

인공와우 전기 청각 인지에 대한 정현파 모델 적용에 관한 연구

Application of sinusoidal model to perception of electrical hearing in cochlear implants

이성민[†]

(Sungmin Lee^{1†})

¹동명대학교 언어치료청각학과

(Received November 9, 2021; accepted December 3, 2021)

초 록: 어음은 복잡한 정현파들의 합으로 구성된다. 본 논문은 정현파 모델을 인공와우 시뮬레이션에 적용하여 인공와우 사용자의 전기 청각 인지 메커니즘을 정상 청력인의 음향 청각 인지 메커니즘과 비교하였다. 이를 위해 정상 청력을 지닌 14명의 성인들이 본 연구에 참여하였다. 2, 4, 6, 8개의 정현파 성분만을 추출하는 정현파 모델이 적용된 문장과 정현파 모델과 인공와우 시뮬레이션(8 채널 보코더)이 함께 적용된 문장에 대한 어음인지도 검사를 실시하였다. 연구 결과, 인공와우 시뮬레이션이 함께 적용된 정현파 문장에서 통계적으로 낮은 어음인지도도를 보였으며, 특히 적은 수(2개)의 정현파 성분에서 큰 차이를 보였다. 본 연구는 정현파 모델에 대한 인공와우 사용자의 전기 청각 인지 패턴의 특성을 제시하며, 인공와우 어음처리 알고리즘의 개발에 대한 기초 자료를 제공한다.

핵심용어: 정현파 모델, 인공와우, 어음인지, 청각장애

ABSTRACT: Speech consists of the sum of complex sine-waves. This study investigated the perception of electrical hearing by applying the sinusoidal model to cochlear implant simulation. Fourteen adults with normal hearing participated in this study. The sentence recognition tests were implemented using the sentence lists processed by the sinusoidal model which extracts 2, 4, 6, 8 sine-wave components and sentence lists processed by the same sinusoidal model along with cochlear implant simulation (8 channel vocoders). The results showed lower speech recognition for the sentence lists processed by the sinusoidal model and cochlear implant simulation compared to those by the sinusoidal model alone. Notably, the lower the number of sine-wave components (2), the larger the difference was. This study provides the perceptual pattern of sine-wave speech for electrical hearing by cochlear implant listeners, and basic data for development of speech processing algorithms in cochlear implants.

Keywords: Sinusoidal model, Cochlear implant, Speech perception, Hearing impaired

PACS numbers: 43.64.Me, 43.71.An, 43.71.Ky

1. 서 론

어음의 발생은 혀에서 배출되는 공기 입자가 성대의 진동을 거쳐 성도를 따라 진행될 때 조음의 위치, 방법, 공명에 영향을 받아서 주파수와 에너지가 결정되는 현상이다. 공기 입자는 희박상과 압축상의 주기를 걸쳐 전파되며 어음은 이러한 주기가 아주 많이 얽혀 역동적으로 변화하는 복합음의 특징을 갖는

다. 푸리에 분석(Fourier analysis)에서 어음은 시간축에서 복잡하게 얽혀있는 정현파들의 합으로 계산될 수 있으며 이러한 변환식을 이용하여 음성의 변환, 분석, 합성 등 다양한 알고리즘에 활용되고 있다.

푸리에 분석을 이용한 음향 가공 모델 중 하나로 정현파 모델이 1986년 McAulay와 Quatieri^[1]에 의하여 제안되었다. 정현파 모델에서는 진폭, 위상, 그리고 주파수 정보를 고속 푸리에 변환을 이용해 어음으로

[†]Corresponding author: Sungmin Lee (smlee8345@gmail.com)

Speech-Language Pathology and Audiology, Tongmyong University, 428 Sinseon-ro, Nam-gu, Busan 48520, Republic of Korea
(Tel: 82-51-629-2134, Fax: 82-51-629-2019)



Copyright©2022 The Acoustical Society of Korea. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

부터 추출하여 지정된 개수의 에너지가 큰 정현파 성분을 추출한다. 추출된 정현파 성분은 시간축 상에서 인접한 프레임으로 연결되어 트렉으로 이어지며, 보간법을 사용하여 이어진 정현파 트렉을 보다 부드럽게 만들어 어음이 단지 몇 개의 지정된 숫자의 정현파 성분만으로 구성되도록 완성된다.

인공와우는 달팽이관이 손상된 고심도난청인의 달팽이관에 이식되어 청신경을 직접 전기자극하여 청각 정보를 전달하는 이식형 청각보조기기이다.^[2] 인공와우의 어음처리전략은 각 제조사에서 추구하는 어음 처리의 목적에 따라서 다양하다. 인공와우 초기 파형의 형태에 초점을 맞춘 파형전략으로는 Compressed Analog strategy(CA), Simultaneous Analog Stimulation(SAS), Continuous Interleaved Sampling(CIS)가 있으며 주파수에 초점을 맞춘 주파수전략으로는 F0/F2 strategy, F0/F1/F2 strategy, spectral peak(SPEAK), Advanced Combination Encoder(ACE) 등이 있다. 이 중 ACE와 SPEAK는 N-of-M strategy에 해당되는데 N-of-M은 M개의 주파수 채널 중 가장 에너지가 강한 N개의 채널만 선택하여 전기신호가 전달되는 개념의 전략을 의미한다. 최근에는 HiResolution Sound Processing, FSP, Fine Structure 4(FS4)와 같은 시간적 미세구조 정보의 전달을 위한 전략도 개발되어 상용화되고 있다.

이처럼 인공와우 신호처리전략은 다양하지만 전체적으로는 비슷한 처리과정을 공유하고 있다. 인공와우의 마이크로 들어온 음향 신호는 전중폭과정을 거쳐 증폭되고 청각기관의 생리적 특성을 반영하여 주파수 대역이 할당된 대역 필터뱅크로 보내어진다. 다음 각 필터에서 추출된 중심 주파수를 갖는 소리 에너지 정보의 포락선이 추출되고 이 포락선 성분을 이용하여 청신경을 자극한다. 즉, 어음은 느리게 변화하는 저주파수 에너지인 포락선 성분과 빠르게 변화하는 고주파수 에너지 정보인 시간적 미세구조 성분으로 구성되어 있지만 인공와우의 신호처리는 제한된 주파수 대역의 에너지가 큰 몇몇 채널의 포락선 성분을 추출하고 시간적 미세구조 정보 대신 양극성 펄스 전기 신호를 이용해 청각정보를 우리의 뇌까지 보내는 것이다.

국내 정현파 모델을 이용한 인공와우 환자들의 어음인지에 대한 임상연구는 보고되어 있지 않다.

2018년 국외에서 Lee *et al.*^[3]은 인공와우에 정현파 모델을 처음 적용하였다. 2명의 인공와우 사용자에게 1개, 3개, 6개의 정현파 어음을 제시하였을 때 선택된 정현파의 개수가 늘어날수록 어음인지도가 높아졌으며, 1개의 정현파를 추출하였을 때는 어음인지도가 크게 낮았으나 3개와 6개의 차이는 크지 않았다. 이후 Lee *et al.*^[4]은 1개, 2개, 6개의 정현파 성분으로 만든 영어권 문장을 사용하여 인공와우 환자로부터 어음인지도와 어음의 음질을 평가하는 임상시험을 수행하였다. 4명의 인공와우 환자가 참여한 연구에서 1개의 정현파가 선택된 모델일 때는 21%, 2개의 정현파일 선택된 어음에 대해서는 61%, 6개의 정현파가 선택된 정현파 모델 문장에 대해서는 64%의 문장인지를 나타냈다. 정상청력을 가진 대조군과 비교하여 유의미하게 어음인지력이 낮았으며 각 조건에 대한 인지적 패턴은 두 그룹이 차이가 없었다. 적은 수의 인공와우 환자로 인하여 통계적 보고가 없었으나 정상청력인과 인공와우를 통한 전기적 어음인지가 비슷한 패턴을 가지는 것으로 보고되었다.

향상된 인공와우 어음처리 알고리즘 개발에 대한 요구, 전기 청각 메커니즘의 특수성, 정현파 모델과 인공와우의 신호처리 방식이 복잡한 전체 어음 에너지 중 에너지가 큰 특정 주파수 성분을 이용한다는 공통점으로 볼 때 인공와우 환자에게 있어서 정현파 모델이 정상인과 비교해서 어떠한 어음인지 패턴을 가지는지에 대한 연구는 필요할 수 있다.

본 연구의 목적은 인공와우의 전기 청각 신호에 대하여 정현파 모델이 적용된 어음의 인지 특성을 분석하고 인공와우 신호처리에 적용 방안을 모색해 보고자 한다. 이를 위해 정현파 모델 처리된 문장과 정현파 모델 처리와 인공와우 시뮬레이션이 함께 적용된 문장 리스트를 정상 청력을 가진 연구 참가자에게 제시하여 어음인지 패턴을 비교 분석하고자 하였다.

II. 연구 방법

2.1 연구 대상

만 19세 이상의 정상청력을 가진 성인 14명(남자 3명, 여자 11명)이 본 임상시험에 참여하였다. 참여자의 평균연령은 24.36(표준편차: 4.63)세 였으며 정상

청력 여부를 확인하기 위하여 다음의 검사들을 시행하였다. 모든 참여자는 Interacoustics의 AT235를 사용하여 선행된 고막운동성검사 결과에서 정상 타입(type A)을 보였다. Interacoustics의 AC40 청력검사기를 사용하여 TDH-39 기도 헤드폰으로 순음 250 Hz~8 kHz를 옥타브 간격으로 25 dB HL로 제시하여 청력 스크리닝을 시행하여 모든 참여자에게서 청력 역치가 정상범위(< 20 dB HL)에 있음을 주파수 별로 확인하였다. 한국어 어음청각검사⁵⁾의 단어인지도검사(Word Recognition Score, WRS)에서 95% 이상의 어음 인지력을 확인하였다. 임상시험 전 모든 대상자들은 연구에 대한 설명을 충분히 이해하였고 모든 연구 절차는 동명대학교 생명윤리위원회의 승인 하에 이뤄졌다.

2.2 자극음

자극음으로 한국어 어음청각검사의 문장인지도 검사(Sentence Recognition Score, SRS) 일반용 리스트를 사용하였다. 어음 인지에 영향을 주지 않는 선에서 정현파 모델의 왜곡과 처리 시간을 줄이기 위해서 원래 원음인 스테레오의 44100 Hz 샘플링속도를 모노의 16 kHz 샘플링속도로 변환 및 다운샘플링 하였다. 8개의 일반용 문장 리스트를 “Sine wave and Sinusoid + Noise Analysis/Synthesis”⁶⁾ Matlab 스크립트를 사용하여 2, 4, 6, 8개의 정현파 모델이 적용되도록 자극음을 만들었다. 본 정현파 모델에서는 256 point의 Short Time Fourier Transform(STFT)를 사용하며 이는 아래와 같이 정의하고 계산된다.

$$\begin{aligned} STFT\{x(n)\}(m, \omega) &\equiv X(m, \omega) \\ &= \sum_{n=-\infty}^{\infty} x_n w_{n-m} e^{-i\omega n}. \end{aligned} \quad (1)$$

Eq. (1)에서 x_n 과 w_n 은 각각 시간축에서 변환되는 입력되는 신호와 window function을 의미한다. STFT에서 입력 신호의 위상과 진폭 변화(m)과 주파수(w)는 X로 표시된다. 하나의 윈도우에 128개의 샘플이 추출되어 그 중 가장 진폭이 큰 정현파 성분이 peak picking algorithm에 의해서 선택되며 나머지 어음에너지는 제거된다. 마지막으로 인접된 윈도우에서 선택

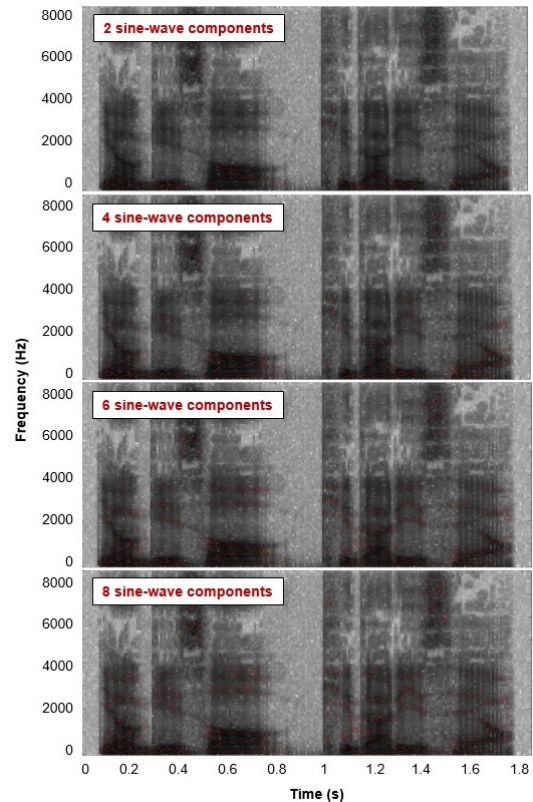


Fig. 1. (Color available online) Spectrograms of 2, 4, 6, and 8 sine-wave sentence. The red dots represent the selected sine-wave components by the sinusoidal model. The sample sentence shown is “Wait here.” in Korean.

된 정현파 성분을 이차보간법을 사용하여 부드럽게 이어지도록 한다. Fig. 1은 샘플 문장 “여기서 기다려주세요”에 대하여 2, 4, 6, 8개의 정현파 성분(빨강)이 선택된 spectrogram이다.

본 연구의 목표 중 하나는 인공와우 환자의 전기 청각과 정상 청력인의 음향 청각을 정현파모델을 적용하여 인지능력을 비교하는 것이다. 인공와우 환자의 전기 청각 메커니즘을 구현하기 위해서 AngelSim V1.08.01 noise vocoder⁷⁾를 사용하였다. 200 Hz~7 kHz 주파수 범위를 8 채널 noise vocoder로 Greenwood function (24 dB/octave filter slope) 주파수 할당⁸⁾을 갖도록 하였다. 정현파 모델 처리된 음원에 다시 한번 이와 같이 noise vocoder를 사용하여 인공와우 시뮬레이션을 적용하여 문장을 만들었다. 마지막 단계로 제작된 모든 자극음의 강도를 같도록 하기 위해서 Adobe

audition 3.0 프로그램을 사용하여 진폭에 대해서 정규화를 시행하였다.

2.3 연구 절차

임상 시험은 방음 설비가 갖춰진 청각부스에서 시행하였다. 제작된 자극음은 컴퓨터를 통해서 AUX 단자로 Interacoustics의 AC40 청력검사기와 경유하도록 하여 최종적으로 연구 대상자로부터 1 m 떨어진 전방스피커를 통해서 제시되었다. 8개의 KSA 문장 리스트는 8개의 음원 조건(4개의 정현파 성분 개수 X 인공와우 모델 처리 여부)에 따라 만들어졌고 이를 중복되지 않도록 무작위로 선택하여 65 dB Sound Pressure Level(SPL)로 제시되었다. 대상자들은 문장을 잘 듣고 들리는 문장을 최대한 따라 말하도록 하여 각 리스트에 40개의 목표 단어 중 정반응을 보인 단어의 수를 백분율로 계산하여 점수화 하였다.

2.4 자료 분석

SPSS 통계프로그램(Ver. 20.0, IBM Inc, Armonk NY, USA)을 사용하여 획득한 개인의 문장인지점수에 대한 기술 통계를 확인하였고 반복측정분산분석(repeated measures ANOVA)을 하여 음원의 조건에 따른 어음 인지도의 차이를 확인하였다.

III. 연구 결과

3.1 소리의 종류에 따른 결과

자극음의 종류에 따른 문장인지도 점수는 Fig. 2와 같다. 정현파 모델 문장에서는 2개, 4개, 6개, 8개의 평균점수가 57.86, 95.18, 98.93, 96.43으로 각각 나타났다. 정현파와 인공와우 모델이 함께 처리된 문장점수의 평균은 16(2개), 72(4개), 84.29(6개), 83.04(8개)로 얻어졌다. 통계분석 결과 정현파 모델을 사용하여 얻은 문장검사 점수가 정현파 모델에 인공와우 모델을 적용하여 얻은 문장검사 점수보다 유의미하게 높은 것으로 나타났다[F(1, 13) = 196.514, p = 0.000].

3.2 정현파 개수에 따른 결과

정현파의 개수도 문장인지점수에 유의미한 영향

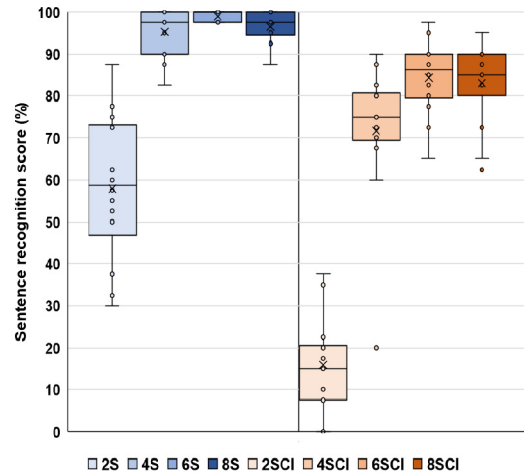


Fig. 2. (Color available online) Sentence recognition scores in percentage for 8 stimulus conditions. The number 2, 4, 6, and 8 indicates the number of sine-wave peaks selected in sinusoidal model. Note, S = sinusoidal model; SCI = sinusoidal model and cochlear implant simulation.

을 주는 것으로 나타났다 [F(3, 39) = 177.257, p = 0.000]. Bonferroni 사후분석 결과 2개의 정현파로 구성된 문장은 4, 6, 8개의 정현파로 구성된 문장보다 통계적으로 더 낮은 어음인지를 보였다(p = 0.000). 사후분석에서 다른 정현파 개수의 짝들은 통계적으로 유의미한 차이를 나타내지 않았다(p > 0.05).

IV. 고찰

본 연구에서는 인공와우의 전기 청각과 정상인의 음향 청각의 정현파 모델에 대한 인지패턴을 정상 청력인을 대상으로 인공와우 시뮬레이션을 사용하여 비교 분석하고자 하였다.

어음을 구성하는 수많은 주파수 성분들 가운데 단지 2개의 진폭을 가지는 정현파 성분만으로도 어음의 50% 이상 이해하는 인간의 어음인지력을 확인 하였다. 이와 함께 단지 4개의 정현파 성분만으로도 100%에 근접하는 어음인지를 가져오는 결과를 얻을 수 있었으며, 이는 100%에 가까운 독일어의 자모음 인지도를 위해서는 1.5 ms 당 최소 2개에서 4개의 정현파 성분이 필요하다는 선행논문과도 일치한다.^[9] 말소리 이해를 위한 음성학적 랜드마크라 할 수 있는 모음 포먼트, 포먼트 전이, 배음 주파수

등의 주요 단서를 정현파 모델에 의해서 단위 시간 당 선택된 4개 이하의 주파수 성분만으로도 반영해 내어 높은 인지능력을 이끌어낼 수 있음을 의미한다. 이를 환언하면 정현파 모델의 인지적 결과는 우리의 말소리가 갖고 있는 높은 잉여성을 설명하며, 본 연구 결과는 신호처리에 있어서 불필요한 정보를 제거하고 필요한 정보를 강조하는 알고리즘 개발 연구에 초석이 될 수 있다. 이러한 시도의 대표적인 예가 정현파 모델을 사용하여 소음억제 효과를 연구한 Kate^[10] 라고 할 수 있다. 그는 정현파 모델이 소음억제 효과를 발생시켜 어음인지도를 향상시킬 수 있는지에 대한 연구하였다. 연구에서 소음상황에서도 정현파 모델이 오히려 자모음 인지에 부정적 영향을 초래하고 신호대잡음비가 약 8 dB ~10 dB정도 나빠지는 것과 같아 소음억제의 긍정적 효과보다는 어음 인지에 부정적 영향을 준다고 보고 하였으나 정현파 모델을 사용하여 어음인지의 개선을 시도한 의미있는 연구라고 할 수 있다.

본 연구에서는 정현파 모델 단독일 때 보다 인공와우 시뮬레이션이 적용된 문장인지 점수가 유의미하게 낮은 점수를 보였다. 정현파 모델 단독일 때는 4개의 정현파에서도 95% 이상의 인지도를 보였으나 인공와우가 적용된 문장에서는 6개가 되어야 정현파의 개수 증가에도 문장인지 점수에 진보가 없는 플래토를 나타냈다. 특히 정현파의 개수가 적은 2개의 상황에서 문장인지에 큰 차이를 보였는데 이는 제한된 채널을 갖고 있는 인공와우 특성을 고려할 때 2개의 주파수 성분에 대한 주파수분석이 적합하게 이루어지지 않아 주파수 분별이 낮아졌음을 의미한다. 그럼에도 불구하고 6개, 8개 정현파에 인공와우 처리된 문장에서 80% 이상의 높은 어음인지도를 보였다. 본 연구에서는 인공와우 채널을 8개로 설정된 모델을 적용하였는데, 이는 8채널 보코더를 사용한 다른 연구들의 정현파 어음이 아닌 일반 어음인지도 결과에서 90% 이상의 수행능력을 보고하였다는 것을 고려했을^[11] 때 상당히 높은 수치이며 정현파 모델의 인공와우 적용 가능성을 보여준다. 정현파 모델에 의해서 선택된 특정 정현파 정보들이 인공와우의 필터뱅크에서 할당된 주파수에서 처리가 되는데, 만약 선택된 정현파가 서로 인접하여 인공와우의 같은 채널

에서 정보가 함께 처리된다면, 정현파 정보들이 각 인공와우 채널에서 따로 처리될 때보다 더 낮은 어음인지도를 가져오게 될 것이다. 따라서 본 연구에서보다 더 많은 인공와우 채널을 사용한다면 정현파가 같은 주파수 필터뱅크에 할당될 확률을 낮춰서 더 높은 어음인지도를 가져올 것으로 사료된다.

인공와우는 보청기로 소리를 들을 수 없는 고심도 청각장애인에게 전기자극을 통해서 소리를 제공하는 획기적인 의료기기이다. 하지만 인공와우를 통한 전기 청각 메커니즘은 건청인들의 음향 청각 메커니즘과 주파수, 강도, 시간적 해결능력에서 차이가 있다. 본 연구에서는 한국어에 대하여 정현파 모델을 적용하여 주파수 제한된 전기적 어음인지 패턴을 확인 하였다. 보다 실질적인 정현파 모델의 적용을 위해 인공와우 시뮬레이션이 아닌 실제 인공와우 환자를 대상으로 보다 정교하게 디자인된 임상 연구가 요구된다.

본 연구에서 얻은 임상적 결과물은 기존의 인공와우가 갖고 있는 제한점들(소음속 어음인지, 음악감상, 메모리 등)을 해결하는데 도움이 될 것이다. 더불어 효과적인 인공와우 어음처리 알고리즘의 개발과 적합한 전기자극을 제시하기 위한 임상 맵을 결정하고 재고하는데 기여할 것이다.

감사의 글

이 논문은 2021년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2021S1A5A80 6455811).

References

1. R. J. McAulay and T. F. Quatieri, "Speech analysis/Synthesis based on a sinusoidal representation," *IEEE Trans. Acoust.* **34**, 744-754 (1986).
2. D. H. Lim, "Speech coding strategies in the scala tympanic electrode array of the cochlear implant system" (in Korean), *Audio*, **1**, 8-13 (2005).
3. S. M. Lee, S. Akbarzadeh, S. Singh, and C. T. Tan, "A speech processing strategy based on sinusoidal speech model for cochlear implant users," *Proc. APSIPA. Convention paper*, 393-397 (2018).

4. S. M. Lee, S. Akbarzadeh, and C. T. Tan, "Neural tracking of band-limited sine-wave speech in normal hearing and cochlear implant listeners," *Int. J. Speech-Lang. Pathol.* **8**, 38-49 (2020).
5. J. H. Lee, S. J. Cho, J. S. Kim, H. S. Jang, D. H. Lim, K. W. Lee, and H. J. Kim, *Korean Speech Audiometry (KSA)* (Hakjisa, Seoul, 2010), pp. 1-448.
6. *Personal Homepage*, <https://www.ee.columbia.edu/~dpwe/resources/matlab/sinemodel/>, (Last viewed December 25, 2021).
7. *Personal Homepage*, http://www.tigerspeech.com/angel_sim/angelsim_about.html, (Last viewed December 25, 2021).
8. D. D. Greenwood, "A cochlear frequency-position function for several species—29 years later," *J. Acoust. Soc. Am.* **87**, 2592-2605 (1990).
9. O. Timms, *Speech processing strategies based on the sinusoidal speech model for the profoundly hearing impaired*, (Doctoral thesis, University Hospital of Zurich, 2003).
10. J. M. Kates, "Speech enhancement based on a sinusoidal model," *J. Speech Hear. Res.* **37**, 449-464 (1994).
11. P. C. Loizou, M. Dorman, and Z. Tu, "On the number of channels needed to understand speech," *J. Acoust. Soc. Am.* **106**, 2097-2103 (1999).

저자 약력

▶ 이 성 민 (Sungmin Lee)



2017년 12월 : University of Memphis 청각학 박사

2020년 2월 : University of Texas at Dallas, Electrical and Computer Engineering 포닥연구원

2020년 3월 ~ 현재 : 동명대학교 언어치료 청각학과 조교수