DOI: 10.1111/nyas.15179

ORIGINAL ARTICLE

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Enhancing speech perception in noise through articulation

Maxime Perron[1,2] | Qiying Liu[1,2] | Pascale Tremblay[3,4] | Claude Alain[1,2,5,6]

[1]Department of Psychology, University of Toronto, Toronto, Ontario, Canada

[2]Baycrest Academy for Research and Education, Rotman Research Institute, North York, Ontario, Canada

[3]CERVO Brain Research Center, Quebec City, Quebec, Canada

[4]École de Réadaptation, Faculté de Médecine, Université Laval, Quebec City, Quebec, Canada

[5]Institute of Medical Sciences, University of Toronto, Toronto, Ontario, Canada

[6]Music and Health Science Research Collaboratory, University of Toronto, Toronto, Ontario, Canada

Correspondence
Maxime Perron, Rotman Research Institute, Baycrest, 3560 Bathurst Street, Toronto, ON, Canada M6A 2E1.
Email: maxime.perron@mail.utoronto.ca

Funding information
William Demant Foundation, Grant/Award Number: 20-1260; Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: RGPIN-2021-02721; Canadian Institutes of Health Research

Abstract

Considerable debate exists about the interplay between auditory and motor speech systems. Some argue for common neural mechanisms, whereas others assert that there are few shared resources. In four experiments, we tested the hypothesis that priming the speech motor system by repeating syllable pairs aloud improves subsequent syllable discrimination in noise compared with a priming discrimination task involving same–different judgments via button presses. Our results consistently showed that participants who engaged in syllable repetition performed better in syllable discrimination in noise than those who engaged in the priming discrimination task. This gain in accuracy was observed for primed and new syllable pairs, highlighting increased sensitivity to phonological details. The benefits were comparable whether the priming tasks involved auditory or visual presentation. Inserting a 1-h delay between the priming tasks and the syllable-in-noise task, the benefits persisted but were confined to primed syllable pairs. Finally, we demonstrated the effectiveness of this approach in older adults. Our findings substantiate the existence of a speech production–perception relationship. They also have clinical relevance as they raise the possibility of production-based interventions to improve speech perception ability. This would be particularly relevant for older adults who often encounter difficulties in perceiving speech in noise.

KEYWORDS
aging, articulation, speech motor, speech perception, speech production

INTRODUCTION

How closely speech perception and production are linked, whether they share common processing mechanisms, and to what extent they influence each other continues to challenge our understanding of the complex mechanisms underlying human language.[1] The fundamental distinction between theories that advocate a close link between speech perception and production and those that defend their relative independence forms the basis of this debate. Early ideas, such as the motor theory of speech perception,[2,3] state that we perceive speech by internally simulating the movements required to produce the speech sounds we hear. This theory assumes common schema between speech perception and production based on articulatory–motor representations.[3] Alternatively, the acoustic theories of speech perception propose a more modular approach, asserting that speech perception and production operate independently and that both rely on different representations, with acoustic properties being the main object of perception.[4]

Evidence from neuroimaging studies suggest a common network enabling speech perception and production.[5–8] For instance, Wilson et al.[6] showed that the upper part of the left ventral premotor cortex was activated during both the production and perception of meaningless syllables in quiet. Moreover, young and older adults presented with degraded or distorted speech sounds show activation in the articulatory system.[7,9,10] The recruitment of the speech motor system under unfavorable listening conditions has been taken as evidence for compensatory mechanisms aimed at decoding impoverished acoustic representations.[9] Additionally, temporary disruption of the articulatory system using inhibitory repetitive transcranial magnetic stimulation (rTMS) impairs speech discrimination in noise[11,12] and in quiet.[13,14] Brisson and Tremblay[15] used excitatory rTMS and found that stimulation of the left ventral premotor cortex improved the ability to discriminate pairs of syllables in noise. The impact of rTMS on this area outweighed its effects on the left posterior superior temporal sulcus, highlighting the importance of the articulatory system in speech-in-noise perception compared with regions usually associated with phonological processing. All the above evidence advocates for a close link between speech perception and production, supporting the notion that the two processes share common representations. However, debate persists over the nature of this relationship. Although some researchers argue that the involvement of the motor system is essential to speech decoding,[11,16] others propose that motor contributions to perception are merely epiphenomenal[17,18]—a by-product of shared neural networks. As an example, consider the dual-stream model of speech processing.[19] Although this model proposes a direct link between motor and auditory processes, it suggests that this link does not play a central role in normal listening conditions.[20,21] Rather, the motor system might have a modulatory effect on speech perception, notably by supporting speech in challenging listening situations.[22]

Stokes et al.[23] designed a dual-task paradigm to explore how motor representations support speech-in-noise perception. Participants were asked to subvocally repeat the word *the* while identifying phonemes in noise. The hypothesis was that subvocal repetition would disrupt the motor system in the same way as inhibitory rTMS, thereby reducing motor resources for speech perception and impairing task performance. Control conditions included mandible movement, foot tapping, and passive listening. Results indicated that phonemic identification was consistently worse than passive listening only in the articulatory suppression conditions, supporting the role of the articulatory system in speech perception. However, as the impact was marginal in terms of effect size, the authors concluded that the motor system plays a minor modulatory role.

Here, we propose that even a modest interaction between perception and production, as suggested by the prior study, could be used to enhance speech processing in adverse listening conditions. This could have significant implications for older adults, who often have difficulty communicating in noisy environments and for whom there is currently no evidence-based clinical intervention. We developed an experimental approach to explore how production and perception interact, focusing on the potential enhancement of perception through syllable repetition. In contrast to Stokes et al.[23] who used articulatory suppression simultaneously with a speech perception task to replicate inhibitory rTMS effects on speech perception, our approach involved syllable repetition prior to a speech perception task to mimic excitatory rTMS benefits through speech motor priming. Theoretical differences arise from distinct mechanisms: Articulatory suppression disrupts speech processes by engaging both motor and perceptual processes simultaneously, whereas speech motor priming involves preactivating and preparing the system prior to a speech perception task, potentially enhancing subsequent perception.

Participants were divided into two groups, each performing separate tasks (hereafter referred to as priming tasks): one involving vocal production of syllable pairs in quiet (Production group) and the other requiring discrimination (same–different judgment) of syllable pairs in quiet (Perception group). Subsequently, all participants were asked to make a same–different judgment for syllable pairs in noise (hereafter referred to as the syllable-in-noise [SiN] task). The latter included syllable pairs from the priming tasks (old pairs) and new syllable pairs to assess near transfer during the main task. If speech perception operates independently of production, individuals who have performed a discrimination task in a quiet environment (Perception group) should perform better at the SiN task, given the similarity between the two tasks. Conversely, if perception and production share representations and resources, the articulatory motor representations activated during production should facilitate perception of the same syllables in noise.

We initially tested our hypotheses in a cohort of young adults (Experiment 1). In Experiment 2, we manipulated the stimulus presentation modality during the priming tasks (auditory vs. visual). Experiment 3 varied the delay between the priming and SiN tasks (no delay vs. 1-h delay). In Experiment 4, we investigated the effects of age on speech motor priming.

METHODS

Experiment 1

Experiment 1 tested the hypothesis that priming the speech motor system through syllable repetition results in subsequent improvement in SiN perception.

Participants

All participants were recruited from Baycrest's participant database and e-mails sent to the Department of Psychology of University of Toronto. To be eligible, participants had to be between 18 and 35 years of age and be native English speakers (or have learned English before the age of 6). Participants with speech or language disorders, significant medical or neurocognitive problems, visual impairments, tinnitus, otological disorders, or past/current experience with hearing aids were excluded. A cohort of 34 young adults was initially recruited, but 4 were excluded: 2 for technical reasons, 1 did not meet the inclusion criteria, and 1 participant withdrew from the experiment due to difficulties in understanding the main task. All participants gave written informed consent and received financial compensation. This study was approved by Baycrest's Research Ethics Board (project no. 21–32).

Half of the participants (n = 15) were assigned to the Perception group (M_{age} = 21.6 ± 2.1, 18–26 years, 7 females) and the other half

TABLE 1 Experiment 1 participants' characteristics.

	Perception (n = 15, 7 females)				Production (n = 15, 11 females)				
Characteristic	Mean	SD	Min	Max	Mean	SD	Min	Max	t-value
Age	22	2.1	18	26	22	3.6	18	31	0.44
Education (years)	16	1.8	12	19	16	3	12	25	0.44
Number of spoken languages[a]	2.3	0.8	1.0	4.0	2.6	1.1	1.0	5.0	0.94
Right ear PTA4[b]	8.9	5.0	1.2	24	6.8	6.6	−3.8	24,0	0.97
Left ear PTA4[b]	7.8	4.8	2.5	16	6.2	4.6	−1.2	12.0	0.93
Better ear PTA4[c]	6.6	3.5	1.2	14	4.3	4.1	−3.8	12.0	1.62
Interaural difference[d]	3.5	4.1	0	16	4.3	2.7	1.2	11	0.65
QuickSIN score[e]	0.4	0.9	−1	1.8	−0.1	1.1	−2.2	1.0	1.52

[a]Number of spoken languages = number of languages spoken, including native language.
[b]PTA4 = pure-tone average thresholds at 0.5, 1, 2, and 4 kHz for each ear individually, measured in decibels (dB hearing level).
[c]Better ear PTA4 = pure-tone average thresholds at 0.5, 1, 2, and 4 kHz for the better ear, measured in decibels (dB hearing level).
[d]Interaural difference = absolute difference between the PTA4 of the left and right ear.
[e]QuickSIN score = QuickSIN measures the ability to perceive sentence in noise. Scores are reported as signal-to-noise ratio (SNR) loss, the difference (in dB SNR) between a listener's performance and that of normal-hearing adult controls. Lower score indicates better speech perception ability.

(n = 15) to the Production group (M_{age} = 22.1 ± 3.6, 18–31 years, 11 females). To ensure that the groups were comparable, we performed a series of independent t-tests on age, years of education, number of languages spoken, pure-tone hearing, and central auditory processing measured with the QuickSIN test. Both groups were also comparable in terms of biological sex, $\chi^2(1) = 1.3$, $p = 0.3$, and musicianship (yes/no), $\chi^2(1) = 0.1$, $p = 0.7$. Table 1 provides a summary of participants' characteristics.

Experimental procedure

Participants were invited to take part in a single session lasting 90 min. The entire procedure took place in a soundproof, double-walled room. Participants first completed a consent form and demographic questionnaires. They then underwent a hearing assessment, including pure-tone audiometry and the QuickSIN test.[24] Following the hearing assessment, participants were randomly assigned to either the Perception group or the Production group using an online flip coin generator (heads for Perception, tails for Production). The online generator guaranteed an equal likelihood of obtaining heads and tails. Pairs were then created so that if the first participant was assigned to one group, the next was assigned to the other. The same process continued for subsequent pairs. Participants performed the priming task corresponding to their assigned group, either repeating syllables (Production) or discriminating syllables in quiet (Perception). Finally, all participants took part in the SiN task.

Hearing assessment

Hearing assessment was performed in a double-walled soundproof room using a clinical audiometer (GSI61, Grason Stadler). Peripheral hearing was assessed by pure-tone audiometry. Each ear was tested separately at 0.5, 1, 2, 3, 4, 6, and 8 kHz, and a pure-tone average (PTA4: 0.5, 1, 2, and 4 kHz) was calculated for both ears. The two groups were matched for the PTA4 in each ear, the PTA4 of the better ear, and the difference between the PTA4 of the two ears (Table 1).

Central auditory processing was assessed using the QuickSIN test.[24] The test consists of lists of six sentences, each containing five key words. Sentences are presented binaurally at 70 dB hearing level with concurrent four-talker babble noise at signal-to-noise ratio (SNR) between 25 and 0 dB, decreasing in 5 dB steps. Participants receive one point for each correctly recalled keyword. The cumulative score per list is calculated. The QuickSIN score is equal to 25.5 minus the cumulative score, representing the SNR loss. Lower scores reflect better speech-in-noise ability. We used four lists. Both groups were matched for QuickSIN score.

Priming tasks

The priming tasks consisted of either the repetition or discrimination of 153 pairs of Canadian–English monosyllabic consonant–vowel–consonant words that were selected from the Massive Auditory Lexical Decision database.[25] Stimuli were recorded by a 28-year-old Canadian phonetics student. Half the pairs were identical (e.g., /tap/-/tap/) and half were different (e.g., /bat/-/pat/). Pairs were either different on the onset (i.e., the first consonant) (33% of different trials), the nucleus (i.e., the vowel) (33%), or the coda position (i.e., the last consonant) (33%). Stimulus files were edited to normalize root-mean-square intensity to 70 dB sound pressure level (SPL). In each trial, two syllables were presented diotically with an inter-stimulus interval of 300 ms. Syllables were presented using Presentation Software (Neurobehavioral System) in quiet.

During the presentation of the syllables, a white fixation cross centered on a dark gray background was presented. Following the presentation of the syllables, a green question mark (?) was presented to indicate to the participants to answer. The Perception group indicated whether the syllables were the same or different using a response box (MilliKey, model SR-5 r2). The Production group was asked to repeat the two syllables without instructions on whether the syllables differed. Participants were given a maximum of 3 s to respond. The inter-trial interval was 1.5 s.

Syllable-in-noise (SiN) task

The SiN task had the same procedure as the priming task used in the Perception group, with the exception that the speech sounds were embedded in multi-talker babble noise. The task comprised the 153 trials from the priming tasks (i.e., old pairs) and an additional 153 trials with syllables not previously presented in the priming tasks (i.e., new pairs). Half of the pairs were identical, and half were different. The same talker recorded syllables from both old and new pairs. Syllables were presented simultaneously with a multi-talker's babble noise of a large group of people (∼30) talking in a large, open room (https://freesound.org/people/mefrancis13/sounds/210611/) at a SNR ($\text{Pressure}_{\text{signal}}/\text{Pressure}_{\text{noise}}$) of −3 dB. Noise files were edited to normalize root-mean-square intensity to 73 dB SPL. The noise was presented throughout the entire trial, starting from the onset of the trial to the participant's response. All participants indicated whether the syllables were the same or different using a response box (MilliKey, model SR-5 r2). Participants performed 12 practice trials.

Performance was measured in terms of accuracy (percentage of correct answers) and reaction time (RT) in milliseconds (ms). Response accuracy was also analyzed within the framework of signal detection theory,[26] using *d*-prime (d') and criterion (c). Sensitivity (d') refers to the ability to accurately discriminate between different pairs when they were different (sensitivity = z(Probability["different" | DIFFERENT]) − z(Probability["different" | IDENTICAL]). Criterion (c) refers to the tendency to favor one response option (same/different) over another, known as response bias. A c value of 0 indicates the absence of bias, meaning that participants have the same probability of choosing one answer or the other. A negative c value indicates a bias in favor of the *identical* choice, whereas a positive c value indicates a bias in favor of the *different* choice.

Statistical analyses

Data and the analysis scripts are publicly available at https://doi.org/10.5683/SP3/USPBLW. Data were analyzed using R version 4.2.3[27] in R Studio.[28] No outliers (values more than three interquartile range) were identified. The variable distributions were visually inspected using histograms and *Q*–*Q* plots. For each dependent measure, linear-mixed models (LMMs) were estimated using the *mixed* function of the *afex* package, which conducts mixed models with *lme4* and computes *p*-values for all fixed effects. Degrees of freedom for the fixed effects were adjusted using Satterthwaite's method to address potential heterogeneity and account for small sample sizes. In each model, Group (Perception, Production) was considered the between-subject variable, Trial type (Old, New) as the within-subject variable, and Participants as the random factor. Significant main effects and interactions were further analyzed using the *emmeans* function for contrasts of estimated marginal means. Effect sizes are expressed as partial eta-squared ($\eta_p{}^2$).

Experiment 2

In Experiment 1, we showed that, compared with active listening, the act of repeating syllables improved subsequent SiN ability. This result is in line with our hypothesis that engaging the articulatory system prior to a speech task improves speech perception, probably due to shared articulatory representations. Based on these results, in Experiment 2, we introduced a switch from auditory to visual stimuli during the priming phase. Our hypothesis was that the modality of stimulus presentation (visual or auditory) should not influence performance, as we considered the critical factor to be the syllable production process.

Participants

Thirty-two additional young adults were recruited and compared with those from Experiment 1. Of the 32, 2 were excluded (1 did not meet the inclusion criteria and 1 had difficulty understanding the main task), leaving 15 participants in the Perception group (M_{age} = 24.8 ± 4.3, 18–33 years, 9 females) and 15 participants in the Production group (M_{age} = 25.5 ± 3.5, 20–33 years, 10 females). We performed power analysis based on the results from Experiment 1 for the predictor "group," using the *simr* package with 1000 permutations. The results yielded a power estimate of 93.7% for detecting a statistical difference at a significance level of 0.05 with a sample of 15 participants per group.

Table 2 provides a summary of the new group of participants included in Experiment 2. The groups were matched in terms of age, education, number of spoken languages, pure-tone hearing, QuickSIN score, biological sex, $\chi^2(1) = 0$, $p = 1.0$, and musicianship (yes/no), $\chi^2(1) = 0$, $p = 1.0$.

Procedure

We followed the protocol of Experiment 1, the only modification being the stimulus presentation modality during the priming tasks. Stimuli were presented as white text against a black background. The presentation duration of the visual stimuli was determined based on the duration of the auditory stimuli in Experiment 1. Syllables were presented one at a time with an inter-stimulus interval of 300 ms. The Perception group had to determine whether the syllables on the screen

TABLE 2 Experiment 2 participants' characteristics.

Characteristic	Perception ($n = 15$, 9 females)				Production ($n = 15$, 10 females)				
	Mean	SD	Min	Max	Mean	SD	Min	Max	*t*-value
Age	25	4.3	18	33	25	3.5	20	33	0.47
Education (years)	16	2.5	12	22	17	2.8	12	22	1.31
Number of spoken languages[a]	2.3	0.8	1	4	2.1	0.8	1	3	0.90
Right ear PTA4[b]	8.2	4.2	1.2	16	7.5	3.1	3.8	15	0.56
Left ear PTA4[b]	8.5	3.6	2.5	14	6.8	4.1	0	14	1.19
Better ear PTA4[c]	7.2	3.8	1.2	14	6.1	3.6	0	14	0.86
Interaural difference[d]	2.2	1.4	1.2	6.2	2.2	1.3	0	3.8	0.17
QuickSIN[e]	0.42	1.2	−2	2.2	0.033	0.77	−1.5	1.2	1.04

[a]Number of spoken languages = number of languages spoken, including native language.
[b]PTA4 = pure-tone average thresholds at 0.5, 1, 2, and 4 kHz for each ear individually, measured in decibels (dB hearing level).
[c]Better ear PTA4 = pure-tone average thresholds at 0.5, 1, 2, and 4 kHz for the better ear, measured in decibels (dB hearing level).
[d]Interaural difference = absolute difference between the PTA4 of the left and right ear.
[e]QuickSIN score = QuickSIN measures the ability to perceive sentence in noise. Scores are reported as signal-to-noise ratio (SNR) loss, the difference (in dB SNR) between a listener's performance and that of normal-hearing adult controls. Lower score indicates better speech perception ability.

were the same or different. The Production group was asked to read the two syllables out loud.

Statistical analyses

We employed the same analytical approach as in Experiment 1, except that Modality (auditory, visual) was included as a between-subject factor in the LMMs. Three outliers (values more than three interquartile range) were detected for accuracy. Analyses were performed both with and without the outliers. The conclusions remained consistent, except for one main effect in two metrics.

Experiment 3

Following the results observed in Experiments 1 and 2, we investigated whether the priming effects lasted beyond an immediate period and persisted after an interval of 1 h. We anticipated that the resources/representations engaged by the production task would fade over time and that the priming effects on perception would be smaller with longer delay between the priming and the SiN tasks. We returned to auditory presentation for the priming tasks, as Experiment 2 showed no significant effect of stimulus modality.

Participants

Thirty-two additional young adults were recruited and compared with the young adults in Experiment 1. The sample for Experiment 3 was also distinct from Experiment 2. Of the 32, 2 were excluded because they performed at chance level (at or below 50%) in the main speech perception task, leaving 15 participants in the Perception group ($M_{age} = 24.4 \pm 4.8$, 19–35 years, 11 females) and 15 participants in the Production group ($M_{age} = 25.9 \pm 4.0$, 20–31 years, 11 females).

Table 3 provides a summary of the new groups of participants included in Experiment 3. The two groups were matched in terms of age, education, number of spoken languages, and pure-tone hearing. They were also matched in terms of biological sex ($\chi^2(1) = 0$, $p = 1.0$) and musicianship (yes/no) ($\chi^2(1) = 0$, $p = 1.0$). However, the Production group showed better central auditory processing ability measured with the QuickSIN (i.e., lower score) than the Perception group.

Procedure

We followed the procedure in Experiment 1, but this time with a 1-h delay between the priming and the SiN tasks. During this 1-h delay, participants performed hearing assessments and completed various questionnaires, which are not analyzed here, to ensure that participants were all engaged in the same activity during the delay.

Statistical analyses

We used the same analytical approach as in Experiment 1, except that Time (no delay, 1-h delay) was additionally included as a between-subject factor in the LMMs. One outlier (values more than three interquartile range) was identified for d' and c. As we found a significant difference in QuickSIN between the two groups, analyses were performed with and without the QuickSIN score as a covariate. The results remained the same. Results are reported without the covariate.

TABLE 3 Experiment 3 participants' characteristics.

Characteristic	Perception (n = 15, 11 females)				Production (n = 15, 11 females)				
	Mean	SD	Min	Max	Mean	SD	Min	Max	*t*-value
Age	24	4.8	19	35	26	4.0	20	31	0.91
Education (years)	17	2.8	13	24	18	2.7	14	23	1.32
Number of spoken languages[a]	2.6	0.8	1.0	4.0	2.2	0.9	1.0	4.0	1.24
Right ear PTA4[b]	7.8	5.8	2.5	21	7.6	5.2	1.2	18.0	0.13
Left ear PTA4[b]	9.9	8.4	0.0	29	6.8	4.7	0.0	16.0	1.24
Better ear PTA4[c]	7.2	6.1	0.0	21	5.8	4.5	0.0	15.0	0.73
Interaural difference[d]	3.2	2.7	0.0	10	2.8	1.8	0.0	6.2	0.60
QuickSIN[e]	0.8	0.7	–0.5	2.0	0.0	1.0	–2.0	1.5	2.48*

[a]Number of spoken languages = number of languages spoken, including native language.
[b]PTA4 = pure-tone average thresholds at 0.5, 1, 2, and 4 kHz for each ear individually, measured in decibels (dB hearing level).
[c]Better ear PTA4 = pure-tone average thresholds at 0.5, 1, 2, and 4 kHz for each ear individually, measured in decibels (dB hearing level).
[d]Interaural difference = absolute difference between the PTA4 of the left and right ears.
[e]QuickSIN score = QuickSIN measures the ability to perceive sentence in noise. Scores are reported as signal-to-noise ratio (SNR) loss, the difference (in dB SNR) between a listener's performance and that of normal-hearing adult controls. Lower score indicates better speech perception ability.
*$p < 0.05$ indicates statistical significance.

Experiment 4

The results of Experiments 1–3 indicate a significant influence of speech production on perception in young adults. In Experiment 4, we explored whether this effect is present in older adults.

Two theories on the role of the motor system in speech perception in aging are considered in the literature.[29] The Motor Compensation Hypothesis suggests that older people make greater use of articulatory motor resources during speech perception to compensate for age-related decline in auditory processing. Conversely, the Auditory–Motor Decline Hypothesis posits that the age-related changes in auditory function result in decreased input to the central auditory system, causing reduced engagement of the speech motor system during perception. In other words, the former hypothesis argues that motor representations aid in compensating for age-related speech perception difficulties, whereas the latter suggests that aging impairs auditory–motor integration, thereby contributing to speech perception difficulties. In support of a motor-mediated compensatory mechanism, Du et al. showed increased activity in frontal speech areas in older adults compared to their younger counterparts, positively correlating with speech-in-noise ability.[9] Conversely, in support of an age-related decline in motor activity during speech-in-noise perception, it has been shown that articulatory suppression impairs speech perception in young but not in older adults.[30] The absence of suppression in older adults suggests that the speech motor system in older adults might not support speech perception to the same degree as it does in young adults. However, the lack of articulatory suppression compared to other dual control tasks could also be related to the cognitive demands inherent in the dual task, which could lead older adults to disengage due to the increased effort required.

Speech motor priming offers a promising way to study the relationship between speech perception and production in older adults without imposing cognitive demands by separating the production and perception tasks. If speech motor representations are involved in a compensatory mechanism for the age-related decline in auditory processing (i.e., Motor Compensation Hypothesis), then we should observe a priming effect in older adults akin to what was observed in young adults across the preceding three experiments. Conversely, if aging results in reduced recruitment of the speech motor network during perception (i.e., according to the Auditory–Motor Decline Hypothesis), we may observe minimal or no speech motor priming effect in older adults.

Participants

Thirty-two older adults were recruited and compared with the young adults in Experiment 1. The eligibility criteria for older adults were identical to those applied to young adults (refer to Experiment 1), with the only exception being that older adult participants had to fall within the age range of 60–90 years. Of the 32, 2 were excluded because they did not meet the inclusion criteria (hearing aids and tinnitus). The final sample was randomly divided into 15 participants in the Perception group (M_{age} = 72 ± 6.2, 62–83 years, 11 females) and 15 participants in the Production group (M_{age} = 69 ± 6.9, 60–86 years, 10 females). The sample size of the older adults was determined in accordance with the sample size of the young adults to maintain consistency and comparability across age groups. Table 4 provides a summary of the older adults. The two older adult groups were matched for all variables.

TABLE 4 Experiment 4 participants' characteristics.

	Perception (n = 15, 11 females)				Production (n = 15, 10 females)				
Characteristic	Mean	SD	Min	Max	Mean	SD	Min	Max	t-value
Age	72	6.2	62	83	69	6.9	60	86	1.27
Education (years)	17	3	13	24	17	1.3	14	19	0.48
MoCA score (/30)[a]	29	1.3	26	30	29	1.1	27	30	1.49
Number of spoken languages[b]	1.5	0.64	1	3	2.1	0.8	1	3	2.02
Right ear PTA4[c]	24	9.7	11	40	20	9.8	12	44	1.08
Left ear PTA4[c]	24	11	8.8	52	20	9.2	6.2	38	0.88
Better ear PTA4[d]	21	9.2	8.8	39	18	9	6.2	38	1.05
Interaural difference[e]	4.8	4.2	0	14	4.7	2.8	1.2	8.8	0.13
QuickSIN[f]	0.97	0.91	−0.75	2.2	0.73	1.3	−0.75	3.8	0.58

[a]MoCA = Montreal Cognitive Assessment scale. The MoCA is a short cognitive test that is scored on a 30-point scale. Higher scores indicate better cognitive functions.
[b]Number of spoken languages = number of languages spoken, including native language.
[c]PTA4 = pure-tone average thresholds at 0.5, 1, 2, and 4 kHz for the better ear, measured in decibels (dB hearing level).
[d]Better ear PTA4 = pure-tone average thresholds at 0.5, 1, 2, and 4 kHz for the better ear, measured in decibels (dB hearing level).
[e]Interaural difference = absolute difference between the PTA4 of the left and right ear.
[f]QuickSIN score = QuickSIN measures the ability to perceive sentence in noise. Scores are reported as signal-to-noise ratio (SNR) loss, the difference (in dB SNR) between a listener's performance and that of normal-hearing adult controls. Lower score indicates better speech perception ability.

Procedure

We followed the protocol of Experiment 1, which included auditory stimuli and no delay between the priming and the main tasks.

Statistical analyses

We employed the same analytical approach as in Experiment 1, except that Age group (Young, Older) was included as a between-subject factor in the LMMs. No outliers (values more than three interquartile range) were identified.

RESULTS

Experiment 1

Results of the effects of speech motor priming in young adults are shown in Figure 1. For accuracy, the LMM revealed main effects of Trial type (b = 1.45, SE = 0.53, $t(28)$ = 2.76, p = 0.01, η_p^2 = 0.21) and Group (b = −2.3, SE = 0.65, $t(28)$ = −3.59, p = 0.001, η_p^2 = 0.32), and no significant interaction between the two factors, b = 0.45, SE = 0.53, $t(28)$ = 0.86, p = 0.40, η_p^2 = 0.03. Participants were more accurate for Old (M = 81%) than New pairs (M = 78%). Accuracy was higher in the Production (M = 82%) than in the Perception group (M = 77%). For d', the LMM revealed only a significant effect of Group (b = −0.14, SE = 0.05, $t(28)$ = −2.67, p = 0.01, η_p^2 = 0.2), in which d' was greater in the Production group (M = 1.92) than in the Perception group (M = 1.63). For RT and c, no effects or interactions were significant.

Experiment 2

Results of the effects of visual stimuli in the priming tasks are shown in Figure 2. For accuracy, the LMM revealed main effects of Trial type (b = 1.2, SE = 0.36, $t(56)$ = 3.4, p = 0.001, η_p^2 = 0.17) and Group (b = −2.78, SE = 0.60, $t(56)$ = −4.63, p < 0.001, η_p^2 = 0.28). Additionally, a main effect of Modality was observed, indicating that the visual modality was associated with lower performance compared to the auditory modality. However, this effect was not significant after removing the three outliers. No significant interactions were found with and without outliers. As in Experiment 1, accuracy was higher for Old (M = 79%) than New pairs (M = 77%) and higher in the Production group (M = 81%) than in the Perception group (M = 75%). For d', the LMM revealed only a significant effect of Group (b = −0.19, SE = 0.04, $t(56)$ = −4.6, p < 0.001, η_p^2 = 0.27), in which d' was greater in the Production group (M = 1.86) than in the Perception group (M = 1.48). For c, only the effect of Modality was significant, but this effect was no longer significant after removing the outliers. For RT, no effects or interactions were significant.

Additional analyses were carried out to compare the accuracy and d' of the two groups in the visual modality condition, to ensure that the observed group difference was not driven by the group difference in the auditory modality condition. Analysis on the estimated marginal means showed that the Visual Production group exhibited

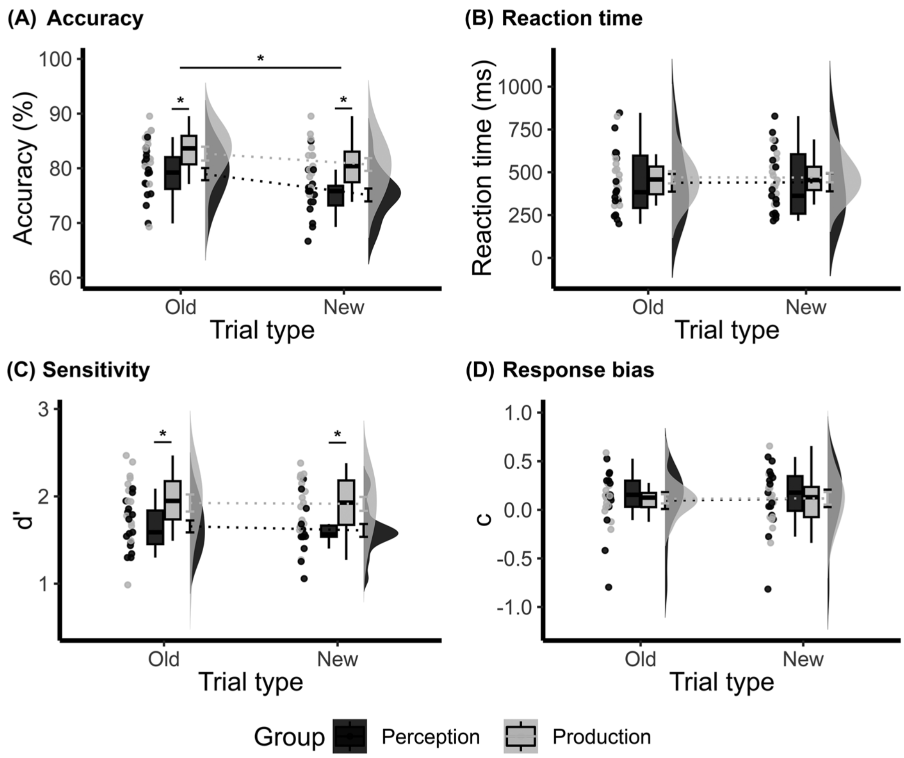

FIGURE 1 Group differences for (A) accuracy, (B) reaction time, (C) sensitivity (d'), and (D) response bias (c) for Experiment 1. For each condition and group, the graphs include a boxplot, individual data points, and a half-violin plot. Error bars represent standard errors of the mean. Dotted lines connect the mean of the two conditions for each group. Asterisks indicate statistical significance ($p < 0.05$).

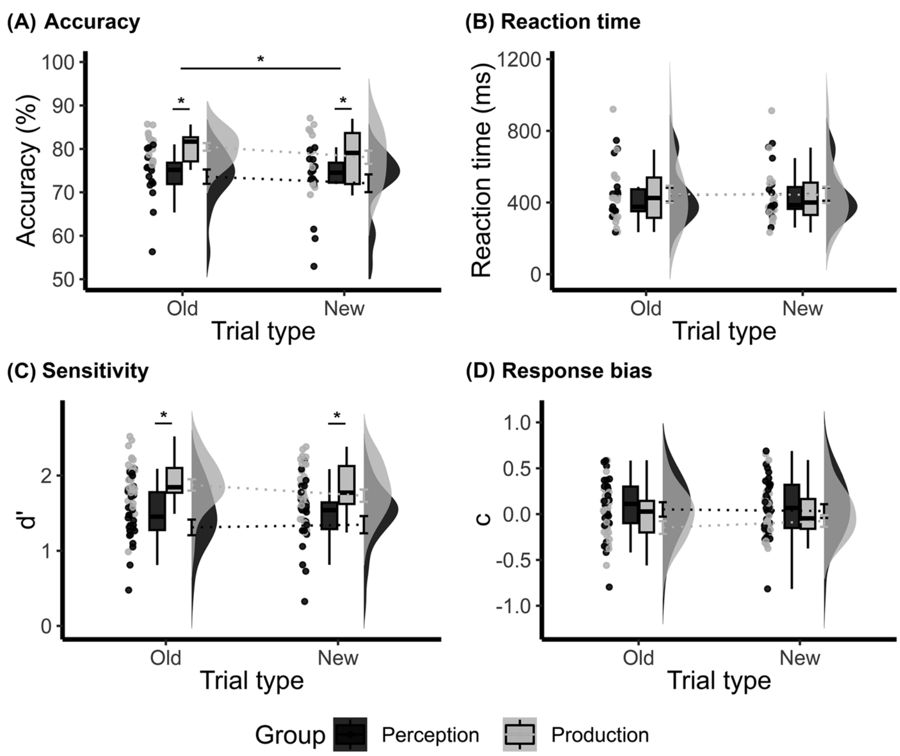

FIGURE 2 Group differences for (A) accuracy, (B) reaction time, (C) sensitivity (d'), and (D) response bias (c) for Experiment 2. Only group comparisons for the visual modality are shown. For each condition and group, the graphs include a boxplot, individual data points, and a half-violin plot. Error bars represent standard errors of the mean. Dotted lines connect the mean of the two conditions for each group. Asterisks indicate statistical significance ($p < 0.05$).

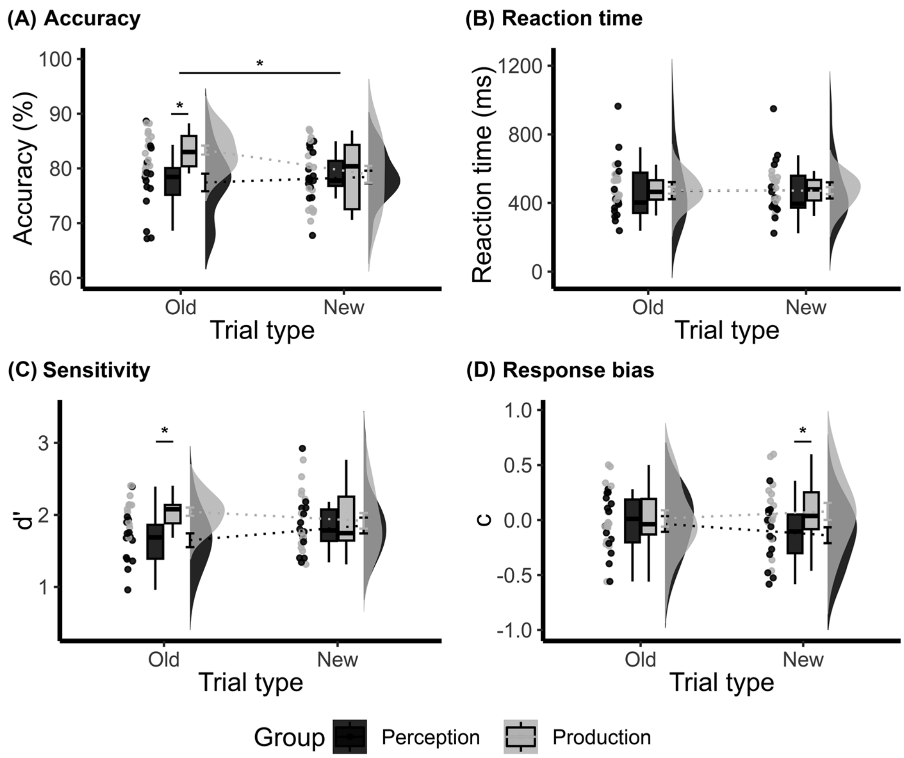

FIGURE 3 Group differences for (A) accuracy, (B) reaction time, (C) sensitivity (d'), and (D) response bias (c) for Experiment 3. Only group comparisons for the 1-h delay condition are shown. For each condition and group, the graphs include a boxplot, individual data points, and a half-violin plot. Error bars represent standard errors of the mean. Dotted lines connect the mean of the two conditions for each group. Asterisks indicate statistical significance ($p < 0.05$).

significantly higher accuracy (M = 79%) than the Visual Perception group (M = 73%) ($t(56) = -3.79$, $p < 0.001$). The Visual Production group also had higher d' ($M = 1.8$) than the Visual Perception group ($M = 1.33$) ($t(56) = -4.02$, $p < 0.001$). Additionally, no difference in accuracy ($t(56) = 1.4$, $p = 0.16$) and d' ($t(56) = 0.98$, $p = 0.33$) was observed between the two Production groups.

Experiment 3

Results of the effects of a 1-h delay between the priming tasks and the SIN task are shown in Figure 3. For accuracy, the LMM revealed main effects of Trial type ($b = 1.15$, $SE = 0.37$, $t(56) = 3.1$, $p = 0.003$, ${\eta_p}^2 = 0.15$) and Group ($b = -1.97$, $SE = 0.5$, $t(56) = -3.9$, $p < 0.001$, ${\eta_p}^2 = 0.21$), as well as a Group × Time × Trial type interaction ($b = 0.91$, $SE = 0.37$, $t(56) = 2.5$, $p = 0.017$, ${\eta_p}^2 = 0.1$). We, therefore, focused on the three-way interaction. Analyses on the estimated marginal means revealed that in the absence of any delay between the priming task and the main task, the Production group exhibited better accuracy for both Trial types (Old: M = 83%; New: M = 81%) in comparison to the Perception group (Old: M = 79%; New: M = 75%) (Old: $t(103) = -2.1$ $p = 0.04$; New: $t(103) = -3.14$, $p = 0.002$). However, when introducing a 1-h delay, the Production group demonstrated better accuracy for Old pairs (M = 83%) compared to the Perception group (M = 77%) ($t(103) = -3.35$, $p = 0.001$). No statistically significant difference in accuracy was found between the Production (M = 79%) and Perception (M = 78%) groups for New pairs ($t(103) = -0.27$, $p = 0.79$). For d', the LMM revealed a significant effect of Group ($b = -0.13$, $SE = 0.04$, $t(55) = -3.4$, $p = 0.001$, ${\eta_p}^2 = 0.18$), as well as a Group × Time × Trial type interaction ($b = 0.5$, $SE = 0.02$, $t(55) = 2.1$, $p = 0.04$, ${\eta_p}^2 = 0.07$). Examination of the estimated marginal means led to the same conclusions as those for accuracy. After the delay, the Production group had better d′ only on Old pairs ($M = 2.05$) compared to the Perception group ($M = 1.65$) ($t(89) = -3.2$, $p = 0.002$). For c, the LMM only showed a significant interaction between the three factors ($b = -0.02$, $SE = 0.01$, $t(55) = -2.1$, $p = 0.04$, ${\eta_p}^2 = 0.07$). After the delay, the Perception group displayed a significantly greater bias toward identifying new syllable pairs as identical ($M = -0.14$) compared to the Production group ($M = 0.08$) ($t(66) = -2.0$, $p = 0.04$), which showed minimal bias. Other comparisons were not significant. For RT, no effects or interactions were significant.

Experiment 4

Results of the effects of speech motor priming in older adults are shown in Figure 4. For accuracy, the LMM revealed a main effect of Age group ($b = 2.6$, $SE = 0.47$, $t(56) = 5.5$, $p < 0.001$, ${\eta_p}^2 = 0.35$), in which older adults (M = 74%) performed lower than young adults (M = 79%). The LMM also revealed main effects of Trial type ($b = 1.2$, $SE = 0.41$,

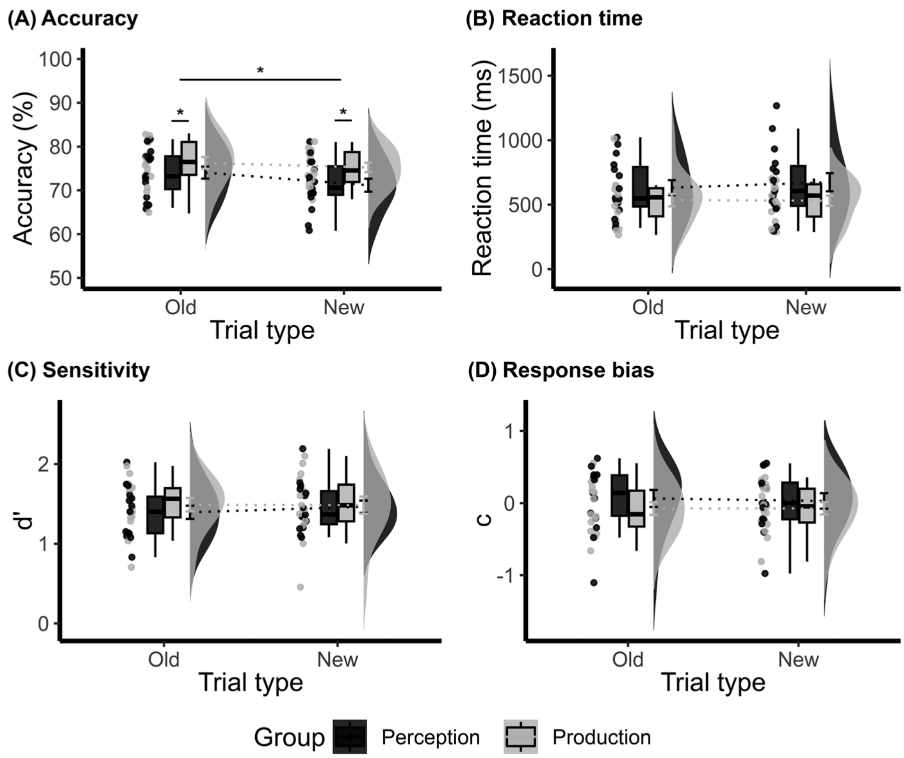

FIGURE 4 Group differences for (A) accuracy, (B) reaction time, (C) sensitivity (d'), and (D) response bias (c) for Experiment 4. Only group comparisons for the older adults are shown. For each condition and group, the graphs include a boxplot, individual data points, and a half-violin plot. Error bars represent standard errors of the mean. Dotted lines connect the mean of the two conditions for each group. Asterisks indicate statistical significance ($p < 0.05$).

$t(56) = 2.9$, $p = 0.005$, $\eta_p^2 = 0.13$) and Group ($b = -1.95$, $SE = 0.47$, $t(56) = -4.1$, $p < 0.001$, $\eta_p^2 = 0.23$). As in Experiments 1 and 2, accuracy for Old pairs ($M = 78\%$) was higher than New pairs ($M = 76\%$), and the Production group had significantly higher accuracy ($M = 79\%$) than the Perception group ($M = 75\%$). No significant interactions with Age group were observed, suggesting that the impact of production on accuracy is consistent across age groups. For d', the LMM revealed a main effect of Age ($b = 0.16$, $SE = 0.04$, $t(56) = 4.1$, $p < 0.001$, $\eta_p^2 = 0.23$) and Group ($b = -0.09$, $SE = 0.04$, $t(56) = -2.2$, $p = 0.03$, $\eta_p^2 = 0.08$). The older adults had lower d' ($M = 1.46$) than the young adults ($M = 1.78$), and the Production group had a higher d' ($M = 1.7$) than the Perception group ($M = 1.5$). No significant interaction with age group was observed, indicating a consistent impact of Production on d' across both age groups. For c and RT, no significant effects or interactions were observed.

Additional analyses were carried out on the estimated marginal means to compare the accuracy and d' of the two groups of older adults to ensure that the observed group difference was not driven by the young adult group. Older adults in the Production group exhibited significantly higher accuracy ($M = 74\%$) than those in the Perception group ($M = 71\%$) ($t(56) = -2.35$, $p = 0.02$). Syllable repetition resulted in an average benefit of 5% for young adults and 3% for older adults. Regarding d', older adults in the Production group had a slightly higher d' ($M = 1.49$) than those in the Perception group ($M = 1.43$). However, this difference was not statistically significant ($t(56) = -0.6$, $p = 0.6$), suggesting that the main effect of Group for d' was likely driven by the group difference in the young adults.

DISCUSSION

Our study builds on previous research that used articulatory suppression to investigate the influence of motor representations on speech perception.[23,30] The general idea is that articulatory suppression uses motor resources and that if motor resources contribute to perception, reducing motor resources should in turn alter perception. Studies have generally shown a marginal reduction in phoneme identification in noise when young adults engage in articulatory suppression. Here, we introduced a paradigm aimed at demonstrating that the production–perception relationship could be leveraged to induce benefits rather than impairments. We hypothesized that if these two processes interact, one can be exploited to enhance the other. Specifically, we investigated whether repeating syllables (production) before discriminating them in noise (perception) could affect perceptual discrimination ability. Our aim was to make motor representations readily available for perception by pre-activating them through syllable repetition.

Speech motor priming increases syllable discrimination in noise

In a series of four experiments, we provide converging evidence that engaging the articulatory motor system through the repetition of syllables improves the ability to discriminate SiN. This benefit in SiN discrimination was observed compared to a control group that performed a phonological priming task requiring same–different judgments using button presses. Interestingly, this priming task was analogous to the SiN task, with the only exception that multi-talker babble was added in the background in the SiN task. Yet the Production group performed better than the Perception group. Engaging the articulatory system was advantageous for subsequently enhancing speech-in-noise perception. The speech motor system is thought to support speech perception by providing top–down information to posterior temporal regions to help disambiguate the noisy or degraded speech signal.[31,32]

The difference in performance between the two groups was not attributed to change in response bias but rather to a difference in sensitivity to phonological details. These results suggest that the motor system is involved in the perceptual processing of speech, especially for phonological discrimination. This is in line with previous neuroimaging studies highlighting the association between white matter structure in the arcuate fasciculus, the scaffolding of the auditory–motor integration network, and sensitivity to phonological details.[33,34] Overall, our results indicate a noteworthy link between perception and production that warrants further exploration, as this link could be leveraged to improve perceptual abilities through motor aspects of speech.

Pooling participants from all four experiments, we identified a 5% improvement in speech perception ability in the Production group compared with the Perception group. This 5% difference is in line with the findings of Stokes et al.,[23] who reported a reduction of around 5% in correct responses in phoneme identification when participants engaged in articulatory suppression. A benefit of 5% could be considered small, but this was observed after only one session of syllable repetition of about 10 min. It is likely that longer and repeated sessions could lead to greater benefits.

Our results also echo a previous rTMS study by Brisson and Tremblay.[15] In their study, increased excitability in the left ventral premotor cortex similarly yielded a 5% enhancement in syllable discrimination ability in noise for both young and older adults. In contrast, the application of rTMS to the left superior temporal sulcus, a region commonly associated with phonological processing,[19] had no discernible effect on syllable discrimination. According to the DIVA (Directions Into Velocities of Articulators) model of speech production,[35,36] the left ventral premotor cortex plays an essential role in housing speech motor representations. The authors therefore suggested that motor representations might have a greater influence in speech-in-noise perception than acoustic/phonological representations. Another interesting observation is the 5% parallel difference observed in our study and their rTMS study. As the paradigm used in the two studies is very similar (i.e., phonological discrimination of syllable pairs in babble noise), this suggests that vocal repetition of syllables prior to a speech perception task could potentially have a comparable effect to the application of excitatory rTMS on the premotor cortex. Although further research is required to directly compare the benefits of both procedures, this has notable clinical implications, as syllable repetition is a much simpler and more practical procedure to administer than rTMS.

Although our study does not directly address the mechanism behind the perceptual benefits of syllable repetition, we can draw parallels with the impact of excitatory rTMS, where brain areas are intentionally stimulated to enhance subsequent processes. Excitatory rTMS is said to stimulate performance via interconnected mechanisms. When magnetic pulses are applied to specific brain regions, rTMS increases neuronal excitability, which in turn promotes increased synaptic transmission and neuronal firing rates. This modulation of synaptic activity can trigger short- or long-term changes in synaptic plasticity, such as long-term potentiation, resulting in improved signal transmission between groups of neurons.[37] Vocal repetition of syllables could exert a similar influence on the motor system. Syllable repetition probably leads to increased motor system activity, which would make it more readily available and adapted to the perceptual processing of speech sounds. Further studies are needed to investigate the neurobiological mechanism by which production influences perception. For instance, studies could measure brain activity during production and assess whether this activity correlates with or explains subsequent perceptual ability.

Near transfer and temporal decay

Another important finding of this study is that when the speech motor priming is conducted immediately before the speech perception task (Experiments 1, 2, and 4), the effect of syllable repetition extended not only to pairs of syllables that were used in the priming tasks (old pairs), but also to new ones that were not presented during the priming task. This suggests that the neural circuits activated during syllable repetition become more sensitive to a broader range of phonologically related stimuli. The DIVA model of speech production[35,36] suggests that speech production begins with the activation of specific cells in the premotor cortex, which project to feedforward articulator velocity maps in bilateral ventral motor cortex. These maps consist of eight antagonistic pairs of cells encoding movement velocities for the articulators, including the upper and lower lips, the jaw, the tongue, and the larynx. Each cell corresponds to specific articulatory gestures for a given speech sound. The concept of distinct neural representations for different articulators gains further support from neuroimaging and rTMS studies showing somatotopy in the motor cortex for speech perception.[8,12] In our study, the syllable pairs differed by only one phoneme, but this difference encompassed diverse articulatory features, likely engaging various articulators during the priming task. This diversity may have primed all articulators, potentially explaining why the priming effect extended to both old and new pairs. Future studies should explore whether priming specific articulators selectively influences the discrimination ability for syllables produced using those articulators. This would support not only our results but also the somatotopy of the motor system for speech perception.

In Experiment 3, the introduction of a 1-h delay between the motor priming and speech perception tasks led to a significant temporal effect. After the delay, the influence of motor priming on discrimination ability was observed exclusively on old pairs, whereas no difference was observed on new pairs. Despite the lack of difference observed in the discrimination of new syllable pairs between groups, the Production group showed significantly less bias than the Perception group in syllable discrimination for these new pairs. The temporal specificity observed suggests that the impact of motor priming decreases over time. Activation patterns induced by motor priming in produced syllables may have a diminishing effect on speech motor patterns associated with non-produced syllables after a 1-h delay. Another explanation could be that engaging in the act of speech production enhances the strength and distinctiveness of the items in memory. As a result, speech-produced items might be better remembered or recognized than non-produced items. This suggests a potential interaction between memory and motor processes in shaping perception, a concept that will be addressed later in the discussion.

Influence of speech motor priming in aging

In Experiment 4, we examined the impact of motor speech priming in older adults. This investigation is important for this demographic because of the frequent difficulties they encounter in perceiving speech-in-noise. Currently, there is a notable absence of evidence-based clinical interventions designed to address this problem.

Here, we observed a consistent influence of syllable repetition across age groups, as evidenced by the absence of any significant Group × Age group interaction for both accuracy and d'. This suggests that speech motor priming remains influential in speech perception in older adults. Our results align with the work of Brisson and Tremblay,[15] where rTMS applied to the ventral premotor cortex similarly improved speech-in-noise perception in young and older adults. Overall, our results are consistent with the Motor Compensation Hypothesis. However, it is important to note that, although we found no significant Group × Age group interaction, we did observe an average improvement of 5% in young adults and 3% in older adults, suggesting a weaker effect in older adults than in young adults. Relatedly, the Group effect for d' in Experiment 4 was likely due to a group difference in young adults and not in older adults. Our results partly support the idea that speech motor priming loses its influence on speech perception during the aging process, which is consistent with the Auditory–Motor Decline Hypothesis.

Overall, our results can be interpreted within the framework of both hypotheses. One possible explanation is that there is greater variability in older adults. The influence of the speech motor system on speech perception may diminish with age, leading to perceptual deficits. However, older people who maintain a higher level of motor activity during perception may experience less decline. Another explanation could be that the influence of motor priming on perception remains the same during aging, but that the effect of speech motor priming decays more rapidly in older adults than in younger adults. This idea is consistent with the results of Experiment 3, which show a temporal decay of the effect in young adults. Nevertheless, our observation that syllable repetition in older adults confers perceptual benefits is particularly encouraging. This suggests that it may be possible to harness the relationship between production and perception to improve auditory communication skills in the aging population. However, it is important to acknowledge that the observed effects in older adults are small. Additionally, considering the results of Experiment 3, which indicate a decline in benefits over time in young adults, it is likely that the effects of speech motor priming disappear in older adults after 1 h. However, we believe that these effects may be enhanced with repeated sessions and longer priming periods (the priming task lasted only 10 min in this study, which was not designed as a clinical intervention study but rather as a proof of concept). Further studies should investigate the impact of multiple longer priming sessions as part of a longitudinal study design. Interestingly, a growing body of evidence supports the idea that singing, which involves vocal motor training and vocal motor engagement over a longer period of time, can improve speech-in-noise ability in aging.[38,39] This additional support reinforces our idea, highlighting the potential efficacy of interventions incorporating motor aspects of speech to improve auditory communication skills throughout the lifespan.

Other potential mechanisms by which production could influence perception

Here, we postulated that syllable repetition acts as a primer for the motor system—facilitating greater accessibility to articulatory representations for perception. However, it is important to recognize the existence of alternative mechanisms that could also contribute to the influence of production on perception.

In the context of repeating syllables, a significant challenge arises when trying to differentiate the impact of articulation from the influence of one's auditory feedback. Speech production models propose that during speaking, the motor cortex sends efferent copies of motor commands to the auditory cortex, allowing for a comparison between expected and actual sound outputs. When errors are detected, corrective motor commands are initiated. Numerous studies have shown that speech production is associated with activity in the posterior temporal areas, whether production is silent (covert) or audible (overt).[40,41] The involvement of posterior temporal regions in production suggests that speech production may contribute to the priming of acoustic and phonological representations, which may also explain the advantages observed in the current study. We can draw parallels with previous studies that have addressed the question of whether the perceptual changes associated with speech motor learning are due to the motor aspects of speech or to the sensory information received during motor learning. In the study by Lametti et al.,[42] participants undertook a speech motor learning task adapting to altered auditory feedback and then undergoing assessments of perceptual changes. The authors noted that perceptual changes resulting from speech motor learning aligned with the phonetic range of the adapted speech production and not that of the altered auditory feedback. Changes in speech perception appear to be primarily associated with motor changes rather than auditory feedback. Building on the work of Lametti et al., here we suggest a similar mechanism. Syllable repetition probably improves perception by pre-activating motor representations rather than relying solely on hearing one's own voice. To distinguish the effects of production and auditory feedback using our paradigm, future studies could compare silent versus vocalized articulation or altered versus unaltered auditory feedback during priming.

Another mechanism that might explain the observed benefits associated with speech production can be derived from findings in the memory literature. A body of research indicates that reading words aloud significantly improves memory compared to silent reading.[43,44] Two hypotheses have been put forward to explain this phenomenon. One suggests that producing items increases their distinctiveness in memory, whereas the other postulates that it increases the strength of items in memory (i.e., familiarity). Although these two hypotheses can explain the results of Experiment 3, in which participants in the Production group performed better on old pairs, they cannot explain why production leads to better discrimination on new and old pairs in the other three experiments (i.e., when the priming is done immediately before the speech perception task). Furthermore, we observed that old pairs were better discriminated than new pairs in both groups, suggesting that syllable production and discrimination lead to a trace of these syllables in memory to the same degree. In a recent study investigating the neural mechanisms associated with the production effect in memory, researchers observed that reading aloud is linked to increased activity in inferior frontal motor areas, including the premotor and motor cortices.[45] This increased activity in motor areas predicted memory performance, suggesting a potential interaction between motor and memory processes.

Finally, another possible cognitive explanation is that participants assigned to the speech production task may have been intrinsically more engaged than those tasked with making perceptual judgments. It is plausible that the act of speech production is more mentally stimulating than active listening. Therefore, this lower level of engagement in the perceptual group could potentially have affected their attentional state and performance during the main task, resulting in poorer performance across all syllable pairs. However, it should be noted that the priming discrimination task required concentration and effort to discriminate syllables, especially as the difference between syllables could not be predicted. Therefore, although it is unlikely that the observed effects of production can be attributed solely to attention and level of engagement, further research is required.

Limitations and future directions

We replicated our results in four separate experiments, reinforcing the robustness of our findings. However, it is important to recognize the main limitation of our study, which lies in its between-subjects design and the relatively small sample sizes within each group. To address this, future studies could adopt a within-subjects design, ensuring that each participant experiences all conditions, thereby increasing statistical power. Furthermore, the group interaction between young and older adults could become significant if the sample size of young and older adults were larger. It is also worth noting that we did not measure baseline performance in the syllable discrimination task, which raises the possibility that our participant groups may not have been perfectly matched in this regard. The integration of pre- and post-priming test sessions to measure performance before and after the priming task could alleviate this limitation. Moving forward, it would be useful for future projects to incorporate a control priming condition devoid of speech stimuli, such as a cognitive priming task. This would enable us to better understand the impact of cognition on the observed benefits and to determine whether our auditory discrimination task has any advantages over a speech-free control task. In addition, a longitudinal study involving several sessions of speech motor priming would enable us to investigate whether syllable repetition leads to sustained improvements in speech perception. It would also be interesting to compare speech motor priming training with other methods aimed at improving speech-in-noise ability, such as musical, cognitive, and auditory training. Nevertheless, our results suggest a significant interaction between speech production and perception that merits further exploration.

Another limitation of the study is that we did not assess the far transfer of production on speech perception. In theory, priming the motor system could improve speech perception under various listening conditions. Although this was not the focus of the current study, we have a first indication that there may be a transfer effect to other tasks. In Experiment 3, participants were subjected to a 1-h waiting period between the priming task and the main task. During this interval, the hearing assessment was incorporated, leading participants to engage in the QuickSIN immediately after completing the priming task. We observed a difference between the two groups in the QuickSIN test, with the Production group having better sentence-in-noise comprehension. This group difference could be due to preexisting group differences or to the priming task, such that the effect of syllable repetition might not be specific to syllables but also to sentences. The priming difference on the QuickSIN could only be observed in Experiment 3, as the auditory assessment was performed prior to the priming task in all other experiments. Future research could involve priming the motor system and exploring its transfer effects in various speech tasks encompassing diverse speech stimuli, covering both lexical and sub-lexical levels.

Finally, we have only investigated the influence of production on speech perception in noise. It would be relevant for future studies to explore the influence of production on speech perception in quiet conditions to determine whether priming the motor system can affect speech perception in favorable circumstances or whether it provides support only in difficult listening situations.

CONCLUSION

In conclusion, our study highlights the importance of investigating the relationship between speech perception and production. Using a novel behavioral paradigm, we demonstrated the efficacy of speech motor priming in improving syllable discrimination in noise with benefits lasting for at least 1 h. The consistent benefits observed across different modalities (auditory or visual) and age groups, particularly in older adults, emphasize the potential for clinical applications. This research opens avenues for further exploration of the interplay between speech perception and production and offers valuable insights into improving speech discrimination, particularly in challenging acoustic environments.

AUTHOR CONTRIBUTIONS

M.P.: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; visualization; writing—original draft preparation. Q.L.: Investigation; writing—review and editing. P.T.: Conceptualization; methodology; writing—review and editing. C.A.: Conceptualization; funding acquisition; methodology; project administration; resources; supervision; writing—review and editing.

ACKNOWLEDGMENTS

We thank all the participants for their contributions. We also thank Signe Lund Mathiesen for her valuable help in creating the graphical abstract. This research was supported by grants to C.A. from the William Demant Foundation and from the Natural Sciences and Engineering Research Council of Canada (NSERC). M.P. was funded by a Canadian Institutes of Health Research (CIHR) graduate scholarship.

CONFLICT OF INTEREST STATEMENT

The authors have declared no conflicts of interest.

ORCID

Maxime Perron https://orcid.org/0000-0001-7015-5858
Pascale Tremblay https://orcid.org/0000-0001-7161-9255
Claude Alain https://orcid.org/0000-0003-4459-1538

REFERENCES

1. McGettigan, C., & Tremblay, P. (2018). Links between perception and production: Examining the roles of motor and premotor cortices in understanding speech. In S.-A. Rueschemeyer & M. G. Gaskell (Eds.), *The Oxford Handbook of Psycholinguistics* (2nd ed., pp. 306–332). Oxford University Press.
2. Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461.
3. Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.
4. Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology*, *55*, 149–179.
5. Tremblay, P., & Small, S. L. (2011). On the context-dependent nature of the contribution of the ventral premotor cortex to speech perception. *Neuroimage*, *57*, 1561–1571.
6. Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*, 701–702.
7. Adank, P. (2012). The neural bases of difficult speech comprehension and speech production: Two Activation Likelihood Estimation (ALE) meta-analyses. *Brain and Language*, *122*, 42–54.
8. Pulvermüller, F., Huss, M., Kherif, F., Moscoso Del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 7865–7870.
9. Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nature Communications*, *7*, 12241.
10. Wong, P. C. M., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., & Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia*, *47*, 693–703.
11. Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, *17*, 1692–1696.
12. D'ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, *19*, 381–385.
13. Sato, M., Tremblay, P., & Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language*, *111*, 1–7.
14. Schomers, M. R., Kirilina, E., Weigand, A., Bajbouj, M., & Pulvermüller, F. (2015). Causal influence of articulatory motor cortex on comprehending single spoken words: TMS evidence. *Cerebral Cortex*, *25*, 3894–3902.
15. Brisson, V., & Tremblay, P. (2021). Improving speech perception in noise in young and older adults using transcranial magnetic stimulation. *Brain and Language*, *222*, 105009.
16. Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, *11*, 351–360.
17. Scott, S. K., Mcgettigan, C., & Eisner, F. (2009). A little more conversation, a little less action—Candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience*, *10*, 295–302.
18. Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, *13*, 110–114.
19. Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.
20. Hickok, G., Costanzo, M., Capasso, R., & Miceli, G. (2011). The role of Broca's area in speech perception: Evidence from aphasia revisited. *Brain and Language*, *119*, 214–220.
21. Hickok, G. (2010). The role of mirror neurons in speech perception and action word semantics. *Language and Cognitive Processes*, *25*, 749–776.
22. Wu, Z.-M., Chen, M.-L., Wu, X.-H., & Li, L. (2014). Interaction between auditory and motor systems in speech perception. *Neuroscience Bulletin*, *30*, 490–496.
23. Stokes, R. C., Venezia, J. H., & Hickok, G. (2019). The motor system's [modest] contribution to speech perception. *Psychonomic Bulletin & Review*, *26*, 1354–1366.
24. Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, *116*, 2395–2405.
25. Tucker, B. V., Brenner, D., Danielson, D. K., Kelley, M. C., Nenadić, F., & Sims, M. (2019). The massive auditory lexical decision (MALD) database. *Behavior Research Methods*, *51*, 1187–1204.
26. Hautus, M., Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide*. Routledge.
27. R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

28. RStudio Team. (2022). *RStudio: Integrated development environment for R*. RStudio, PBC.
29. Panouillères, M. T. N., & Möttönen, R. (2018). Decline of auditory-motor speech processing in older adults with hearing loss. *Neurobiology of Aging*, *72*, 89–97.
30. Slade, K., Beat, A., Taylor, J., Plack, C. J., & Nuttall, H. E. (2023). The effect of motor resource suppression on speech perception in noise in younger and older listeners: An online study. *Psychonomic Bulletin & Review*, *31*(1), 389–400.
31. Osnes, B., Hugdahl, K., & Specht, K. (2011). Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *Neuroimage*, *54*, 2437–2445.
32. Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *PNAS*, *111*, 7126–7131.
33. Tremblay, P., Perron, M., Deschamps, I., Kennedy-Higgins, D., Houde, J.-C., Dick, A. S., & Descoteaux, M. (2019). The role of the arcuate and middle longitudinal fasciculi in speech perception in noise in adulthood. *Human Brain Mapping*, *40*, 226–241.
34. Perron, M., Theaud, G., Descoteaux, M., & Tremblay, P. (2021). The frontotemporal organization of the arcuate fasciculus and its relationship with speech perception in young and older amateur singers and non-singers. *Human Brain Mapping*, *42*, 3058–3076.
35. Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language & Cognitive Processes*, *26*, 952–981.
36. Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, *25*, 408–422.
37. Jannati, A., Oberman, L. M., Rotenberg, A., & Pascual-Leone, A. (2023). Assessing the mechanisms of brain plasticity by transcranial magnetic stimulation. *Neuropsychopharmacology*, *48*, 191–208.
38. Perron, M., Vaillancourt, J., & Tremblay, P. (2022). Amateur singing benefits speech perception in aging under certain conditions of practice: Behavioural and neurobiological mechanisms. *Brain Structure and Function*, *227*, 943–962.
39. Dubinsky, E., Wood, E. A., Nespoli, G., & Russo, F. A. (2019). Short-term choir singing supports speech-in-noise perception and neural pitch strength in older adults with age-related hearing loss. *Frontiers in Neuroscience*, *13*, 1153.
40. Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*, 673–682.
41. Okada, K., & Hickok, G. (2006). Left posterior auditory-related cortices participate both in speech perception and speech production: Neural overlap revealed by fMRI. *Brain and Language*, *98*, 112–117.
42. Lametti, D. R., Rochet-Capellan, A., Neufeld, E., Shiller, D. M., & Ostry, D. J. (2014). Plasticity in the human speech motor system drives changes in speech perception. *Journal of Neuroscience*, *34*, 10339–10346.
43. Macleod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*, 671–685.
44. Forrin, N. D., Macleod, C. M., & Ozubko, J. D. (2012). Widening the boundaries of the production effect. *Memory & Cognition*, *40*, 1046–1055.
45. Bailey, L. M., Bodner, G. E., Matheson, H. E., Stewart, B. M., Roddick, K., O'neil, K., Simmons, M., Lambert, A. M., Krigolson, O. E., Newman, A. J., & Fawcett, J. M. (2021). Neural correlates of the production effect: An fMRI study. *Brain and Cognition*, *152*, 105757.

How to cite this article: Perron, M., Liu, Q., Tremblay, P., & Alain, C. (2024). Enhancing speech perception in noise through articulation. *Ann NY Acad Sci.*, 1–15. https://doi.org/10.1111/nyas.15179