A Literature Review of Neural Tracking to Speech Envelope from the View of Audiology

Sungmin Lee

Department of Speech-Language Pathology and Audiology, Tongmyong University, Busan, Korea

Received: January 19, 2021

Revised: March 24, 2021

Accepted: April 8, 2021

Correspondence:

Sungmin Lee, PhD
Department of Speech-Language
Pathology and Audiology,
Tongmyong University, 428
Sinseon-ro, Nam-gu, Busan 48520,
Korea
Tel: +82-51-629-2134
Fax: +82-51-629-2019
E-mail: slee18@tu.ac.kr

Despite the significant contribution of hearing assistive devices, medications, and surgery to restoring auditory periphery, a large number of people with hearing loss still struggle with understanding speech. This leads many studies on speech perception to move towards the central auditory functions by looking at associated brain activities using macroscopic recording tools such as electroencephalography (EEG). Up until a few years ago, however, limitation has been given to the brain scientists who attempted to investigate speech perception mechanisms using the EEG. In particular, short duration of speech segments has inevitably been used to elicit auditory evoked potential, even though they were too brief to be considered as speech. Today, however, advance in neural engineering and better understanding of neural mechanism have better facilitated brain scientists to perform studies with running stream of continuous speech and expand the scope of EEG studies to include comprehension of more realistic speech envelope. The purpose of this study is to review literatures on neural tracking to speech envelope to discuss it in Audiology perspective. This review article consists of seven subjects including introduction, neural tracking theories, neural tracking measure, signal processing & analysis, literature review on neural tracking associated with hearing loss, application of neural tracking to audiology, and conclusion. We noted that neural tracking has potential to be used in clinical sets to objectively evaluate speech comprehension for people with hearing loss in the future.

Key Words: Neural tracking, Speech comprehension, Electroencephalography, Audiology.

INTRODUCTION

난청의 정도 및 유형을 객관적으로 평가하기 위한 전기생리학 적 방법으로 청성유발전위(auditory evoked potential, AEP)가 임상에서 널리 사용되고 있다. 특히 청성뇌간반응(auditory brainstem response)과 청성지속반응(auditory-steady-state response, ASSR)은 순음청력역치(pure-tone average)와 상관관계가 높고 검사 신뢰도와 민감도가 높아서 신생아 및 유소아를 비롯한 의사소통과 협조가 어려운 환자의 역치 추정 에 있어서 유용하게 사용되고 있다(Vasileiou et al., 2018). 하지만 이러한 신경계의 말초 영역에서 측정되는 검사들은 소리의 인지(cognition)가 아닌 탐지(detection) 영역에 초점을 맞추는 검사로 청능재활(aural rehabilitation)의 궁극적 목표인 어

음이해능력(speech intelligibility)을 반영하지 못한다는 단점 이 있다. 언어적 레벨과 관련이 높은 AEP로 N400과 P600반응 이 있다. N400반응은 음 제시 후 약 400 ms에 나타나는 내인 적(endogenous) 음전위반응으로 의미론적 부조화(semantic incongruity)와 관련되어 있으며(Federmeier & Kutas, 1999), P600반응은 음 제시 후 약 500 ms에 나타나는 내인 적 양전위반응으로 구문론적 부조화(syntactic incongruity)와 관련이 있다(Kaan et al., 2000). 이와 같은 후기반응(late latency response)은 언어와 관련된 전위반응(linguistic potential)으로 인간의 언어능력과 관련성이 있음이 입증되었지만 반응의 반복성, 검사 신뢰도, 자극음의 제한성 등의 한계로 임상적으로 사용되지 못하고 있는 실정이다.

대뇌피질 뉴런의 활동은 청각각을 통해 전달되는 말소리의 느린 시간적 envelope을 따르는 특징을 갖는다. 즉, 대뇌 피질 의 신경활동은 지속적으로 시간 축에서 변화되는 언어 자극의 envelope 정보에 약 180 msec의 잠복기 후 역동적으로 반응

한다(Aiken & Picton, 2008). 이러한 현상을 설명하는 과학적 용어로 'cortical entrainment', 'neural tracking', 'neural synchrony', 'neural entrainment', 'envelope following response', 'cortical synchronization' 등이 혼용해서 사용되고 있다. 본 연구에서는 '신경추적(neural tracking)'이라는 용어로 통일하여 설명하고자 한다. 신경추적 현상은 이미 1990년대 초에 보고되었던 사실이지만 뇌 과학과 신호처리

(signal processing) 기술의 발달과 더불어 최근에는 기계학습(machine learning) 또는 신경망(neural network) 기술의 적용까지 더해져 보다 진보된 뇌파 측정과 분석으로 인해 오늘날 많은 신경과학자들에게 큰 관심을 받고 있다. 신경추적이 가장 활발하게 연구되고 있는 주제는 소음속에서 공간청취(spatial hearing)이다. 인간은 대화자 잡음의 복잡한 상황 속에서 원하는 소리(신호)에 선택적으로 집중하여 인지할 수 있는 청각 처리

Figure 1. Three samples of speech envelope (blue) overlapping with EEG response (red) that represent high (A), moderate (B), and low (C) cross-correlations (Lee et al., 2020). The EEG activities were obtained from CI subject 1 in response to the speech processed with 1 sinusoidal component and the low-pass filter at 1 kHz. The lag indicates time lags between stimulus onsets and moments that lead the maximum correlations by cross-correlation analysis. Permission is obtained to use this figure from authors. EEG: electroencephalography.

능력을 가진다. 이러한 선택적 듣기능력(selective attention)으로 알려진 '칵테일 파티 효과(cocktail party effect)'에 대한 정확한 배경은 아직 명확히 밝혀지지 않은 영역으로 남아있다. 신경추적에 대한 흥미로운 점 하나는 주변 잡음과 어음신호가 동시에 제시된 상황에서 집중하지 않은 잡음보다 선택적으로 집중한 어음신호에 더욱 일치하는 신경추적을 보인다는 것이다. 이와 관련된 많은 실험적 연구들이 뇌파로부터 제시 어음을 재구성(reconstruction)하였을 때 집중한 어음에서 더 높은 신경추적 수치가 나타남을 보고하고 있다(Ding & Simon 2012a; Fuglsang et al., 2020; O'Sullivan et al., 2015; Power et al., 2012; Vanthornhout et al., 2018; Zion Golumbic et al., 2013). Figure 1 (Lee et al., 2020)은 인공와우 환자로부터 얻은 신경추적반응의 예로 순음모델링과 저주파수 여파된 짧은 문장의 어음 envelope(파랑)와 이에 대한 뇌파반응(빨강)을 나타낸다. 어음신호와 뇌파신호의 상관관계(c)의 정도에 따라 분류(Fig. 1: A, 높음; B, 보통; C, 낮음)하였고, 이는 신경추적의 정도를 의미한다고 할 수 있다. Lag는 교차상관관계분석(cross-correlation)에서 가장 높은 상관관계가 나타난 시간지연(time lag)을 표시하고 있다. 신경추적의 정도는 다양한 변수에 의해서 영향을 받으며 같은 외부 조건이라도 다음과 같이 피검자사의 내적 요인에 따라서 차이가 나타날 수 있다.

이 분야 이외에도 신경추적은 오늘날 정상 뇌의 청각 처리 영역에 머물지 않고 인지를 비롯한 상위 처리능력으로 적용성을 확장하고 있다. Power et al.(2016)은 난독증 아동을 대상으로 한 연구에서 난독증 아동의 0~2 Hz의 신경추적이 정상 아동보다 떨어짐을 보고하였으며, Di Liberto et al.(2018)은 이러한 비정상적인 신경추적은 대뇌 우반구의 음성 특징 추적능력의 결핍으로 보고있다.

시간 축에서 진폭이 변화하는 느린 저주파수 어음성분인 envelope이 어음 인지와 아주 큰 관련성이 있음(Shannon et al., 1995)을 고려할 때 시간적 envelope을 따르는 신경추적 현상은 어음 인지에 대한 대뇌 신경활동의 메커니즘을 규명하기 위한 기초과학적인 관점뿐만 아니라 난청인의 청각 및 어음 자극 인지와 관련된 청능재활의 관점에서도 그 의미가 크다고 하겠다. 본 연구의 목적은 최근 청각자극에 대한 뇌파분석 패러다임으로 많이 연구되고 있는 신경추적법에 대해 소개하고 관련 연구 문헌을 고찰하여 이를 근거로 청각학적 관점에서 고찰해 보고자 한다.

MATERIALS AND METHODS

신경추적 해석 이론

그렇다면 신경추적 현상이 말소리 이해에 있어서 의미하는 바는 무엇인가? 인간의 대뇌피질에서 나타나는 저주파수 신경

추적과 언어 이해 및 처리에 대한 메커니즘과의 관계에 대해서는 아직 뚜렷하게 정립되어 있지 않다. 비언어(non-speech) 자극음으로 신경추적을 관찰한 연구(Lalor et al., 2009) 또는 왜곡된 언어를 자극음으로 신경추적을 관찰한 연구(Howard & Poeppel, 2010; Nourski et al., 2009)를 고려하면 신경추적이 단지 음향 자극의 물리적 변화를 반영하여 나타나는 상향식(bottom-up) 정보 처리 현상임을 암시하고 있으나 이와 반대로 신경추적이 중추적 언어 레벨에서 발생하는 주의집중(attention) 및 인지와 관련된 하향식(top-down)기능과 관계 깊음을 보여주는 연구(Ding & Simon, 2012a; Petersen et al., 2017; Vanthornhout et al., 2019)도 많이 보고되었다. Ding & Simon(2014)이 정리한 네 가지 신경추적에 대한 주요 가설을 간략히 설명하면 다음과 같다.

시작추적이론(onset tracking hypothesis) (Howard & Poeppel, 2010)이 제시하는 신경추적의 의미는 신경추적은 역동적으로 변화하는 음향신호의 큰 변화 지점(edge)의 시작(onset) 또는 끝(offset)을 반영하여 나타난다는 것이다. 말소리의 진폭 변화는 음절 단위로 이뤄지며 이 변화 지점에 대한 신경반응이 시간 축에서 축적되어 연속적으로 나타나는 것이 신경추적으로 기록될 수 있다. 이는 결국 사건유발전위(event-related potential, ERP)의 합으로 표현된다고 할 수 있는 것으로 언어 특수성을 보이는 반응이 아닌 물리적 음향 변화에 따른 뇌파의 변화로 설명하는 이론이다. AEP의 후기반응 중 하나인 P1-N1-P2 (compound auditory evoked potential)가 변화하는 소리반응에 나타나는 것으로 알려진 음향변화전위(acoustic change complex)가 이어지는 말소리 변화 시 마다 연속적으로 일어나는 반응으로 설명될 수 있겠다.

수집된 특징추적이론(collective feature tracking hypothesis) (Ding & Simon, 2012b)은 제시된 어음이 우리 뇌의 청각시스템에 들어가면 다양한 음향적 특징, 예를 들면 음조(pitch), 음원의 위치 정보(source of location), 주파수/시간 변화(spectral/temporal change) 등을 해독하고 주/보조 청각 영역(primary/associative auditory area)에서의 종합적 신경반응의 동시적 반응이 이러한 주파수/시간 정보의 특징을 반영한다는 것이다. 따라서 시작추적이론과 수집된 특징추적이론은 말소리 이해에 있어서 대뇌피질의 역할이 어음신호를 분석하고 해독(decoding)하는 현상으로 정리하였다.

또 다른 신경추적을 해석한 시도로써 음절분석이론(syllabic parsing hypothesis)이 있다. 이 이론은 신경추적현상이 단지 어음의 음향적 특징을 수동적으로 따르는 것이 아니라 어음의 음절 단위로 단절시켜 분석하는 데 중요한 역할을 한다는 것이다(Giraud & Poeppel, 2012). 이는 보통 말소리에 있어서 음절의 속도(rate)가 약 5~8 Hz 주기로 나타나고 음운 정보를 분석

하는데 중요한 역할을 하는 대뇌의 theta파의 주기 역시 약 4~8 Hz로 비슷한 주파수를 갖는다는 이유로 설명된다. 이 주기는 어음 envelope의 인접한 꼭지점(peak)이 나타나는 주기와도 같다. 즉, 음절분석이론은 신경추적이 단순하며 수동적인 음향정보를 따르는 뇌파 현상이 아닌 언어 이해를 위해 어음을 음절로 단위화하는 작업으로 설명한다.

마지막으로 감각선택이론(sensory selection hypothesis)은 신경추적이 배경잡음으로부터 어음신호를 구분하는 역할을 한다는 것이다. 변화하는 어음의 리듬성 음향 특성의 시간적 동시성은 어음의 envelope에 투영되며 뇌가 잡음으로부터 신호인 어음을 분별하도록 도움을 준다는 것이다. 이 이론은 배경잡음을 사용한 복잡한 듣기 환경이 주어진 연구에서 어음신호에 대한 신경추적이 배경잡음에 대한 신경추적보다 강하게 나타나는 연구들이 이를 뒷받침한다(Ding & Simon, 2012a; Ding & Simon, 2013). 감각선택이론은 또 다시 우리의 뇌에서의 신경추적이 어음신호의 음향 특징을 잡아내기 위한 시간적 일치성을 분석하는 과정이라는 것(Ding & Simon, 2012a)과 잡음 속에서 어음신호와 관련된 음향 특성을 갖는 시간을 선택적으로 구분하도록 도와주는 역할을 한다는 것으로 나뉜다(Schroeder & Lakatos, 2009).

지금까지 간략하게 정리한 네 가지 이론 모두 어음 인지모델과 연관된 신경활동의 특성을 타당성 있게 설명하고 있다. 즉, 어떠한 어음 인지모델을 기준으로 설명하는 것이냐에 따라 해석이 달라질 수 있으나, 어음의 특정 리듬에 동시성(synchrony)을 보이는 신경활동은 말소리를 분석하고 이해하는 데 큰 기여를 하는데 공통적인 견해를 가지는 것으로 보인다. 향후 연구에서는 신경추적의 정확한 기전(neural generator)과 말소리 이해와의 명확한 관계에 대한 연구가 더욱 필요하다고 하겠다.

신경추적 측정

신경추적의 측정은 어음 에너지의 빠른 변화에 따른 신경 반응을 측정해야 하기 때문에 대개 높은 시간 분해력(temporal resolution)을 갖고 있는 다채널 뇌전도(electroencephalography, EEG) 또는 뇌자도(magnetoencephalography)시스템을 사용하여 측정된다. 본 연구에서는 측정이 비교적 간편하고 비용이 적게 발생하여 많이 사용되고 있는 EEG 측정과 관련된 내용들을 주로 다루고자 한다. EEG 측정에서는 보통 64채널 이상의 두피전극(scalp electrode)을 international 10~20 전극시스템에 따라 부착하여 측정한다. 이러한 다채널시스템은 각 전극이 위치한 뇌 영역에 대한 국소적 신경반응을 확인하여 뇌 위치에 따른 신경반응 분석이 가능하게 할 뿐만 아니라 독립성분분석(independent component analysis)을 통한 잡파(artifact) 제거가 용이하고, 신경반응과 어음신호를 해독기

(decoder)를 통한 기계학습(machine learning) 또는 교차 상관관계분석(cross correlation)을 이용하여 분석할 때 많은 데이터를 활용하여 제시 어음의 envelope과 유발되는 뇌파의 envelope과의 상관성을 정립하는 데 유리한 측면이 있다.

신경추적의 가장 큰 장점 중 하나는 기존에 전통적으로 사용되어왔던 짧은 음이 아닌 실제 일상생활에서 사용되는 말소리와 비슷한 긴 어음을 사용하여 언어 이해능력을 신경생리학적으로 확인하는 데 가장 근접한 측정이라는 것이다. 피검사자에게 측정에 대한 설명이 끝나면 피검사자를 조용한 사운드부스(sound booth)에 위치시키고 어음신호가 이어폰 또는 스피커를 통하여 제시된다. 신호 어음으로는 하나의 문장을 수백 번에서 수천 번 반복하여 제시하거나(Lee et al., 2020), 더욱 실제 대화와 가깝도록 환경을 만들기 위해 짧은 이야기 또는 대화를 제시한다(Di Liberto et al., 2015; Kong et al., 2015; O'Sullivan et al., 2015; Vanheusden et al., 2020). 측정 전 보편적으로 시행하는 EEG 측정과 마찬가지로 측정 환경 주변의 전기적/음향적 잡파와 소음을 최대한 통제한다. 모든 전극의 저항(impedance) 값이 낮게(10 kΩ 이하) 그리고 균일하게 유지될 수 있도록 전극 부착에 주의가 필요하다. EEG 측정 시 신호대잡음비(signal-to-noise ratio, SNR)를 높이기 위해서는 피검사자의 협조가 필요하다. 검사가 진행되는 동안 피검사자에게 눈 깜박임과 근육 움직임을 최소화하도록 요청하고 측정이 진행되는 동안 검사자는 피검사자의 상태를 지속적으로 모니터링하여 피검사자가 비협조적이거나 집중력이 떨어졌을 경우 적절한 동기 부여 또는 휴식을 제공해야 한다.

신경추적의 연구는 궁극적으로 행동학적인 어음 이해능력과 신경생리학적인 신경추적 결과 값의 상관관계를 계산하여 청각적 어음 인지의 메커니즘을 설명하거나 자극음에 따른 신경반응 특성을 확인하고자 한다. 따라서 객관적이고 전기생리학적 신경추적뿐만 아니라 주관적이고 행동학적인 수행평가를 함께 측정한다. 행동학적 어음 인지에 대한 측정을 위해서는 단어인지 검사, 문장인지검사, 음질평가, 이야기를 듣고 물음에 답하기 등의 평가 방법이 주로 사용된다. 검사 시 연구 설계에 따라 행동학적 어음 인지 측정과 EEG를 통한 전기생리학적 신경추적이 동시에 진행되기도 하며(Decruy et al., 2019) 두 검사를 따로 진행하기도 한다(Anderson et al. 2011; Ding & Simon, 2013; Vanthornhout et al., 2018). 따로 측정될 경우 피검사자는 어음이 제시되는 동안 어음신호에 집중을 요구받는 능동적 패러다임(active paradigm)과 음소거가 설정된 비디오를 보거나 책을 읽도록 하여 어음에 집중이 필요 없도록 디자인된 수동적 패러다임(passive paradigm)으로 나뉠 수 있다. 신호 어음에 집중을 요구하는 능동적 패러다임의 경우 피검사자가 제시음에 집중을 하도록 하기 위한 방법으로 어음 제시가 끝난 후 제시된 내용

과 관련된 간단한 퀴즈를 풀도록 하는 방식이 많이 사용된다.

신호처리 및 분석

어음 제시 시간 동안 측정된 뇌파는 컴퓨터에 저장되고 필요 시 전처리(pre-processing) 과정으로 기준전극을 다시 설정하기 위한 referencing, 데이터 용량을 축소시키기 위한 다운샘플링(down-sampling), 파형을 표준화하기 위한 baseline correction 등을 시행한다. 이와 함께 SNR을 높이기 위하여 근육반응(muscle activity), 눈 깜박임반응(electro-ocular artifact), 심장 박동(heart beat), 저주파 전기적 잡파 등을 수동적으로 또는 다양한 제거 알고리즘 등을 사용하여 제거한다. 전처리된 EEG반응은 약 1~40 Hz 대역통과필터(band-pass filter)를 사용하여 envelope 신호만을 추출한다. 어음에 대한 신경추적의 비교분석을 위해서는 EEG 신경반응뿐만 아니라 어음신호의 envelope 정보도 신호처리 과정을 통하여 추출되어야 하는데 이는 주로 힐버트 트랜스폼(Hilbert transform)을 이용해 이뤄진다. 신호처리 과정에 있어서 주의해야 할 점은 EEG와 어음신호에서 처리되는 저주파수대역(약 40 Hz)과 다운샘플링(주로 128 Hz) 값이 동일하도록 처리가 이루어져야 두 가지 추출한 envelope변수의 벡터(vector)값의 비교분석이 가능하다는 것이다. 추출된 어음의 envelope과 신경반응의 envelope에 대한 상관관계분석은 과거에는 시도내위상일

치(inter-trial phase coherence)분석을 사용하거나 교차상관관계분석을 이용하여 어음의 envelope과 EEG의 envelope를 시간 축에서 이동시켜 가장 큰 상관관계를 갖는 값을 찾는 방식을 주로 이용하였으나, 최근에는 기계학습과 신경네트워크 기술의 발달에 따라 temporal response function (TRF) (Ding & Simon, 2012a)을 이용한 모델링(modeling) 기법이 더 많이 사용되고 있는 추세이다. TRF는 인간의 뇌를 선형불변시스템(linear time-invariant system)으로 가정하고 입력 신호로부터 특정 시간에 반응하는 뇌파반응값을 이용한다. TRF는 포워드 모델(forward model)과 백워드 모델(backward model)로 구분될 수 있는데 포워드 모델은 자극음의 특징을 이용하여 EEG를 예측하며, 백워드 모델은 이와 반대로 EEG 데이터로부터 자극음을 재구성한다. 포워드 모델은 자극 어음을 단일변수(univariate) 입력값으로 취급하고 각 채널의 반응을 모델링하기 때문에 다변수(multivariate)반응으로 볼 수 있는 뇌파반응을 고려해야 할 경우에는 EEG로부터 어음을 재구성하는 방식인 백워드 모델을 더 많이 사용한다. 백워드 모델의 적용을 위해서는 어음 envelope을 재구성하기 위한 해독기를 훈련(training)시키는 과정이 필요하기 때문에 이를 위해서 EEG 기록세션을 따로 구성하여 데이터를 수집하거나, 실험세션 내에 획득한 데이터를 분할하여서 최적의 값을 찾도록 설계하기도 한다. 신경추적을 위한 해독기로 Crosse et al.(2016)이 개발한 Matlab toolbox

Figure 2. Flowchart of summarized neural tracking measure and analysis. EEG: electroencephalography.

기반의 mTRF가 주로 사용된다. 어음신호와 EEG에서 발생하는 시간 차이(time lag), 채널의 수 등의 변수를 고려하여 가공된 어음신호의 envelope과 EEG반응으로부터 재구성된 어음신호 envelope 사이에 최소 제곱근 오차(least squares error)값이 최소화(minimized)되도록 설계되고 최적 모델을 갖도록 각 채널에 가중치를 부여하여 신경추적이 어음의 envelope를 따르는 정도를 정량화한 결과값을 도출해낸다. 자세한 수학적 및 알고리즘에 대한 설명은 본 연구에서는 생략한다. 도출한 신경추적값은 제시된 음에 대한 신경생리학적 지표로서 대표성을 가지며 이와 함께 또는 따로 측정된 행동학적 결과값(e.g., 어음인지점수)과 비교분석하여 객관적/생리적 결과와 주관적/행동학적 결과의 상관관계를 분석한다. 많은 연구(Akbarzadeh et al., 2020; Kong et al., 2015; Lee et al., 2020; Vanthornhout et al., 2018)에서 신경추적과 행동학적 어음 인지도 사이에 높은 상관관계가 있음이 보고되고 있다. Figure 2는 신경추적 연구의 측정 및 분석 과정을 대략적으로 요약한 블록 다이어그램(block diagram)으로 실질적으로는 연구 목표 및 설계에 따라 차이가 있을 수 있다.

청력손실과 신경추적에 관한 연구

지금까지 소수의 연구들이 청력손실을 가진 난청인을 대상으로 신경추적 실험을 진행하였다. Petersen et al.(2017)은 27명의 60세 이상 미도에서 중고도 난청인을 대상으로 보청기 착용 후 신경추적 패턴을 분석하고자 하였다. 짧은 이야기가 조용한 상황, 다른 경쟁이 되는 이야기가 주어진 소음 상황에서 스피커를 통해 제시되었으며 피검자가 보청기를 통해서 제시 이야기를 듣는 동안 EEG를 측정하였다. EEG를 통해 측정된 뇌파값은 제시 어음의 envelope과 교차상관계분석을 통해서 신경추적 정도를 도출하였다. 연구 결과에서 집중하지 않도록 요구받은 경쟁 어음 잡음에 비해 집중하도록 요구받은 어음신호에서 더 큰 신경추적이 관찰되었으며 청력손실이 클수록 집중하지 않은 어음에 대한 신경추적이 떨어진다고 보고하였다. 잡음의 크기가 작을수록 신경추적이 높아지는 정도가 청력손실이 클수록 낮았는데 이는 청력손실이 중추의 집중과 관련된 메커니즘에 영향을 주게 되어 집중한 소리와 집중하지 않은 소리의 신경추적에 큰 차이가 없음을 의미한다고 해석하였다. 이 연구는 보청기로 증폭된 소리라도 어음에 대한 신뢰도 있는 신경추적이 가능함을 보여주는 첫 번째 연구였으며 보청기로 인한 소리 변화효과는 신경추적에 미세한 영향을 주는 것으로 해석되었다.

하지만 최근에는 Petersen et al.(2017)의 연구 결과와 반대로 정상인보다 노인 난청인에게서 집중한 소리에 대하여 더 큰 신경추적반응을 얻은 연구(Fuglsang et al., 2020)가 보고되었다. 이러한 두 연구 결과에 대한 불일치를 설명하고자 Fuglsang et

al.(2020)은 연구 설계의 차이를 이유로 제시하였다. Peterson et al.(2017)의 연구에서는 공간적으로 같은 위치에서 특정 어음 인지역치(speech reception threshold)를 기준으로 제시하여 말소리 이해에 대한 난이도가 더 어려웠고, 비슷한 어음 이해력을 맞추기 위하여 난청 정도에 따라 소음에 대한 신호 어음의 크기도 함께 증가시켰기 때문에 집중에 대한 변수만을 적절히 고립시켜 측정할 수 없었다는 것이다. Fuglsang et al.(2020)은 난청 환자 그룹의 선택적 듣기에 대한 신경추적 결과가 높게 나타났음에도 불구하고 행동학적 경쟁 소음 하에 어음인지점사에서 큰 어려움을 보였다는 것은 노인성 난청인이 갖는 소음 하 어음 인지능력 저하의 문제는 선택적 듣기능력만으로 설명할 수 없음을 의미한다고 보고하였다.

이밖에 Vanheusden et al.(2020)은 보청기의 증폭으로 인한 음향학적 변화가 가져올 수 있는 영향에 대하여 연구하고자 경중도 난청인 17명을 대상으로 신경추적을 시행하였다. 보청기에서 이뤄지는 신호처리는 시간적 envelope 정보를 왜곡시킬 수 있으나 보청기 착용 전/후 신경추적 결과에는 큰 변화가 없는 것으로 보고되었다. 이는 진보된 해독기 모델 기술로 인해 보청기로 인한 envelope의 왜곡에도 신뢰성 있는 측정이 가능함을 의미하며 향후 보청기 기술에 다양한 적용이 가능함을 시사한다.

사실 난청과 관련된 신경추적의 연구는 보청기보다는 인공와우(cochlear implant)와 관련되어 더 많이 보고되고 있다. 인공와우가 단순히 소리를 증폭하는 것이 아닌 말초 청신경을 직접 자극하여 소리를 대뇌피질까지 전달한다는 점에서 신경추적을 연구하는 신경과학 및 뇌과학자들 사이에서 더 흥미로울 수 있다. 특히 인공와우가 더욱 주목을 받는 이유는 인공와우와 신경추적법 모두 소리의 envelope 정보를 주요 소스(source)로 이용하고 있다는 것이다. 인공와우는 필터뱅크(filter-bank)를 이용해 말소리의 주파수별 envelope 정보를 추출하여 전기적 이중상 펄스(electrical bipolar pulse) 자극을 말초 청신경에 전달한다. 얼마나 미세한 envelope의 변화까지 전달하느냐는 자극속도(stimulation rate)를 얼마나 빠르게 하느냐에 의해서 결정된다. 따라서 어음 envelope을 추적하는 신경추적은 어음 envelope을 주된 정보로 제공하는 인공와우 시스템을 평가하는 데 적합한 방법으로 사료될 수 있다. 인공와우 환자에게 신경추적을 적용하기 위한 제한점으로써 과거부터 지적되어 왔던 문제 중 하나는 인공와우 어음처리기와 신경자극에서 발생하는 전기적 잡파(electrical artifact)가 순수한 뇌파신호를 얻는 것을 방해한다는 것이다. 하지만 오늘날 이를 극복할 수 있는 기술들이 개발되면서 인공와우 환자로부터 신경추적을 성공적으로 측정하고 있다. 소개된 인공와우 잡파 제거 방법으로는 먼저 잡파를 측정한 후 뇌파신호에서 빼주는 방법(Deprez et al., 2017; Friesen & Picton, 2010), 주기성을 띄며 발생하는 전기 잡파를

제거한 뇌파신호를 얻기 위해 제시음에 짧은 갭(gap)을 넣는 방법(Somers et al., 2019), 주성분분석(principal component analysis) (Martin, 2007), 공분산(covariance)을 사용하는 독립성분분석(Martin, 2007) 등이 있다. 인공와우 환자들을 대상으로 진행된 연구(Akbarzadeh et al., 2020; Lee et al., 2020; Nogueira et al., 2019; Verschuere et al., 2019)에서도 마찬가지로 신경추적이 어음 인지능력과 높은 상관관계가 있음을 보고하였다.

DISCUSSIONS

신경추적의 청각학적 활용

청각학(Audiology) 영역에서 신경추적의 임상적 활용성은 높은 잠재력을 갖고 있다고 할 수 있다. 과거 고전적인 AEP 측정 및 분석법은 짧은 어음신호에 단속적으로 일어나는 ERP분석을 활용하는 패러다임에 초점이 맞춰져 있었다(Lee & Bidelman, 2017; Tremblay, et al., 2006). 이러한 접근법은 어음 변화에 대한 탐지 및 인지를 반영하는 중요한 방법이지만 짧고 작은 단위의 어음 변화에 대한 신경반응의 변화를 설명할 뿐 실질적으로 복잡하고 지속적인 특성을 갖는 일상생활의 말소리에 대한 종합적 이해능력을 측정하기에는 한계가 있다. 100 Hz 이하의 주파수변조음을 사용하는 ASSR은 현재 임상적으로 사용되고 있는 청성유발전위 중 신경추적법과 유사점이 가장 많은 측정법이라고 할 수 있다. 특히 10~40 Hz 이하의 느린 지속반응은 자극음의 반복성과 변조 주파수에 일치하여 연속적으로 일어나는 대뇌피질 뉴런의 동시성(synchrony)을 측정하였다는 점에서 신경추적과 공통성을 갖고 있다고 할 수 있다. 하지만 ASSR은 측정 및 분석 방법에서 신경추적과 차이를 보이고 특히 임상적 역할이 청각 역치 추정을 위한 단서를 제공하는 것에 국한되어 있다. 많은 학자들이 언급하고 있는 것과 같이 신경추적은 어음 이해력을 객관적으로 측정할 수 있는 임상도구으로써 충분한 가능성을 갖고 있다고 볼 수 있다(Decruy et al., 2019; Somer et al., 2019; Vanthornhout et al., 2018).

향후 지속적인 신경추적 기술의 혁신은 청력보장구의 적합 및 평가에도 기여가 가능할 것이다. 보청기 적합 공식(hearing aid fitting formula) 또는 어음향상(speech enhancing) 알고리즘의 개발은 의사소통이 어려운 환자에게 객관적이고 자동적인 방법으로 최적 적합을 제공할 것이고, 임상에서 보청기 착용 후 신경추적 측정을 통한 어음이해도의 예측은 보청기 평가와 실이 측정(real ear measurement)에서 난청 환자의 최적 적합을 제공하는 데 도움을 줄 것이다.

특히 앞서 지적한 바와 같이 인공와우와 신경추적은 어음 envelope를 스스로 이용하는 공통점을 갖고 있기에 활용 가

능성이 더 클 수 있다. 인공와우의 또 다른 장점은 와우내전극(intra-cochlear electrode)이 청신경을 자극하는 기능뿐만 아니라 신경전기반응을 측정하는 기능도 가질 수 있다는 것이다. 이를 활용하여 실제 임상에서는 전극의 저항(impedance)과 청신경의 전기복합활동전위(electrical compound action potential, ECAP)를 측정하여 인공와우 시스템을 통한 각 전극으로의 전기신호의 전달 성공 유무를 판단하고 매핑(mapping)에 활용한다. 향후 신경추적 기술이 발전하여 EEG 시스템이 인공와우에 탑재가 되고 와우내전극을 통한 기록이 가능해 진다면 인공와우 환자의 수술 중 또는 매핑 시 어음이해력을 측정하는 혁신적이고 효율적인 모듈로서 작용할 수 있을 것이다.

CONCLUSION

신경추적은 우리의 대뇌피질의 신경 활동이 저주파수 어음의 envelope에 따라 수반되는 현상으로 어음이해도와 높은 상관관계를 보인다. 현재까지는 신경추적과 청력손실에 관한 연구가 아직 걸음마 단계에 있으며, 측정 과정이 복잡하고 검사기기의 가격이 높으며 분석 기법이 일반 임상가가 사용하기에는 어려운 한계점이 있기 때문에 임상에서 사용하기에 부적절한 단계에 있다. 하지만 향후 기술적 진보로 신경추적의 상용화를 위한 난제들(잡파와 소음, 많은 필요 전극의 수, 측정기기의 복잡성, 긴 측정시간, 신호처리 및 분석의 어려움)이 해결되고, 사용자 친화적인 상용화된 장비와 함께 가이드라인이 보급된다면 청각장애인의 어음이해도를 평가할 수 있는 임상적 지표로써, 청력보조기기의 최적 적합 및 알고리즘의 개발을 위한 객관적 측정 모듈로써 임상청각학에서의 활용을 기대해 볼 수 있을 것이다.

중심 단어 : 신경추적, 어음 이해, 뇌전도, 청각학.

Ethical Statement

N/A

Acknowledgments

N/A

Declaration of Conflicting Interests

There are no conflict of interests.

Funding

This work has supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021S1A5A806455811).

ORCID iD

Sungmin Lee <https://orcid.org/0000-0002-3848-7833>

REFERENCES

- Aiken, S. J. & Picton, T. W. (2008). Human cortical responses to the speech envelope. *Ear and Hearing*, 29(2), 139-157.
- Akbarzadeh, S., Lee, S., & Tan, C. T. (2020). The implication of sound level on spatial selective auditory attention for cochlear implant users: Behavioral and electrophysiological measurement. *Interspeech*, 2020, 2497-2501.
- Anderson, S., Parbery-Clark, A., Yi, H. G., & Kraus, N. (2011). A neural basis of speech-in-noise perception in older adults. *Ear and Hearing*, 32(6), 750-757.
- Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. *Frontiers in Human Neuroscience*, 10, 604.
- Decruy, L., Vanthornhout, J., & Francart, T. (2019). Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. *Journal of Neurophysiology*, 122(2), 601-615.
- Deprez, H., Gransier, R., Hofmann, M., van Wieringen, A., Wouters, J., & Moonen, M. (2017). Template subtraction to remove CI stimulation artifacts in auditory steady-state responses in CI subjects. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(8), 1322-1331.
- Di Liberto, G. M., O'Sullivan, J. A., & Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Current Biology*, 25(19), 2457-2465.
- Di Liberto, G. M., Peter, V., Kalashnikova, M., Goswami, U., Burnham, D., & Lalor, E. C. (2018). Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. *Neuroimage*, 175, 70-79.
- Ding, N. & Simon, J. Z. (2012a). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America*, 109(29), 11854-11859.
- Ding, N. & Simon, J. Z. (2012b). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*, 107(1), 78-89.
- Ding, N. & Simon, J. Z. (2013). Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *Journal of Neuroscience*, 33(13), 5728-5735.
- Ding, N. & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Neuroscience*, 8, 311.
- Federmeier, K. D. & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41(4), 469-495.
- Friesen, L. M. & Picton, T. W. (2010). A method for removing cochlear implant artifact. *Hearing Research*, 259(1-2), 95-106.
- Fuglsang, S. A., Märcher-Rørsted, J., Dau, T., & Hjortkjær, J. (2020). Effects of sensorineural hearing loss on cortical synchronization to competing speech during selective attention. *Journal of Neuroscience*, 40(12), 2562-2572.
- Giraud, A. L. & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511-517.
- Howard, M. F. & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology*, 104(5), 2500-2511.
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15(2), 159-201.
- Kong, Y. Y., Somarowthu, A., & Ding, N. (2015). Effects of spectral degradation on attentional modulation of cortical auditory responses to continuous speech. *Journal of the Association for Research in Otolaryngology*, 16(6), 783-796.
- Lalor, E. C., Power, A. J., Reilly, R. B., & Foxe, J. J. (2009). Resolving precise temporal processing properties of the auditory system using continuous stimuli. *Journal of Neurophysiology*, 102(1), 349-359.
- Lee, S., Akbarzadeh, S., & Tan, C. T. (2020). Neural tracking of band-limited sine-wave speech in normal hearing and cochlear implant listeners. *International Journal of Speech and Language Pathology and Audiology*, 8, 38-49.
- Lee, S. & Bidelman, G. M. (2017). Objective identification of simulated cochlear implant settings in normal-hearing listeners via auditory cortical evoked potentials. *Ear and Hearing*, 38(4), e215-e226.
- Martin, B. A. (2007). Can the acoustic change complex be recorded in an individual with a cochlear implant? Separating neural responses from cochlear implant artifact. *Journal of the American Academy of Audiology*, 18(2), 126-140.
- Nogueira, W., Dolhopiatenko, H., Schierholz, I., Büchner, A., Mirkovic, B., Bleichner, M. G., et al. (2019). Decoding selective attention in normal hearing listeners and bilateral cochlear implant users with concealed ear EEG. *Frontiers in Neuroscience*, 13, 720.
- Nourski, K. V., Reale, R. A., Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., et al. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *Journal of Neuroscience*, 29(49), 15564-15574.
- O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., et al. (2015). Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cerebral Cortex*, 25(7), 1697-1706.
- Petersen, E. B., Wöstmann, M., Obleser, J., & Lunner, T. (2017). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *Journal of Neurophysiology*, 117(1), 18-27.
- Power, A. J., Colling, L. J., Mead, N., Barnes, L., & Goswami, U. (2016). Neural encoding of the speech envelope by children with developmental dyslexia. *Brain and Language*, 160, 1-10.
- Power, A. J., Foxe, J. J., Forde, E. J., Reilly, R. B., & Lalor, E. C. (2012). At what time is the cocktail party? A late locus of selective attention to natural speech. *European Journal of Neuroscience*, 35(9), 1497-1503.
- Schroeder, C. E. & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1), 9-18.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303-304.
- Somers, B., Verschueren, E., & Francart, T. (2019). Neural tracking of the speech envelope in cochlear implant users. *Journal of Neural Engineering*, 16(1), 016003.
- Tremblay, K. L., Billings, C. J., Friesen, L. M., & Souza, P. E. (2006). Neural representation of amplified speech sounds. *Ear and Hearing*,

- 27(2), 93-103.
- Vanheusden, F. J., Kegler, M., Ireland, K., Georga, C., Simpson, D. M., Reichenbach, T., et al. (2020). Hearing aids do not alter cortical entrainment to speech at audible levels in mild-to-moderately hearing-impaired subjects. *Frontiers in Human Neuroscience*, 14, 109.
- Vanthornhout, J., Decruy, L., & Francart, T. (2019). Effect of task and attention on neural tracking of speech. *Frontiers in Neuroscience*, 13, 977.
- Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z., & Francart, T. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. *Journal of the Association for Research in Otolaryngology*, 19(2), 181-191.
- Vasileiou, A., Pavlos, M., Antonios, T., Ioannis, X., Vlastarakos, P. V., & Nikolopoulos, T. P. (2018). Correlation of ASSR hearing thresholds with ABR hearing thresholds in children. *Journal of Hearing Science*, 8(3), 35-39.
- Verschueren, E., Somers, B., & Francart, T. (2019). Neural envelope tracking as a measure of speech understanding in cochlear implant users. *Hearing Research*, 373, 23-31.
- Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron*, 77(5), 980-991.