대체의사소통도구 개발을 위한 학령기 아동 및 성인의 핵심어휘 조사. 『언어청각장애연구』, 2003, 제8권, 제2호, 93-110. 최근 심한 표현언어장애아동을 위한 보완·대체 의사소통 프로그램에 대한 관심이 매우 높아지고 있다. 본 연구는 보완·대체 의사소통 도구 개발의 첫 단계인 어휘데이터베이스를 구축하기 위해 일반인들이 많이 쓰는 어휘를 수집하고 그 중에서 고빈도 어휘 및 핵심어휘를 찾아내는 데 그 목적이 있다. 이를 위해 학교에서의 쉬는 시간과 일상생활에서 자발적인 발화를 수집하였다. 일상생활과 관련하여 음식점, 쇼핑센터, 미용실, 영화관, 비디오가게, 병원, 은행, 학교, 가정의 9가지 대화상황을 선정하였다. 연구자료는 학령기 아동 24명과 일반인 20명의 학교 및 상황별 발화를 전사하여 총 11,092개의 어휘를 가지고 분석하였다. 결과처리는 낱말의 출현빈도에 근거한 핵심어휘 찾기 및 구조어와 내용어의 비율 분석, 그리고 상위 고빈도 어휘가 전체어휘를 대표하는 비율 측정으로 이루어졌다. 연구 결과 아동, 청소년, 그리고 성인의 핵심어휘들 간에는 53%가 공통된 것으로 나타났으며 구조어와 내용어의 비율은 평균적으로 34.8%:65.2%로 나타났다. 또한 상위 고빈도 25개 어휘는 전체어휘의 34%를, 50개의 어휘는 전체어휘의 58%를 차지하는 것으로 나타났다.

핵심어: 보완대체의사소통, 표현언어장애, 핵심어휘, 구조어, 내용어

I. 서 론

최근 지역사회 통합 및 정상화 그리고 통합교육 등의 흐름과 더불어 장애인의 사회적 참여와 일반인과의 사회적 상호작용은 다양한 상황에서 극적으로 증가하고 있다. 장애인들이 교육, 사회, 종교, 여가, 직업 등에서 점점 성공적으로 참여하게 됨에 따라 이들의 의사소통 요구 또한 변화할 필요성이 제기되고 있다. 이러한 변화의 필요성을 증가시키고 현실화하는데 영향을 미치는 가장 영향력 있는 요인중의 하나가 테크놀로지의 발달이다. 테크놀로지의 급격한 발달은 심한 의사소통 장애를 지닌 사람들의 성공적인 의사소통 수행에 크게 영향을 주고 있으며 이러한 경향은

¹본 연구는 보건복지부 IMT-2000 출연금 기술개발지원사업(과제번호: 01-PJ11-PG9-01HT00-0019)과 인천대학교 멀티미디어 연구센터의 지원에 의하여 수행되었음.

본 연구의 자료수집을 도와 준 이주연, 최정아, 하선영, 이영미 선생에게 감사사를 드린다.

특히 구어를 사용하여 자신의 의사소통 요구를 충족시킬 수 없는 사람들을 위한 보완·대체 의사소통(Augmentative and Alternative Communication: 이하 AAC) 분야에서 특히 두드러지고 있다. AAC체계의 도움을 받는 심한 언어장애에는 발달장애(예: 뇌성마비, 정신지체, 자폐 관련 장애, 말실행증 등)나 후천성 뇌손상 또는 진행성 장애(예: 근위축성 측삭 경화증, 다발성 경화증, 파킨슨 증후군, 심한 실어증 등)가 주로 포함되는데 이들은 말, 언어, 쓰기 장애를 포함한 심한 표현 의사소통 장애를 보인다(Beukelman & Mirenda, 1992).

AAC분야에서 꾸준히 연구가 이루어진 부분의 하나가 어휘 선택에 관한 연구들이다. AAC 도구 개발에서 최초의 어휘 항목 선택은 극단적으로 중요하다고 한다(Yorkston et al., 1989). Carlson(1981)은 AAC도구가 성공적으로 사용되지 못한다면 우선 먼저 어휘를 살펴야 한다고 제안하였다. 의사소통에 어려움을 갖는 장애인이 AAC체계를 사용하여 효과적으로 대화할 수 있는지의 여부는 얼마나 적절한 어휘를 선택했느냐에 따라 그 성패가 좌우될 수 있기 때문이다. 빈약하게 선택되거나 부적절한 어휘는 성공적인 의사소통의 주된 방해물 중의 하나일 수 있으며 AAC 사용자의 좌절을 증가시키는 요인이 될 수 있다(Carlson, 1981).

AAC도구 사용자들은 자발적으로 말을 하거나 메시지를 철자로 표현할 수 없는 경우가 많아 의사소통에 필요한 어휘들을 다른 사람들이 선택해 주어야만 하는 경우가 대부분이며 이는 어린 장애아동의 경우에 특히 그렇다(Beukelman, Jones & Rowan, 1989).

따라서 AAC도구를 구성하는 어휘선택과 관련하여 AAC팀에 소속된 연구자들의 포괄적이고 세심한 노력과 주의가 요구되며 이를 위해서는 AAC도구 사용자의 개별적 요구와 관심사항, 의사소통 상황, 어휘의 기능성 등 다양한 요인들이 고려되어야만 한다.

AAC사용자에게 특정상황에서 요구되는 모든 어휘항목들을 선택할 수 있도록 할 수 있는 지식과 경험을 가진 사람은 드물다. 따라서, 다양한 자원들을 통해 어휘정보를 얻어야 한다. 일반적으로 어휘 선택을 위해 주로 사용되는 자원이 핵심어휘(core vocabulary)와 부수어휘(fringe vocabulary)이다.

핵심어휘는 다양한 사람들이 일반적으로 사용하는 어휘로서 매우 빈번하게 발생하는 낱말이나 메시지를 말하며 주로 구조어(structure word: 대명사, 조동사, 접속사, 전치사 등)로 구성된다. 핵심어휘 관별을 위한 방법에는 주로 세 가지 방법이 사용된다(Beukelman & Mirenda, 1998). 우선 성공적인 AAC사용자들의 어휘사용패턴에 기초한 낱말목록을 이용하는 경우로 2-3주간 AAC사용자들을 관찰하여 이들이 사용한 모든 어휘를 기초로 하여 가장 빈번하게 사용된 낱말을 관별하는 방법이다. 두 번째 방법은 개별 AAC사용자의 어휘사용패턴을 기초로 낱말 목록을 모으는 것으로 일정 시간동안 사용자가 산출한 모든 낱말들을 녹음하는 방법이 주로 사용된다. 세 번째 방법은 일반 대화자들간의 어휘사용패턴을 이용하는 것으로 해당 상황에서 일반인들이 나누는 대화를 녹음하는 방식을 주로 취하며 대부분의 연구들이 이 방법을 따르고 있다.

연구들에서 사용한 핵심어휘 판별 기준은 다양하다. 그러나 어휘의 발생 빈도를 기준으로 하여 빈도가 높은 어휘가 핵심어휘로 선택되어야 한다는 점에서는 일치를 보이고 있다(Balandin & Iacono, 1999; Beukelman, Jones & Rowan, 1989; Beukelman et al., 1984; Marvin, Beukelman & Bilyeu, 1994; Stuart, Beukelman & King, 1997; Yorkston et al., 1990). 해당 어휘의 출현이 얼마만큼의 빈도로 나타나야 하는가와 관련해서 연구들이 제시하는 기준은 다양하다. 그러나 주로 1000낱말 당 0.5회의 출현을 많이 사용하고 있다(Balandin & Iacono, 1999; Marvin, Beukelman & Bilyeu, 1994; Stuart, Beukelman & King, 1997). 반면에, 빈도 수치를 제시하지 않은 연구들도 있으며 발화에서 나타난 모든 어휘들을 종합하여 전체적인 누가 빈도를 계산하고 가장 높은 빈도의 어휘를 중심으로 상위 25-500개의 어휘를 선택하는 연구도 있다(Beukelman et al., 1984). 또한 연구 대상자의 50 % 또는 그 이상의 사람들이 공통적으로 사용한 낱말수의 견지에서 핵심어휘를 논의한 연구도 있다(Balandin & Iacono, 1999).

부수어휘(Beukelman & Mirenda, 1998)는 개별적인 의사소통 요구를 위해 제공되는 어휘 낱말이나 메시지를 말하며 주로 내용어(content word: 명사, 동사, 형용사, 부사 등)로 구성된다. 핵심어휘와 달리 상대적으로 많은 양의 낱말들이 포함되어 있을 뿐 아니라 사용자 개인에 관련된 낱말들을 포함한다. 즉, 자신의 이름, 특정 장소, 활동, 선호도 등이 포함된다. 따라서 핵심어휘 목록에서는 나타나지 않는 사고나 메시지의 표현을 가능하게 한다. 부수어휘는 AAC사용자 자신이나 이들과 이들의 의사소통 상황을 잘 아는 정보 제공자들(부모나 가족, 고용주나 직장동료, 친구 등)에 의해 추천되어야 한다.

어휘 수집에 대한 연구들에서 가장 관심을 갖는 부분은 물론 핵심어휘와 부수어휘이다. 이외에도 연구들에서 보여지는 분석대상을 살펴보면 어휘의 다양도(Type Token Ratio, 이하 TTR) (Balandin & Iacono, 1999; Beukelman, Jones & Rowan, 1989; Marvin, Beukelman & Bilyeu, 1994)와 구조어와 내용어의 비율 분석(Marvin, Beukelman & Bilyeu, 1994)을 들 수 있다. 후자와 관련한 연구(Beukelman, Jones & Rowan, 1989)에서 저자들은 50개의 가장 빈번하게 나타나는 낱말들에 대한 아동간 어휘 공통성을 보면 주로 구조어였다고 한다. Marvin, Beukelman & Bilyeu (1994)는 고빈도 상위 2,000개의 낱말을 분석한 결과 20 %가 구조어이고 80 %가 내용어였다고 한다. 따라서 내용어에 대한 핵심어휘가 의사소통을 위해 기능적이라고 하였다. 이와는 대조적으로 Arvidson & Lloyd(1997)는 AAC사용자의 대다수가 자신의 도구 속에 제한된 어휘(대부분이 내용어)를 저장하고 있는데 이는 자연스러운 의사소통을 방해할 수 있으며 전형적인 대화자가 되기 위해서는 좀 더 많은 구조어를 필요로 한다고 하였다.

또한 상위 고빈도 어휘가 전체어휘를 얼마나 대표할 수 있는가 하는 것도 연구의 초점이 된다. 통합 유치원 상황에서 또래의 일반 아동들이 사용하는 어휘의 사용 빈도수를 조사한 Beukelman, Jones & Rowan(1989)은 가장 빈번하게 발생하는 어휘들은 전체어휘에서 차지하는 비율이

매우 높음을 발견하였다. 1,000날말당 5회 이상 발화된 250개의 어휘를 분석한 결과 25개의 가장 빈번하게 발생하는 어휘는 전체어휘의 45.1%를 차지하였으며, 가장 빈번하게 발생하는 상위 50개의 어휘는 60%를, 상위 250개의 어휘는 전체 어휘의 85%를 차지하였다. 따라서 전체 의사소통 샘플의 대부분을 매우 적은 수의 상위 고빈도 어휘가 차지함을 알 수 있었다.

AAC도구 개발이 활발한 외국의 경우 AAC도구 개발을 위한 적절한 어휘 선택 노력은 무엇보다도 많은 관심을 받아왔다. 그러나 국내의 경우에는 이 분야에 대한 연구가 매우 적은 실정이다(김영태, 2002). 뇌성마비 초등학생을 대상으로 AAC체계를 위한 기초어휘 조사(박은혜, 1996), 정신지체 학생의 지역사회 기능에 필요한 기능적 어휘 목록 개발 연구(박승희, 1999), 그리고 AAC체계 적용을 위한 상황중심 핵심어휘 개발 연구(이정은·박은혜, 2000)가 이루어진 정도이다. 박승희(1999)의 연구는 정신지체인들의 지역사회 기능에 필요한 어휘 목록을 실제 지역사회에 나아가 표지판이나 로고, 메뉴판, 안내문 등을 중심으로 문자언어 대상의 어휘 조사가 이루어졌다.

박은혜(1996)의 연구는 9명의 초등 저학년 뇌성마비아동을 대상으로 하였으며 이들의 부모와 교사 그리고 동일 연령대의 일반아동과 이들의 부모로부터 설문조사를 하여 어휘를 수집하였다. 다만 일반아동의 경우에는 가정상황을 제외하고 학교상황에서는 직접 녹음을 하여 발화를 수집하는 방법을 사용하였으나 분석과정에서 녹음발화와 설문발화가 명확히 구분되지 않고 분석에 통합되어 그 특징이 두드러지지 않고 있다.

이정은·박은혜(2000)의 연구는 특수교사와 장애아동의 학부모 55명을 대상으로 설문조사를 실시하였다. 설문에 참여한 대상자들이 인식하는 의사소통 상황의 우선 순위를 결정하고 그 상황에서 구체적으로 필요하다고 생각되는 어휘를 조사하였다. 자료의 분석은 공통어휘 및 핵심어휘를 판별하는데 주안점을 두었다.

국내의 연구는 그 수에서도 제한적이지만 특히 장애아동을 대상으로 하고 있으며 주로 뇌성마비아동들이 포함되어있다. 또한 설문조사를 사용한 경우가 대부분이었으며 일반인들을 대상으로 의사소통 현장에서의 발화를 녹음하거나 수집한 경우가 거의 없어 아쉬움이 있다.

음성출력이 가능한 휴대용 AAC도구 개발을 위한 위탁연구로 수행된 본 연구의 목적은 AAC프로그램 개발의 첫 단계인 어휘데이터베이스 구축을 위한 어휘의 수집과 분류에 있다. 따라서 AAC도구의 최종 사용자인 언어장애인의 의사소통을 효율적으로 돕고 성공적인 대화자가 되도록 하기 위하여 이들이 주로 접하는 사회적 상황과 이들 상황에서 일반인들이 사용하는 어휘에 대한 정보를 수집하여 앞으로 사용할 잠재적인 어휘에 대한 충분한 자료를 제공하는 데 도움이 되고자한다.

따라서 본 연구에서는 학령기 아동 및 일반인들이 어떠한 방식으로 어떠한 어휘를 사용하여 의사소통 하는가를 분석하기 위하여 (1) 초등학생과 중·고등 학생들을 대상으로 학교 쉬는 시

간에 보이는 발화를 분석하여 연령간 공통적인 핵심어휘를 찾아내고자 하였으며, (2) 고등학생 및 성인이 여러 상황에서 공통적으로 사용하는 상황간 핵심어휘를 찾아내고자 하였다.

II. 연구방법

1. 연구대상

본 연구에서 분석한 발화들은 다음과 같은 연구 대상자들로부터 수집되었다.

- 가. 초등학교 3-4학년아동: 서울지역의 일반 D초등학교에 재학중인 남녀 12명(남6, 여6)의 아동들로 담임 교사에 의하여 정상 발달을 보이는 것으로 보고되었으며 특별한 감각장애나 언어/학습 문제 그리고 행동/정서상의 문제를 보이지 않는 아동들이 참여하였다.
- 나. 중학생: 서울지역의 E중학교와 Y중학교에 재학중인 학생들로 학교장의 추천을 받아 연구부장의 도움으로 선정된 12명(남6, 여6)의 남녀 학생들이 연구에 참여하였다.
- 다. 고등학생: 서울지역의 E고등학교와 Y고등학교에 재학중인 학생들로 학교장의 추천을 받아 연구부장의 도움으로 선정된 12명(남6, 여6)의 남녀 학생들이 연구에 참여하였다.
- 라. 성인: 상황별 발화를 제공한 참여자들은 연구자들이 발화를 수집하기 위하여 찾아가는 장소에서 만난 정상 성인 20명(남8, 여12)이었다.

처음에 연구에 참여한 총인원은 56명이었으나 녹음 상태가 나빠 발화전사가 어렵거나 전사 가능한 발화가 매우 적은 경우는 제외하였다. 따라서 최종적으로 연구에 참여한 인원은 초등 8명(남4, 여4), 중등 8명(남5, 여3), 고등 8명(남4, 여4), 성인 20명으로 총 44명이었다.

2. 발화 수집

초등학생의 경우에는 연구자의 조언에 따라 담임 교사가 직접 자유놀이 시간 동안에 이루어지는 아동들의 대화를 녹음하였다. 중학생과 고등학생의 발화는 연구자들의 조언에 따라 연구부장이 직접 학생들을 대상으로 녹음기 조작과 적절한 녹음 방법을 훈련한 후 자유시간 동안에 이루어지는 학생들 간의 대화내용을 직접 녹음하게 하는 방법을 취하였으며 나중에 녹음기와 녹음 테이프를 수거하였다. 연구에 사용된 녹음기는 SONY TCM-100이었으며 모든 연령의 학생들을 대상으로 동일한 기기가 사용되었다. 발화의 녹음기간은 2002년 4월-5월에 주로 이루어졌으며 녹음에 소요된 시간은 학교 사정에 따라 1일에서 2주일 정도까지 다양하였다.

아동의 발화를 낱말로 분석한 결과 아동마다 나타나는 낱말의 양이 다양하였다. 따라서 가장 적은 수의 낱말을 보인 아동을 기준으로 낱말수를 통일하였다. 그 결과 각 아동당 최초의 305개 낱말이 선택되었으며 최종적으로 총 7,320개의 낱말이 분석에 포함되었다.

상황별 발화는 연구자들이 지역사회 공공장소(미용실, 영화관, 병원, 은행, 비디오가게, 백화점, 옷가게)를 다니면서 일반인들이 하는 말을 녹음하여 전사하였다. 이때 사용된 녹음기는 학교상황 녹음을 위해 사용한 것과 동일한 기기가 사용되었다. 가정상황의 경우는 5가족의 TV보기, 옷입기, 식사하기 등의 활동 속에서 나오는 발화를 수집하였다. 이들 과정을 통해 얻어진 발화는 낱말로 분석했을 때 총 3,090개로 각 상황별 낱말수는 영화관 178개, 병원 155개, 미용실 351개, 은행 275개, 비디오가게 158개, 쇼핑 281개, 음식점 520개, 가정 1,208개였다.

3. 발화 전사

발화의 전사는 정확도를 높이기 위하여 녹음과 함께 실제 상황에서 기록한 자료를 토대로 녹음 내용을 들어가면서 두 명의 전사자(언어병리학 석사, 언어병리학 박사)가 함께 작업하였다. 학교 상황의 경우에는 현장에서 기록한 자료가 없었기 때문에 두 명의 전사자가 여러 번에 걸쳐 정확한 전사를 위하여 녹음된 발화를 청취하였다. 이들 두 명의 전사자는 사전에 전사원칙을 숙지한 상태였으며 의견에 일치할 보이지 않는 발화는 전사에서 제외하였고 낱말의 반복은 ‘강조’의 경우가 아닌 경우에는 제외하였다. 부정확하여 알아들을 수 없는 경우에는 그 발화 전체를 전사에서 제외하였다. 우물쭈물하거나 오류를 보였어도 전사자들이 예측 가능하여 상호 일치를 보인 경우에는 정확한 낱말로 바꾸어 전사하였다.

4. 발화 분석

본 연구의 목적은 AAC개발을 위한 내용어 분석에 있으므로 분석에 사용한 낱말은 어휘형태소이며 문법형태소는 제외하였다. 각 낱말은 다음과 같은 전사원칙에 의하여 입력되었다.

- 가. 동사의 경우에는 시제의 구분 없이 원형으로 기재하였다. 예를 들면, ‘했었다’는 ‘하다’로 표시하였다.
- 나. 축약된 말, 즉 ‘하면’과 같은 경우에는 이를 두 낱말로 분리하여 ‘하다’와 ‘그러면’으로 기재하였다.
- 다. 존대어의 경우에도 평어의 형태로 바꾸어 기재하였다. 예를 들면, ‘저’ 또는 ‘제가’와 같은 말을 ‘나’로 기재하였다.

입력된 낱말은 다음의 측면에서 분석되었다.

- 가. 빈도 분석: 낱말이 출현한 빈도를 중심으로 고빈도 순위를 산출하였다.
- 나. 핵심어휘 선정: 연령집단 간(초, 중, 고), 상황 간(9상황)에 50-60% 이상의 연령이나 상황에 공통적으로 사용되고 10회 이상의 빈도를 보이는 어휘로 선정하였다.
- 다. 구조어와 내용어의 비율 분석: 상위 고빈도 어휘를 중심으로 10, 20, 30, 40, 50개의 어휘 간에 나타나는 구조어(대명사, 관형사, 접속사, 감탄사)와 내용어(명사, 동사, 형용사, 부사)의 비율을 산출하였다.
- 라. 고빈도 어휘가 전체어휘에서 차지하는 비율 분석: 학령기 아동의 상위 10, 25, 50, 75개의 고빈도 어휘가 전체어휘에서 차지하는 비율을 계산하여 어느 정도의 적은 어휘가 전체어휘를 대표할 수 있는가를 살펴보았다.

III. 연구결과

1. 학령기 아동의 발화분석 결과

가. 고빈도 어휘 분석

각각의 집단은 아동 8명당 305개의 낱말씩 총 2,440개의 낱말이 포함되었다. 고빈도 어휘의 제시는 각 집단별로 20회 이상 나타난 어휘만을 대상으로 하였다(<표 - 1> 참조). 고빈도 어휘는 일반아동들이 의사소통 상황에서 공통적으로 사용할 가능성이 큰 것으로 해석할 수 있다. 총 2,440개의 어휘 중 20회 이상의 고빈도로 나타난 어휘는 초등부 집단에서 13개 낱말, 중학부 집단에서 21개 낱말, 그리고 고등부 집단에서 23개의 낱말이었다. 이들 중 세 집단에서 모두 공통적으로 사용된 낱말은 ‘하다’, ‘나’, ‘너’, ‘야’, ‘것’, ‘안’, ‘되다’, ‘왜’, ‘보다’의 9개로 이는 초등학생들이 보인 13개의 어휘 중 69%를 차지하는 것으로 나타났다. 또한 나머지 4개의 어휘 중 3개 즉, ‘그래’, ‘다’, ‘아니다’는 두 집단 이상에서 공통적으로 사용하는 것으로 나타나 핵심어휘로 선정될 가능성이 크다고 할 수 있다.

이들 집단의 어휘다양도는 <표 - 2>에 제시되어 있다. 각 집단별 어휘다양도는 초등학생이 .44 (.38 - .50)로 중학생의 .52 (.49 - .60), 고등학생의 .52 (.44 - .59)에 비해 낮은 것으로 나타났다.

<표 - 1> 집단별 고빈도 어휘

초등부	빈도	중학부	빈도	고등부	빈도	전체	빈도	전체	빈도
하다	90	하다	109	하다	87	하다	286	잘	43
나	57	나	95	것	77	나	214	아(감)	42
너	43	가다	53	야	64	것	157	그렇다	41
야	35	것	50	애	64	너	138	이렇게	39
것	30	안	47	나	62	야	129	응	36
안	24	보다	46	너	50	있다	120	모르다	33
되다	24	너	45	보다	44	보다	110	말다	32
다	24	있다	41	그	42	안	109	어디	32
아니다	22	무엇	39	있다	39	애	104	여기	31
그래	22	주다	31	안	38	무엇	94	좋다	30
왜	21	오다	30	무엇	37	되다	74	맞다	29
그러면	21	야	30	그렇다	37	아니다	68	그런데	25
보다	20	되다	28	가다	34	왜	67	너무	24
		아니다	26	아(감)	31	주다	66	더	24
		알다	25	진짜	26	다	57	어떻게	24
		왜	24	그런데	25	알다	55	좀	23
		없다	24	알다	23	오다	55	빨리	22
		애	23	주다	22	이거	53	싫다	22
		그래	20	왜	22	없다	46	나오다	21
		이거	20	되다	22	그거	44	또	21
		좋다	20	이거	21	먹다	44	못	21
				먹다	20	그러면	44	죽다	21
				다	20	우리	43	많다	20

<표 - 2> 각 집단 별 어휘다양도

집단(N = 8)	전체 낱말수	다른 낱말수	어휘다양도
초등학생	2,440	1,068	.44
중학생	2,440	1,275	.52
고등학생	2,440	1,279	.52

나. 핵심어휘 분석

핵심어휘는 초등부, 중학부, 고등부 집단 모두에서 10회 이상의 빈도를 보이고 두 집단 이상

에서 공통적으로 사용되는 어휘로 선정하였다. 그 결과 <표 - 3>에서와 같이 31개의 고빈도 어휘 중 27개의 어휘가 핵심어휘로 선정되었다. 각 연령에서 모두 나타나는 다소 높은 기준을 쓴 이유는 이 발화표본이 24명의 아동들의 표본을 종합한 것이기 때문이다.

학령기 아동의 상위 10개의 핵심어휘를 살펴보면, ‘하다’, ‘나’, ‘것’, ‘너’, ‘야’, ‘있다’, ‘보다’, ‘안’, ‘애’, ‘가다’로 대부분의 어휘가 100번 이상의 고빈도로 사용됨을 알 수 있다.

핵심어휘를 품사별로 살펴보면 우선 동사가 가장 많았다. 특히 가장 많이 사용된 핵심어휘 1위도 동사 ‘하다’가 차지하였으며 286회 이상의 빈도로 산출되었다. 예상과 달리 명사에 해당하는 낱말은 ‘것’과 ‘애’ 둘 뿐이었다. 또한 대명사 어휘가 명사 어휘보다 더 다양하게 나타났는데 즉, ‘나’, ‘너’, ‘이거’, ‘그거’, ‘우리’, ‘어디’, ‘무엇’, ‘여기’가 핵심어휘에 포함되었다. 핵심어휘에 포함된 접속사는 ‘그러면’ 하나였다.

<표 - 3> 학령기 아동의 전반적 고빈도 어휘 및 핵심어휘

순위	낱말	빈도	순위	낱말	빈도	순위	낱말	빈도
1	하다	286	13	아니다	68	23	그러면	43
2	나	214	14	왜	67	23	우리	43
3	것	157	15	주다	66	23	잘	43
4	너	138	16	다	57	24	아(감)	42
5	야	129	17	알다	55	25	그렇다	41
6	있다	120	17	오다	55	26	이렇게	39
7	보다	110	18	이거	53	27	응	36
8	안	109	19	진짜	49	28	모르다	33
9	애	104	20	그래	48	29	말다	32
10	가다	95	21	없다	46	29	어디	32
11	무엇	94	22	그거	44	30	여기	31
12	되다	74	22	먹다	44	31	좋다	30

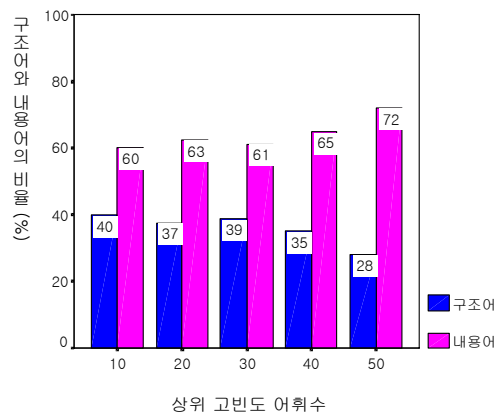
음영된 낱말은 핵심어휘임.

다. 구조어와 내용어의 비율 분석

<표 - 1>에 나타난 전체(초, 중, 고) 빈도를 중심으로 최초의 고빈도 어휘 50개에 대한 구조어와 내용어의 비율을 분석한 결과는 <그림 - 1>과 같다. 구조어는 대명사, 접속사, 관형사 등 연속발화를 위한 틀로서 사용되는 낱말을 말하며 내용어는 각 발화의 전형적인 의미를 제공할 수 있는 것으로 명사, 동사, 형용사, 부사를 말한다(Bowen, Madsen & Hilferty, 1985). <그림 - 1>에서 볼 수 있는 바와 같이 상위 고빈도 어휘에서 내용어가 차지하는 비율은 구조어에 비해 더 높았

다. 이러한 현상은 상위 10개에서 40:60이던 비율이 상위 50개의 어휘에서는 28:72로 더 벌어지는 것으로 나타났다.

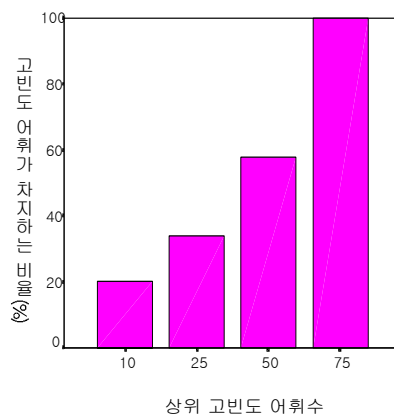
구조어와 내용어의 비율 분석에서 나타나는 품사별 출현 빈도를 살펴보면 동사가 1,378, 명사가 790, 대명사가 676, 부사가 582, 형용사가 289회로 동사가 가장 많이 사용되는 것으로 나타났다.



<그림 - 1> 구조어와 내용어의 비율

라. 전체 어휘 중 고빈도 어휘가 차지하는 비율 분석

전체 발화에서 상위 고빈도 어휘가 차지하는 정도를 살펴보기 위하여 상위 빈도 10, 25, 50, 75개의 어휘를 중심으로 분석을 하였다. 이를 그림으로 나타내면 <그림 - 2>와 같다.



<그림 - 2> 전체어휘 중 고빈도 어휘가 차지하는 비율

얼마나 적은 어휘로 전체어휘를 대표할 수 있는가는 핵심어휘 선정의 중요성을 반증하는 것일 수 있다. 앞의 그림에서 살펴보았듯이 최초의 고빈도 어휘 10개는 전체 어휘의 20%를 차지한다. 상위 고빈도 어휘 25개는 전체어휘의 34%를 설명해 주며 상위 50개의 어휘는 전체어휘의 58%를 설명해 준다. 마지막으로 본 연구의 빈도순위는 57위에 이르면 연구에 포함된 모든 전체어휘를 포함하게 되는 것으로 나타났다. 결과적으로 상위 소수의 고빈도 어휘는 전체어휘의 많은 부분을 설명해 줄 수 있음을 알 수 있다.

2. 상황별 발화분석 결과

가. 상황별 고빈도 발화 분석

각 상황에서 전체 낱말수 중 상위 2%에 해당하는 낱말들을 제시한 결과는 <부록 - 1>에 나와 있다. 이에 따르면, 이들 성인의 상황별 고빈도 어휘 대부분이 아동들 발화에서의 고빈도 어휘에 포함되는 것들을 알 수 있다. 예를 들어, ‘하다, 되다, 주다, 있다, 아니다, 무엇, 안, 이거, 나’ 등은 아동들의 고빈도 및 핵심 어휘에 포함되는 것들이다.

그 외 상황별로 많이 나타난 어휘들은 (1) 영화관상황에서는 시간과 관련된 ‘분’이나 인원수와 관련된 ‘명’이 고빈도 어휘에 포함되었고, (2) 병원상황에서는 ‘아프다, 가렵다, 심하다’ 등의 상황과 관련된 어휘들이 포함되었다. (3) 미용실상황에서는 ‘머리, 자르다’, (4) 은행상황에서는 ‘넣다’, (5) 비디오가게상황에서는 ‘보다, 재미있다’, (6) 쇼핑상황에서는 ‘괜찮다, 입어보다’, (7) 가정상황에서는 ‘엄마’ 그리고 (8) 음식점상황에서는 ‘빅맥세트, -개, 하나’가 상황적인 고빈도 어휘로 보인다.

<표 - 4> 각 상황별 어휘다양도

상황	총 낱말수	다른 낱말수	어휘다양도
음식점	520	152	.29
학교	646	340	.53
쇼핑	281	88	.31
가정	1,208	341	.28
미용실	351	164	.47
영화관	178	68	.38
비디오가게	158	71	.45
병원	155	101	.65
은행	275	122	.44

9가지 상황(음식점, 학교, 쇼핑, 가정, 미용실, 영화관, 비디오가게, 병원, 은행)에서의 청소년(고등학생) 및 성인의 발화를 수집하여 각각의 상황에서 나타나는 낱말 유형 및 빈도를 분석하였다. <표 - 4>에서와 같이 각 상황에서 수집한 낱말수(token)는 152-646개였는데, 그 중 다른 낱말의 형태(type)는 71-341개로 어휘다양도(TTR)는 .29 - .65를 보여주었다.

나. 핵심어휘 분석

상황간 핵심어휘를 찾기 위하여 우선 전체 상황에서 10회 이상 나타난 어휘들을 <부록 - 2>에, 9상황 중 5상황 이상에서 적어도 1회 이상 공통적으로 나타난 어휘들을 <표 - 5>에 정리해 보았다. 그 결과 5상황(56%) 이상에서 적어도 1회 이상 공통적으로 사용되었으며 그 빈도가 10회를 넘는 낱말들이 <표 - 5>의 음영으로 표시된 34개의 낱말들이다. 특이한 것은 9상황(100%) 모두에서 나타난 어휘 ‘하다, 되다, 이거, 여기’는 모두 대명사, 대동사, 또는 지시어에 해당하는 것이었다. 이들 34개 낱말들 중 일반아동들의 발화에서도 핵심어휘로 선정된 낱말은 18개로 53%가 공통된 것이었다. 즉, 이들 34개 어휘들은 특정 상황에만 쓰이는 것이 아니라 어떤 상황에서도 가장 공통적으로 많이 쓰이는 어휘라고 할 수 있으며, 그 중 18개는 연령과 상황에 상관없이 공통적으로 가장 많이 쓰이는 어휘라고 할 수 있다.

<표 - 5> 56% 이상의 상황에서 나타난 어휘들

5상황		6상황		7상황		8상황		9상황	
어휘	빈도	어휘	빈도	어휘	빈도	어휘	빈도	어휘	빈도
것	32	나	95	오다	25	주다	126	하다	143
가다	27	보다	76	많다	23	네	81	되다	119
얼마	15	무엇	49	잘	17	안	72	이거	91
왜	14	같다	30	그리고	17	있다	57	여기	38
사다	12	그거	28	시간	14	아니다	39		
(나)오다	11	다	24			좀	38		
모르다	11	맞다	14			그냥	29		
그러나	10	몇	12			없다	17		
감사하다	9	그렇다	12						
수고하다	9	어떻게	8						
또	9								
빨리	6								
얼마나	5								

음영된 낱말은 9상황 중 5상황 이상에서 10회 이상 나타난 낱말

IV. 논의 및 제언

본 연구는 AAC체계에 내장될 어휘를 결정하기 위해 학령기 아동 및 일반인들을 대상으로 학교와 다양한 지역사회 상황에서 오고가는 자발적인 대화를 중심으로 어휘를 수집하였다. 수집된 발화는 낱말 단위로 분석되어 빈도가 결정되고 연령집단간, 상황간 핵심어휘가 선정되었으며 각 집단의 어휘다양도가 분석되었고, 학령기 아동들을 대상으로 구조어와 내용어의 비율, 상위 고빈도 어휘가 전체어휘에서 차지하는 비율이 분석되었다. 따라서 빈도분석 및 핵심어휘 선정결과, 구조어와 내용어의 비율 분석 결과, 그리고 상위 고빈도 어휘가 차지하는 비율 분석 결과를 중심으로 논의를 진행하고자 한다.

첫째, 학령기 아동의 전반적 고빈도 어휘 및 핵심어휘 분석결과 명사의 사용이 예상과 달리 매우 적었다. 이는 박은혜(1996)의 연구 중 일반아동의 녹음 결과를 중심으로 분석한 부분과 일치한다. 또한 핵심어휘에 포함된 동일 어휘의 비율이 상대적으로 높았다. 이는 이들 두 결과가 동일한 방법을 썼기 때문으로 보인다. 또한 아동의 핵심어휘와 상황간 핵심어휘를 살펴보면 53%가 공통된 것으로 나타났다. 이는 연령과 상황에 상관없이 공통적으로 사용되는 핵심어휘 선정의 타당성을 입증하는 선행연구(Beukelman et al., 1984; Beukelman, Jones & Rowan, 1989; Marvin, Beukelman & Bilyeu, 1994; Stuart et al., 1993; Yorkston et al., 1989; Yorkston et al., 1990)들과 일치하였다. 또한 모든 집단에서 나타난 핵심어휘는 대명사, 동사, 지시어 등이었는데 이는 이들 유형의 낱말 사용이 연령간 상황간에 공통성이 더 높고 다양한 상황간에 핵심어휘로서의 유용성이 높다는 Marvin, Beukelman & Bilyeu(1994)과 Yorkston et al.(1989)의 연구와 일치한다.

둘째, 구조어와 내용어의 비율과 관련하여 이 둘간의 비율이 중요시되는 이유는 기존의 AAC체계 내에 제한된 어휘가 들어있음으로 인해서 의사소통에 방해가 되었다는 점과 밀접하게 연관되어 있다. 즉, AAC도구들이 주로 내용어만을 배타적으로 포함하고 있어 사용자의 의사소통 의도와 언어수준에 걸맞는 문장을 산출할 수 없는 경우가 많고 주로 전보식 문장을 종종 사용하게 된다고 한다(Marvin, Beukelman & Bilyeu, 1994). AAC체계에 내용어와 구조어의 낱말 유형을 충분히 넣어주면 어린 아동의 경우에는 언어발달을 촉진할 뿐 아니라 아동의 초기 의사소통 선호도를 뛰어넘는 언어적 상호작용을 보인다는 것이다. 만약 일반아동들을 대상으로 수집한 자료가 내용어에 비해 구조어의 비율이 높다면 이러한 문제는 특히 설득력을 갖는다고 할 수 있다. 본 연구에서 살펴본 내용어와 구조어의 비율은 상위 고빈도 50개 어휘를 분석했을 때 빈도가 높을수록 구조어 비율이 높은 반면 어휘수가 증가할수록 내용어의 비율이 증가하는 경향을 보여 주었다. 상위 20개 어휘에서 나타나는 구조어와 내용어의 비율은 37%와 63%이었다. 반면에 50개 어휘에서 나타나는 비율은 23%와 72%로 나타났다. 따라서 구조어와 내용어의 비율은 내용어가 배타적으로 많은 것으로 나타났다. 이는 Marvin, Beukelman & Bilyeu(1994)의 연구 결과와 유사

하였다. 그러나 핵심어휘내에서는 구조어의 비율이 내용어의 비율에 비해 높은 것으로 나타났다. 또한 본 연구에서는 어휘 형태소만을 대상으로 하였으므로 구조어와 내용어의 비율에 대한 판단에 있어서 유의가 필요하다. 따라서 AAC체계내의 어휘에는 내용어만큼 구조어도 풍부하게 저장되어야 하며 품사별로 봤을 때도 명사류만의 배타적인 저장은 지양되어야 할 것이다.

셋째, 고빈도 어휘가 전체어휘에서 차지하는 비율 분석과 관련하여 상위 25개의 어휘는 전체어휘의 34%, 50개 어휘는 58%를 차지하는 것으로 나타났다. 이는 Beukelman, Jones & Rowan(1989)의 25개의 가장 빈번하게 발생하는 어휘는 전체 어휘의 45.1%를 차지하였고, 50개의 어휘는 60%를 차지하였다는 연구와 Beukelman et al.(1984)의 25개의 고빈도 어휘가 전체 샘플의 35%를 차지하였고 200개의 어휘가 전체어휘의 65%를 차지하였다는 연구 결과들과 유사함을 알 수 있었다. 제시된 비율이 연구마다 조금씩 차이가 나는 것은 아마도 전체어휘수에서의 차이에 기인하는 것으로 보인다. 전체어휘가 더 많을 경우 소수의 상위 어휘가 전체어휘를 대표하는 비율이 증가할 것으로 보인다. 결국 소수의 고빈도 어휘가 전체어휘를 대표할 수 있다는 점에서는 모든 연구들이 공통성을 보여 상위 고빈도 어휘와 핵심어휘를 개발하는 데 좀 더 많은 관심과 노력을 기울여야 할 것이다.

일반아동을 대상으로 한 본 연구 결과의 시사점은 결국 빈번하게 발생하는 핵심어휘의 존재는 AAC체계를 설계하고 만들 때 중요하게 고려되어야 할 요인이며 AAC체계를 만들 때 극단적으로 많은 핵심어휘를 필요로 하지 않을 뿐 아니라 개개인의 의사소통 요구에 따라 부가적인 내용어들이 융통성 있게 포함될 수 있어야 한다는 점이다.

본 연구는 학교 상황과 지역사회 상황에 직접 나가 자발적인 대화를 녹음하여 그 자료를 중심으로 고빈도 어휘 및 핵심어휘를 제시한 최초의 연구로서 그 의의가 있다. 본 연구의 결과를 바탕으로 앞으로의 연구를 위한 제언을 하면 우선 본 연구는 분석에 포함된 전체 낱말의 수가 외국 연구들에 비해 상대적으로 적었다는 것이다. 따라서 더 많은 어휘를 포함한 연구가 지속적으로 이루어짐으로써 AAC도구에 저장될 수 있는 핵심어휘 목록이 다양하게 개발될 필요가 있을 것이다. 둘째, 구조어와 내용어 분석과 관련하여 본 연구에서는 어휘 형태소에만 국한한 분석을 실시하였다. 따라서 우리말에서 중요한 기능을 갖는 문법 형태소(예: 조사)에 대한 고찰은 하지 않았다. 그러므로 본 연구의 결과를 일반화하는데는 주의가 필요하며 이를 고려한 후속연구가 이루어져야 할 것이다.

참 고 문 헌

김영태(2002). 언어장애인이 사용하는 어휘의 빈도 및 중요도 연구. 정보통신부 IMT-2000 기술개발지원 Project 최종보고서.

- 박은혜(1996). 보완·대체 의사소통체계를 위한 기초어휘 조사: 뇌성마비 초등 저학년을 중심으로. 『특수교육논총』, 13(1), 91-115.
- 박은혜·이정은(2000). 장애아동을 위한 웹기반 보완·대체 의사소통도구 개발연구. 『특수교육학연구』, 35(3), 17-44.
- 박승희(1999). 정신지체 학생의 지역사회 기능에 필요한 기능적 어휘 목록 개발 연구. 『재활복지』, 3(1), 23-57.
- 이정은·박은혜(2000). 보완·대체 의사소통체계 적용을 위한 상황 중심 핵심어휘 개발 연구. 『재활복지』, 4(1), 96-121.
- Arvidson, H. & Lloyd, L.(1997). Vocabulary selection. In L. Lloyd, D. Fuller & H. Arvidson (Eds.), *Augmentative and alternative communication: A handbook of principles and practices*. Boston: Allyn & Bacon.
- Balandin, S. & Iacono, T.(1999). Crews, wusses. and whopapas: Core and fringe vocabularies of Australian meal-break conversations in the workplace. *Augmentative and Alternative Communication*, 15, 95-109.
- Beukelman, D., Yorkston, K., Poblete, M. & Naranjo, C.(1984). Frequency of word occurrence in communication samples produced by adult communication users. *Journal of Speech and Hearing Disorders*, 49, 360-367.
- Beukelman, D., Jones, R. & Rowan, M.(1989). Frequency of word usage by nondisabled peers in integrated classrooms. *Augmentative and Alternative Communication*, 5, 243-248.
- Beukelman, D. & Mirenda, P.(1992). *Augmentative and alternative communication: Management of severe communication disorders in children and adults*. Baltimore: Paul H. Brookes.
- Beukelman, D. & Mirenda, P.(1998). *Augmentative and alternative communication: Management of severe communication disorders in children and adults*(2nd ed.). Baltimore: Paul H. Brookes.
- Browen, J., Madsen, S. & Hilferty, A.(1985). *TESOL: Techniques and procedures*. Boston: Heinle & Heinle.
- Carlson, F.(1981). A format for selecting vocabulary for the nonspeaking child. *Language, Speech, and Hearing Service in Schools*, 12, 240-248.
- Marvin, C., Beukelman, D. & Bilyeu, D.(1994). Vocabulary-use patterns in preschool children: Effects of context and time sampling. *Augmentative and Alternate Communication*, 10, 224-236.
- Stuart, S., Vanderhoof, D. & Beukelman, D.(1993). Topic and vocabulary use patterns of elderly woman. *Augmentative and Alternative Communication*, 9, 95-110.
- Stuart, S., Beukelman, D. & King, J.(1997). Vocabulary use during extended conversations by two cohorts of older adults. *Augmentative and Alternative Communication*, 13, 40-47.
- Yorkston, K., Honsinger, M., Dowden, P. & Marriner, N.(1989). Vocabulary selection: A case report. *Augmentative and Alternative Communication*, 15, 101-108.
- Yorkston, K., Beukelman, D., Smith, K. & Tice, R.(1990). Extended communication samples of augmentative communicators II: Analysis of multi-word sequences. *Journal of Speech and Hearing Disorders*, 55, 225-230.

<부록 - 1> 상황별 고빈도 어휘: 전체 낱말 빈도수 중 2% 이상 출현한 고빈도 어휘

빈도	영화관	빈도	병원	빈도	미용실	빈도	은행	빈도	비디오
15	네	8	네	26	하다	22	주다	15	보다
13	시	6	아니다	14	네	21	네	9	주다
9	되다	5	가렵다	13	머리	10	이거	8	것
9	주다	4	같다	13	주다	8	넣다	7	아줌마
7	꺼	4	아프다	8	안	7	다	7	재미있다
7	분	4	오다	7	그냥	6	되다	6	그거
6	자리	4	조금	7	차르다	6	여기	5	나오다
6	장	3	되다	7	조금	6	하다	5	무엇
5	그거	3	몸					5	있다
5	있다	3	심하다					4	가져오다
5	하다							4	그럼
4	그러면							4	
4	명								
4	없다								
4	좋다								

빈도	쇼핑	빈도	가정생활	빈도	음식점	빈도	학교상황
39	이거	82	되다	52	주다	23	나
16	괜찮다	59	나	38	하다	12	그런데
13	어떻다	51	엄마	21	무엇	12	안
9	수 예) 33	47	하다	18	-개	12	입다
9	언니	43	보다	15	좀		
7	아니다	36	안	14	네		
7	입어보다			13	여기		
6	다른데			12	빅맥세트		
6	더			12	하나		
6	보다						
6	있다						

<부록 - 2> 전체 상황에서의 고빈도 어휘: 10회 이상 사용된 어휘

순위	빈도	낱말	순위	빈도	낱말	순위	빈도	낱말
1	132	하다	21	22	머리*	41	13	어떻다*
2	123	주다	22	22	아니다	42	12	그애*
3	107	되다	23	21	같다	43	12	그*
4	105	나	24	21	그거	44	12	많이
5	80	보다	25	21	입다*	45	12	빅맥세트*
6	76	이거	26	20	좀	46	12	아줌마*
7	72	네	27	18	()개*	47	12	자리*
8	54	무엇	28	18	더*	48	11	말다*
9	54	안	29	18	오늘*	49	11	(나)오다
10	51	엄마*	30	18	하나*	50	11	없다
11	44	있다	31	17	그냥	51	10	콜라*
12	42	먹다*	32	17	나가다*	52	10	그래서*
13	31	그리고	33	17	싫다*	53	10	그러나
14	28	가다	34	17	아프다*	54	10	그렇다
15	26	웃*	35	16	괜찮다*	55	10	그리다*
16	32	것	36	16	막*	56	10	말*
17	24	그런데*	37	15	다	57	10	맨날*
18	24	여기	38	14	언니*	58	10	바지*
19	23	그러면*	39	13	시			
20	22	너*	40	13	아빠*			

*5상황(56%) 미만에서만 나타난 어휘들

ABSTRACT

School-Aged Children and Adults's Core Vocabulary for the Development of an Augmentative and Alternative Communication Tool

Young Tae Kim (Dept. of Special Education & Interdisciplinary Program
of Communication Disorders, Ewha Womans University)

Hyeon Ju Park (Interdisciplinary Program of Communication Disorders,
The Graduate School, Ewha Womans University)

Hong Ki Min (Dept. of Information and Telecommunication
Engineering, University of Incheon)

Recently the interest in the augmentative and alternative communication program for expressive language disorders is increasing. The purpose of this present research is to collect vocabulary used by common people to build vocabulary data bases, which is the first step for a development of the augmentative and alternative communication tool, and to find the highly frequent vocabulary and the core vocabulary. To do this, we collected spontaneous speech samples in break times of the school and in our daily lives. We selected nine conversation situations related to our daily living places such as restaurants, shopping centers, beauty shops, movie theaters, VCR rental shops, clinic centers, banks, schools, and homes. The analyses were carried out for about 11,092 vocabulary items sampled from 24 school-aged children and 20 adults. The results of the research showed that there is a difference of 53 % of common core vocabulary items between school-aged children and adults. The ratio of structure words vs content words was 34.8 % : 65.2 %. Also, the highly frequently occurring first 25 words represented 34 %, and the first 50 words represented 58 % of the total communication samples.

Key Words: augmentative and alternative communication, expressive language disorder, core vocabulary, structure word, content word

▶ 게재 신청일: 2003년 6월 9일

▶ 게재 확정일: 2003년 7월 16일

▶ 김영태(교신 저자): 이화여자대학교 특수교육학과 및 언어병리협동과정 교수, e-mail: youngtae@ewha.ac.kr

▶ 박현주(제 2 저자): 이화여자대학교 언어병리학 협동과정 박사과정, e-mail: phj8747@hanmail.net

▶ 민홍기(제 3 저자): 인천대학교 정보통신공학과 교수, e-mail: hkmin@incheon.ac.kr