OPINION

Timing of speech envelope entrainment to the brain while speaking, hearing, and self-listening

Jack Ma

Ma J. Timing of speech envelope entrainment to the brain while speaking, hearing, and self-listening. J Neuropathol. 2022; 2(5):44-5.

ABSTRACT

The dynamics of speech envelope tracking during speech production, listening, and self-listening are examined in this study. Participants engage in a paradigm where they listen to normal speech (Listening), create natural speech (Speech Production), then listen to the recording of their own speech (Self-Listening), all while their brain activity is being monitored by an EEG. We employed a Gaussian copula mutual information measure to evaluate the link between information content in the EEG and auditory signals after time-locking the collection of EEG data and the recording and playback of auditory signals. We determined several latencies for maximal speech envelope tracking during speech production and speech perception in the 2 Hz-10 Hz frequency range. Maximum speech tracking occurs roughly between the vocalization phase of speech production and the auditory presentation phase of perception communication.

Key Words: *Brain activity*

INTRODUCTION

Speakers and listeners exchange information during conversation by using speech to transmit signals from one brain to another. Due to transmission delays from the speaker to the listener and other physical constraints that mediate speech communication, the speaker's brain activity should lead the listener's in accordance with the concept of the speech chain. Given the speed of sound for discussions over 1 m^{-2} m in distance, transmission delays from the speaker to the listener are among these that can be anticipated. These physical delays, however, are insignificant when compared to the probable cerebral delays brought on by the speaker's and the listener's neural mechanisms for comprehension and speech planning, respectively.

It would seem natural that there will be a measurable delay in conveying a concept from a speaker to a listener during conversation given the inherent nature of these delays. Perhaps surprisingly, presenters and listeners frequently exhibit instantaneous inter-brain synchronization without detectable phase-lag in neurological observations. In other words, it is clear that both the speaker and the listener's brains are active simultaneously without presuming that the listener's brain impulses arrive later than those of the speaker. The concept of a voice chain is called into question by observations of instantaneous inter-brain synchronization.

We need to rule out a few methodological issues, though, before we can say that linguistic inter-brain synchronization during speaking and listening, measured with high temporal resolution approaches, indicates mutual predictive processing. The phenomenon of brain entrainment to speech is the main area of worry. Here, the auditory signal or speech envelope's overall loudness slowly modulates while the brain "tracks" it. The reasoning for how measured inter-brain synchronization is influenced by brain entrainment to speech is simple: Inter-brain coupling may simply be a byproduct of simultaneous brain entrainment to speech sounds in the speaker and listener. This is because there will be coupling between the listener's cortex and the speech, and because the speaker is also hearing their own speech (both through air and bone conduction).

That is, because both brains are concurrently tracking the speech envelope at the same latency, immediate inter-brain coupling may take place. The notion that inter-brain synchronization is caused by concurrent brain entrainment to speech sounds is refuted if we find that the speaker and listener experience different delays in speech tracking. Instead, it might back up prognostic inter-brain synchronization accounts.

We looked at the EEG latency at which the strongest brain tracking of speech envelopes during speech creation, listening, and self-listening is seen. In order to link brain activity to the slow amplitude modulation frequencies provided by the speech envelope, which are essential for speech intelligibility, we demonstrated a precise timeframe for speech tracking in speakers and listeners. The results show that, on average, peak neuronal tracking occurs during listening about 110 ms following speech presentation. The same subject (Self-

Editorial office, Journal Of Neuropathology, United Kingdom

Correspondence: Jack Ma, Editorial office, Journal Of Neuropathology, United Kingdom, E-mail: neuropathology@pulsusinc.com

Received: 2 September 2022, Manuscript No PULNP-22-5405; Editor assigned: 4 September 2022, PreQC No. PULNP-22-5405 (PQ); Reviewed: 17 September 2022, QC No. PULNP-22-5405 (Q); Revised: 18 September 2022, Manuscript No. PULNP-22-5405 (R); Published: 26 September 2022, DOI: 10.37532/pulnp.2022.2(5).44-5

Listening) or a separate participant (Listening), two circumstances that have been demonstrated to cause equal temporal lobe activation in functional imaging studies, both produced equivalent latencies when listening to speech.

Since there is little time between participants' production and the replay for the Self-Listening condition, various layers of language representation (such as syntax, lexical items, prosody, and phonetics) are putatively predictable. Despite these variations in prediction, the electrophysiological findings presented support equal brain timing for processing speech from others and speech created by the individual. This finding refutes the claim that better prediction speeds up brain speech recognition, however prediction may have an impact on the entrainment's potency. For brain entrainment to speech during enhanced top-down focal attention, there have been reports of improvements in the strength of speech tracking with no change in latency. It's interesting to note that when compared to listening, self-listening showed bigger magnitudes of speech tracking (larger GCMI values) and more between-subject variance. The neural sources associated with listening were located in areas of the brain that had previously been linked to speech tracking, such as the inferior frontal and auditory regions, where intelligibility influences cerebra-acoustic coherence, and the precentral gyrus, where top-down connectivity to auditory areas has been linked to speech comprehension. Our findings are consistent with prior publications in that there is a positive delay in coupling between these regions and the speech envelope. Additionally, there is evidence that the phonetic information is processed 114 ms after the speech signal.

The current study may have some limitations due to the possibility that muscle activity related to articulatory motions of the tongue, lips, or occipitalis muscle could muddle EEG responses in the speech production condition. Despite our best efforts to remove non-neural signals from the EEG data using the ASR algorithm and AMICA technique, we would argue that the increased magnitude of GCMI in speech production as compared to speech perception—rather than the differential timing—may be explained by these muscle artefacts. More in-depth analysis of data from anatomically (or functionally) constrained auditory and motor brain regions, which can provide evidence for source-specific features of neural entrainment, would be beneficial for future electrophysiological studies addressing the time course of brain entrainment to speech during production and listening.