

Phonological working memory but not beat gestures affect L2 word stress perception

Luise Böttcher, Marieke Hoetjes, and Lieke van Maastricht
Radboud University

Acquiring second language (L2) prosody, specifically stress, is challenging for learners. While beat gestures reinforce prosodic cues in a native language (L1), their role in L2 stress perception and its interaction with cognitive individual differences is underexplored. This study investigates if beat gestures facilitate L2 stress perception and how this relates to learners' visuospatial, phonological, and executive working memory (WM). Twenty-four native Dutch speakers completed an L2 Stress Perception Task. Mixed-effects modeling revealed that participants with better phonological WM showed higher stress identification accuracy, corroborating its importance in L2 acquisition. However, no benefit of beat gestures was found for either stress identification accuracy or reaction times. Executive and visuospatial WM showed no main effects, nor did any of the WM components interact with the gesture condition. Our findings highlight the importance of incorporating cognitive individual difference measures, as they may have a greater impact than the gesture effect itself.

Keywords: beat gesture, working memory, lexical stress, second language, individual differences

Introduction

Beat gestures are related to prosodic prominence and impact stress perception in a native language (L1; Bosker & Peeters, 2021; Bujok et al., 2024; Krahmer & Swerts, 2007; Rohrer et al., 2024). Determining which syllable in a word is most prominent is essential for speech comprehension (Jesse & McQueen, 2014; Reinisch et al., 2010) and particularly relevant for learners of Spanish, where words can have different meanings solely based on their stress placement (e.g., 'aCEPto' means *I accept*, while 'acePTÓ' means *(s)he accepted*, capital letters indicating lexical stress). Using beat gestures is therefore a plausible approach to support learners

in the word stress of a second language (L2).¹ However, previous studies on the effect of beat gestures have focused on production alone and yielded inconclusive results (Gluhareva & Prieto, 2017; Llanes-Coromina et al., 2018; van Maastricht et al., 2019a; van Maastricht et al., 2022). Research on beat gestures in L2 stress *perception* is currently lacking and relevant for this reason alone. But it is also relevant because it might help us understand previous production findings, as correct perception may be a prerequisite for correct production. Furthermore, cognitive individual differences in working memory (WM) influence L2 learning, yet research disentangling the differential roles of WM components in multimodal L2 is still scarce. Our study aims to advance our understanding of how language learners process speech and co-speech gestures in an L2 by focusing not only on effects that occur or do not occur, but by providing insights into how effects differ by including three different WM measures.

Word stress

The term prosody refers to those sounds of a language that supersede the phoneme level, such as rhythm, phrasing, intonation, and stress, which can be cued by several acoustic correlates, such as duration, amplitude, or fundamental frequency (Arvaniti, 2020). Prosody is a determining factor in both the form and content of communicative messages. For example, in English, phrasal intonation can distinguish a question ('Nijmegen is in the Netherlands?') from a statement ('Nijmegen is in the Netherlands.') by using a rising pitch instead of a falling pitch, or highlight a contrastive word by stressing it ('Nijmegen is in the Netherlands, but LEIPZIG is not'). Within a word, syllables can stand out relative to their neighbors, aiding L1 listeners in processing the message (faster) (Reinisch et al., 2010). However, correctly perceiving and producing the prosody of an L2 is typically challenging for learners. They struggle to make the right syllables stand out in a word or to determine which syllables were stressed in a word they heard (e.g., van Maastricht et al., 2016b, 2016a, 2019b, 2021). In languages with variable stress, where a word's lexical meaning can differ depending solely on the location of the stressed syllable, perceiving word stress is essential for speech comprehension, as is the case in, for example English ('CONtract', *a written/spoken agreement* vs. 'conTRACT', *to decrease in size or number*) (Zhu & Mok, 2022, p.203) and Spanish ('HAblo', *I speak* vs. 'haBLÓ', *(s)he spoke*) (van Maastricht et al., 2021, p.82).

1. Throughout this article, any foreign or additional languages are abbreviated as L2, irrespective of the order in which they were acquired after the L1.

Beat gestures

Aside from acoustic correlates, lexical stress can also be cued manually. *Beat gestures*, quick biphasic up and down or back and forth hand movements, increase the perceived prominence of co-occurring words (Krahmer & Swerts, 2007; Leonard & Cummins, 2011; McNeill, 1992) and serve to highlight stressed syllables in the L1. Bosker and Peeters (2021) and Rohrer et al. (2024) respectively tested Dutch and Spanish natives in a Two-Alternative Forced Choice Task with disyllabic lexical stress minimal pairs (e.g., ‘PLAto’ vs. ‘plaTEAU’ in Dutch or ‘BAIllo’ vs. ‘baiLÓ’ in Spanish). They manipulated the acoustic stress correlates to create a continuum ranging from clear stress on the first syllable to clear stress on the final syllable. Participants saw a speaker produce the target word in a sentence while they produced a beat gesture aligned to the first or second syllable. Participants in both studies were biased to perceive stress on the first syllable when they saw a beat gesture on that syllable and vice-versa, showing that beat gestures affect word stress perception in the L1.

Other relevant work, on the use of beat gestures in teaching L2 prosody, showed that both observing (Gluhareva & Prieto, 2017) and producing (Llanes-Coromina et al., 2018) beat gestures during rhythmic training were associated with better general measures of L2 pronunciation when compared to control groups without beat gestures. The measures in both studies evaluated pronunciation through holistic ratings, for example, of accentedness or comprehensibility. The prosody production implicitly plays a role in these ratings, but only two studies to date have focused specifically on L2 word stress production: Both van Maastricht et al. (2019a) and van Maastricht et al. (2022) presented learners with a short L2 stress training in different gesture conditions (metaphoric or beat gestures, gesture observation or production) and a control group. Participants in all groups improved their stress production in Dutch-Spanish cognates after training, irrespective of the gesture condition. Unlike Gluhareva and Prieto (2017) and Llanes-Coromina et al. (2018), these studies therefore show no evidence that beat gestures were helpful for L2 stress production.

While studies on beat gestures are inconclusive, it appears that hand-clapping training does improve the pronunciation of L2 words. Like beat gestures, hand-clapping can highlight prosodic features in L2 teaching. Zhang et al. (2020) and Baills and Prieto (2021) found that both Chinese adolescents and Catalan children learning French, who observed and performed hand-clapping during training improved significantly more than the controls in target-like final vowel lengthening, an acoustic correlate of word stress in French. This result was partially replicated by Baills et al. (2018) and Baills and Prieto (2021), who reported accentedness ratings rather than acoustic measures and found that observing and

imitating hand-clapping significantly benefited Catalan but not the Chinese children.

Taken together, these findings indicate that beat gestures can help some learners improve overall pronunciation and that hand-clapping, which is similar to beat gestures in some respects, supports target-like stress production. Nevertheless, research using beat gestures for L2 stress has not shown clear results yet. Knowing whether the prosodic highlighting of word stress using gestures helps learners notice stress placement in perception could explain these conflicting results. Moreover, some studies have shown that the benefits of gestures might not be the same for all learners.

Working memory

One of the individual factors that is known to affect both speech and gesture perception is WM (Wen et al., 2020; Özer & Göksun, 2020). WM refers to the cognitive system that we use to remember information over a short time to either keep information active or to mentally manipulate it, for example, when translating a message from your L1 to your L2. To do this, you must remember the original utterance while retrieving the words in the other language, then combine them into a correct L2 sentence while not letting surrounding noise distract you. WM is one of the main cognitive functions determining how well individuals learn an L2 (Dörnyei & Skehan, 2008; Hummel, 2009; Jackson, 2020; Juffs & Harrington, 2011; Li, 2019; Linck et al., 2014; Skehan, 1989; Wen & Li, 2019).

The Phonological/Executive model

While the traditional model of WM by Baddeley and Hitch (1974) and variations thereof are widely used, the *Phonological/Executive (P/E) Model* (Wen, 2019, 2023) has been developed for language acquisition research specifically. The *P/E Model* contains different levels, which describe WM components, finer-grained functions related to the components, the involved portion of long-term memory (LTM), and real-life skill and domain correspondences, see Figure 1.

Phonological working memory (PWM) in the *P/E Model* has two functions: Phonological short-term storage and articulatory rehearsal. According to the *P/E Model*, PWM provides the basis for the parts of language (learning) that are based on sound and chunking, underlying the acquisition and development of lexical, collocational, grammatical, and discourse knowledge. Therefore, PWM is also called a *language acquisition device* (Wen et al., 2019, p. 202). While acknowledged as a component of WM (e.g., in Baddeley and Hitch's model from 1974), visuospatial working memory (VWM), performing the same function as PWM but for visual input, is not yet included in the *P/E Model* due to the lack of L2 research.

The executive working memory (EWM) component in the *P/E Model* comprises three domain-general executive functions, namely memory updating, task-switching, and inhibitory control and is responsible for regulating and coordinating attention and control processes in L2 comprehension and production, sub-skill learning, and interactions. Wen argues that, therefore, EWM is best described as a *language processing device* (Wen, 2019, pp.199–200).

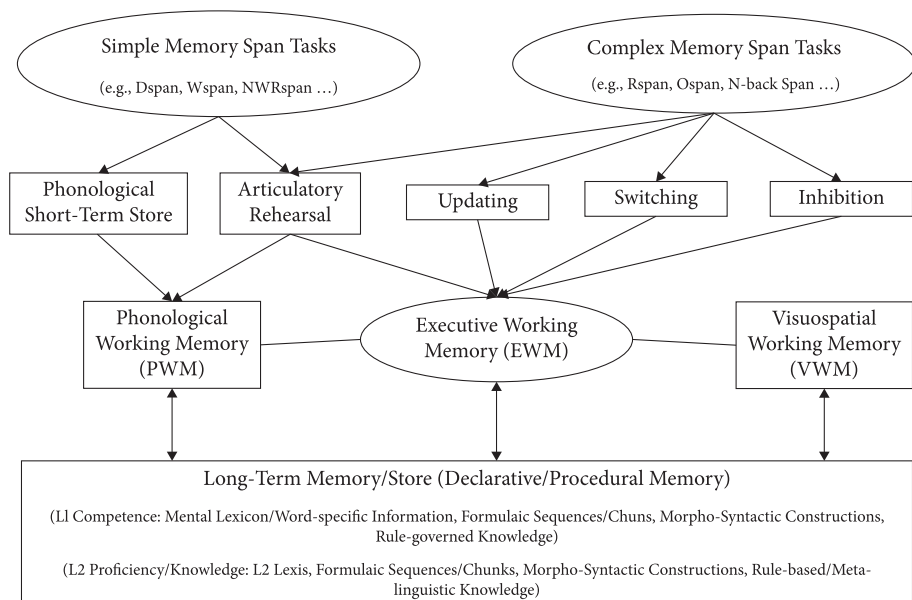


Figure 1. Phonological/Executive model (adapted from Wen et al., 2019)

Working memory in multimodal and L2 stress processing

Although there is previous work (e.g., Ping & Goldin-Meadow, 2010) showing that gestures may lighten the cognitive load, there is also reason to suggest that visual and verbal information may compete for attentional resources in multimodal comprehension (Hoetjes & van Maastricht, 2020; Kelly & Lee, 2012). Retaining and processing gesture and speech is essential for integrating them later. Thus, speakers with strong cognitive resources, either in the form of PWM, VWM, or EWM, have an advantage in language perception. Studies using picture probe classification tasks with congruent and incongruent iconic gestures demonstrated the crucial role of visuospatial resources for benefiting from gestures: VWM but not EWM nor PWM was related to congruency effects in general and under high VWM load (Wu & Coulson, 2014), and congruency effects were reduced under VWM but not PWM load (Wu et al., 2022). For example, Özer and

Göksun (2020) used a mismatch paradigm with a video action prime (video of a daily action) and a target video with speech and gesture. In incongruent trials, either speech or gesture was not related to the action prime, creating trials with high visual load (gesture unrelated to action prime) and high verbal load (speech unrelated to action prime). Higher VWM was related to better accuracy in both types of unrelated trials, but especially under high visual load (gesture unrelated), while higher EWM increased accuracy only in the high verbal load condition (speech unrelated). The authors suggest that different WM components underlie gesture and speech processing and that some people might benefit more from gestures than others. Age-related EWM-decline may explain why older adults benefit less from gestures in multimodal comprehension (Cocks et al., 2011; Schubotz et al., 2021; Thompson, 1995; Thompson & Guzman, 1999): Older participants' weaker EWM might be related to insufficient resources for speech-gesture integration, especially in demanding situations, like with background noise.

Research on beat gestures and L2 learning that includes WM measures is still scarce and has not yet disentangled the roles of PWM, VWM, and EWM. Van Maastricht et al. (2022) included individual EWM differences in a study on the effect of gestures on L2 stress production, but did not find a significant main effect of EWM, nor an interaction of EWM with gesture condition, likely due to low statistical power. Studies on rhythmic pronunciation training with hand-clapping only included PWM: Baills et al. (2018) found a significant effect of PWM on general accentedness in the Chinese adolescents trained on French words but not the Catalan children, both in the hand-clapping group and in the control group. However, Zhang et al. (2020) and Baills and Prieto (2021) found no significant effect of PWM on final syllable durations (the acoustic correlate of word stress) for the same participants. The different results for Chinese and Catalan participants may be explained by the overall task-related cognitive load related to the cognate status of the words for Catalan but not Chinese children.

PWM in L2 word stress is understudied, but Stepanov et al. (2020) reported positive correlations of L2 prosodic discrimination with both PWM and EWM measures in mono- and bilingual children at a general sentence level. EWM benefits were also found in an ERP study by Schwab et al. (2020) for reducing stress deafness in training and in a behavioral study by Darcy et al. (2015) for L2 word stress repetition. Better EWM may allow learners to encode and retain stress information more accurately. However, studies on particular functions of EWM did not find a significant relation with performance in L2 stress tasks (Honbolygó et al., 2019; Mora & Darcy, 2023). These inconsistent findings may be due to the use of more specific EWM tasks, but may also be related to the L2 tasks used, or participants' characteristics.

Taken together, research indicates that the three WM components play differential roles in language processing and language learning. VWM and EWM seem important for benefiting from visual input in speech comprehension, while PWM and EWM appear crucial for speech processing. The specific contributions of distinct WM components – especially to multimodal L2 stress processing – remain unclear and warrant further investigation.

The present study

In the current study, Dutch native speakers are asked to identify the stressed syllable in Spanish words which either have a beat gesture aligned to the stressed syllable (Gesture condition, GC), or not (No-Gesture condition, NGC). We aim to answer the following two research questions:

1. Does perceiving beat gestures benefit L2 word stress perception?
2. Does the effect of perceiving beat gestures on L2 word stress perception depend on PWM, VWM, and/or EWM?

We expect higher accuracy in stress detection in the GC because L1 research has shown that syllables with beat gestures are more likely to be perceived as stressed (Bosker & Peeters, 2021; Bujok et al., 2024; Rohrer et al., 2024). In addition, based on previous research (Bujok et al., 2025), we expect faster reaction times (RTs) in the GC because the additional information provided by beat gestures might facilitate and therefore speed up the decision. We expect better and faster performance of participants with better PWM in both conditions, as PWM is relevant for the short-term storage of speech signals (Wen, 2019; Wen, 2023) and has been shown to determine sensibility to the speech input for multimodal integration with iconic gestures (Özer & Göksun, 2020; Wu & Coulson, 2014). We expect participants with better VWM to be more accurate and faster in the GC, because VWM is essential for benefiting from co-speech gestures (Özer & Göksun, 2020; Wu & Coulson, 2014; Wu et al., 2022). Finally, given its role in language processing (Wen, 2019, 2023), gesture-speech integration (Özer & Göksun, 2020; Schubotz et al., 2021), and L2 stress processing (Darcy et al., 2015; Schwab et al., 2020), we expect participants with better EWM to perform better and faster in both the GC and NGC.

Methods

All supplementary online materials about the data and experiment can be found through this link: <https://doi.org/10.34973/kebe-4n7o>.

Participants

Dutch natives without previous experience with Spanish were recruited through personal contacts and flyers at Radboud University and through its research participation system. Only individuals with no or corrected hearing or vision impairments and without language impairments performed the experiment. The study was approved by the Ethics Assessment Committee of the Faculty of Arts (2024–6975). Twenty-four participants were tested, one of which one had to be excluded due to technical problems while testing. The final sample contained twenty-three individuals (13 female, 9 male, 1 other), aged between 19 and 70 years ($M=29.3$, $Mdn=25$, $SD=13.53$). All were native speakers of Dutch, but six participants (26.09%) reported having at least one more native language. Other native languages were: Chinese, Croatian, Czech, English, Limburgish, Luxembourgish, and Russian. On average, participants reported five languages that they have current or had previous knowledge of ($Min=3$, $Max=7$, $Mdn=4$). Among the participants, 14 (60.89%) indicated that they have played or are currently playing an instrument, including singing. Neither additional (native) languages nor musical activity had a significant influence on L2 word stress perception (see Section Procedure: Mixed-effects modeling).

Materials

L2 Stress perception task

Task design

A dual-task paradigm was used, see Figure 2. The primary task was to watch a video clip of a Spanish native articulating a Spanish word and to indicate, after the secondary task, which syllable carried the word stress. The secondary task was to decide whether simultaneously presented animal pictures and sounds matched or not. This distraction task was included to introduce a period of 10 seconds during which the auditory and visual input from the primary task had to be retained in WM while inhibiting interfering information in both modalities. Upon returning to the primary task, participants had maximally 5 seconds to indicate which syllable was stressed. Otherwise, they would get feedback that they were too slow, and the next trial would start. Wu and Coulson (2014) and Wu et al. (2022) previously used similar dual-task designs when investigating the differential effects of WM. Responses to both tasks were given by button presses with the dominant hand on a button box. Participants completed two blocks of 40 trials each. In between blocks, the WM tests were administered.

Stimulus words

Twenty trisyllabic Spanish verbs with a lexical stress contrast between the first person singular in present tense ('preSENto', *I present*) and the third person singular in the simple past tense ('presenTÓ', *(s)he presented*) were used, resulting in 40 distinct verb forms. To avoid confounding the effect of the presence of a beat gesture with the effects of both subject and verb meaning, a fully balanced experimental design was used in which each verb form was presented to the participants twice, once with a beat gesture time-aligned with the stressed syllable and once without, leading to a total of 80 trials, divided into two lists. If a list contained the first-person form (2nd syllable stressed) with a gesture, then the third-person form (3rd syllable stressed) was presented without a gesture in this list, and vice versa. The verb forms paired with the gestures in list 1 appeared without the gestures in list 2, and vice versa. The order of the verb forms was randomized. To control for order or learning effects, half of the participants received list 1 first, while the other half received list 2 first. This way, participants saw all possible verb forms, in both subject conjugations, with and without a gesture in a random order. Using Spanish Dutch cognates allowed us to test beginning learners for their stress perception without taxing cognitive resources for meaning processing.

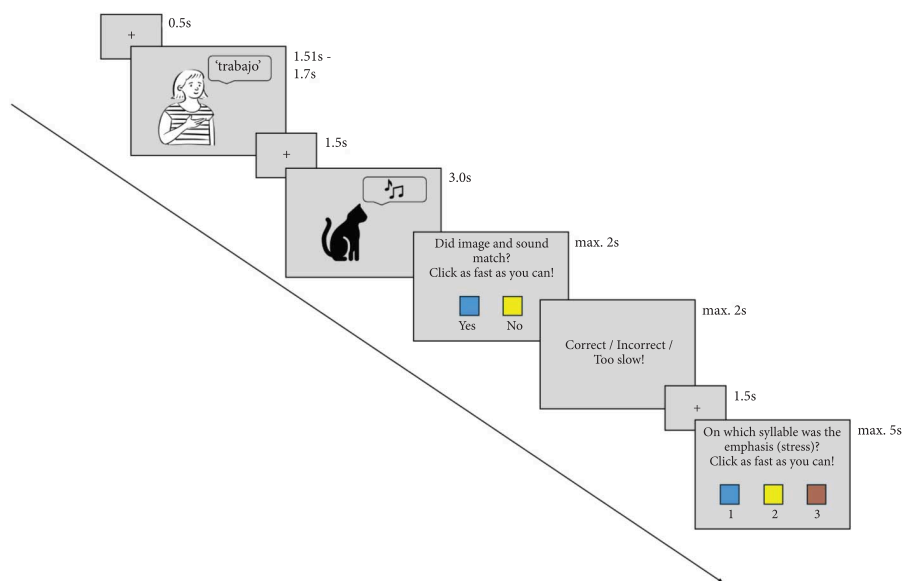


Figure 2. L2 stress perception task with distraction task

Note. The duration of the distraction task feedback screen depended on the participants' RT. Together, distraction task response screen and feedback always had a total duration of 4 seconds.

Stimulus videos

A female Spanish native was recorded producing the stimuli while sitting at a table with her hands in front of her. In the GC videos, she was instructed to perform a simple and natural beat gesture with her right hand to accentuate the stressed syllable, such that the stroke of the gesture was aligned with the stressed syllable. See Figure 3 for an example. In line with prior studies (Li et al., 2020; Hoetjes & van Maastricht, 2020), the audio tracks from the NGC videos were dubbed over the corresponding GC videos. This was done to control for potential acoustic differences in stress production between both conditions (cf. Pouw et al., 2021). The screen size was 367.9×614.0mm and the video was about one third in height and width of the computer screen, meaning the dimensions were about 122.63×204.67mm.



Figure 3. Example Video for Gesture Condition: ‘aCEPto’ (I accept)

Working memory tests

PWM was assessed using a Nonword Repetition Task created for highly educated Dutch speakers (de Bree, 2007). The test contained 16 nonwords of different lengths (four to six syllables). Participants’ repetitions were recorded and scored manually. A syllable was considered correct if all its segments were correctly produced in the participant’s speech. We calculated a percentage score representing the proportion of correctly repeated syllables per participant.

VWM was assessed using the digital Corsi Block Tapping task of Presentation® Cognitive Experiments III (v3, 1:2 staircase configuration; Neurobehavioral Systems, Inc., 2024a). Participants saw nine identical black squares (blocks) that were randomly distributed. Single blocks were highlighted sequentially. Afterwards, participants reproduced the order by clicking on the blocks. They completed 14 trials with the length of the list (i.e., the number of blocks to remember) being adapted based on performance. We calculated the mean span (MS) per participant, following Woods et al. (2011).

EWM was assessed using the digital Operation Span task from Presentation® Cognitive Experiments IV (v2, 1:2 staircase configuration; Neurobehavioral Systems, Inc., 2024b). In each trial, simple equations and letters were alternately presented on the screen. Participants indicated by button presses whether the equation was correct or incorrect. At the beginning of the test, participants were

instructed to memorize the letters that followed the equations. After a certain number of equation-letter iterations, participants were prompted to enter the letters in the correct order using the computer keyboard. The span length refers to the number of letters to remember and adapted automatically. Participants completed 14 trials.² We calculated a maximum span length (ML) per participant.

Procedure

After reading an information sheet on the study and signing a consent form, participants were individually tested in a soundproof booth. They filled in a background questionnaire partially adapted from the Language Background Questionnaire (Sabourin et al., 2016) covering general demographic information and their language background. Then, they performed the first part of the L2 stress perception task, followed by the three WM tasks, and finally the second part of the L2 stress perception task.

Statistical analysis

Data preprocessing and analyses were done using R (v4.1.3; R Core Team, 2022) in RStudio (v2023.12.1.402; Posit Team, 2024).

Data preprocessing

All WM scores were transformed into z-scores for better comparability. RTs from the L2 stress perception Task were log-transformed for the analysis. The following trials were excluded from the L2 stress perception Task: Trials with missing responses ($n=4$), trials with RTs longer than 3s ($n=13$), and all trials of the item ‘ayuDÓ’ (no-gesture) because it had a faulty video ($n=23$), leaving 98% of the data for analysis.

Mixed-effects modeling

Accuracy

Using the *lme4* package (Bates et al., 2015), a generalized mixed-effects (glmer) analysis with logistic regression was performed. First, a model was fitted based on the theoretical assumptions. The dependent variable was stress perception Accuracy, coded as 0 (incorrect) or 1 (correct). The fixed effects in this model were

2. Due to technical problems, six participants took the test in a different, non-adaptive setting (similar to the assessment in the Wechsler Adult Intelligence Scale). According to individual inspections, the scores were still appropriate and therefore kept for the analyses.

the main effects of the Gesture Condition (GC vs. NGC), the z-scores of the WM tests, and the interaction of the Gesture Condition with each of the scaled WM test scores. The model included random intercepts for Participant and Word. Random slopes for Participant and Word caused model divergence and were thus excluded. The following control variables were tested for their effects on model fit: First, cognate status of the word (with English or Dutch, yes vs. no), trial number (1–40), session (list 1 vs. 2), and stressed syllable (2nd vs. 3rd) were added simultaneously. Second, music instrument (yes vs. no), age, number of languages reported, multiple native languages (yes vs. no), Spanish course/learning (yes vs. no), student (yes vs. no) were added. Only age had a significant effect and was thus included in the final model: Accuracy Score $\sim$ Gesture Condition * EWM + Gesture Condition * PWM + Gesture Condition * VWM + Age + (1|Participant) + (1|Word).³

Reaction times

Using the *lme4* package (Bates et al., 2015), a linear mixed-effects (lmer) analysis was performed, with the log-transformed RTs (in s) as the dependent variable, and as fixed variables: Gesture Condition (GC vs. NGC), the z-scores of the WM tests, and the interaction of the Gesture Condition with each of the scaled WM test scores. The procedure was similar to the analysis of accuracy in that it was fitted based on the theoretical assumptions, in this case with only a random intercept for Participant. Random slopes for participants and items, and random intercepts for items were excluded as they caused model divergence. The same control variables as in the analysis of accuracy were entered individually and stepwise, and their effects were evaluated with likelihood ratio tests. Stressed Syllable (2nd vs. 3rd) and Trial Number were the two control variables with significant effects on RTs and were therefore included in the final model: RT $\sim$ Gesture Condition * EWM + Gesture Condition * PWM + Gesture Condition * VWM + Stressed Syllable + Trial + (1|Participant).

3. As it is not possible to use REML = FALSE in *glmer()*, generalized mixed-effects models cannot be compared with likelihood ratio tests to determine significant effects of predictors. Therefore, control variables were not added stepwise and *p*-values provided by *summary()* (R Core Team, 2022) and *report()* (Makowski et al., 2021) were relied on.

Results

Descriptives

The descriptive statistics of the L2 stress perception Task are shown in Tables 1 and 2. The most common errors were that participants indicated that the first syllable was stressed in words with stress on the second syllable (33.55% of incorrect trials), and that the second syllable was stressed in words with stress on the third syllable (29.99% of incorrect trials). Standard deviations indicate large variations in the RTs. Table 3 reports the descriptive statistics of the three WM tests. Large ranges and standard deviations indicate individual variation among participants.

Table 1. Accuracy in the L2 stress perception Task

Response	No-Gesture Condition (NGC)		Gesture Condition (GC)	
	<i>n</i>	%	<i>n</i>	%
Incorrect	229	25.85	239	26.18
Correct	657	74.15	674	73.82
Total	886	100	913	100

Table 2. RTs (in ms) in the L2 stress perception Task

Response	No-Gesture Condition (NGC)				Gesture Condition (GC)			
	<i>n</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>
Incorrect	229	497	381	360	239	457	373	364
Correct	657	382	337	228	674	385	339	260

Table 3. Working memory measures

WM Measure	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>
PWM: Nonword Repetition (percentage)	0.35	0.85	0.67	0.72	0.13
VWM: Corsi Block Tapping (MS score)	4.5	8.4	6.6	6.7	1.0
EWM: Operation Span (ML score)	2	11	6.7	7	2.1

Mixed-effects models

Accuracy

The model revealed no significant main effect of Gesture Condition (NGC vs. GC), nor of VWM or EWM scores on stress perception Accuracy. However, older participants ($b=0.06$, $SE=0.03$, $p=.04$) and those with higher PWM scores ($b=0.15$, $SE=0.41$, $p=.02$) identified more stressed syllables correctly (see Figure 4). The model’s total explanatory power was substantial (conditional $R^2=.45$), yet the part related to the fixed effects was small (marginal $R^2=.07$). All parameter estimates for the model are shown in Table 4. The significant effect of age seemed to be driven by four participants older than 40 years. Excluding these participants, the effects of age and PWM on accuracy were no longer significant.

Table 4. Parameter estimates of the final Mixed-Effects model for accuracy

Parameter	<i>b</i> Est	<i>SE</i>	<i>z</i>	<i>p</i>
Fixed Effects				
Intercept	−0.11	0.88	−0.13	.90
Age	0.06	0.03	2.06	.04*
Gesture Condition	−0.003	0.13	−0.02	.98
VWM Score	−0.13	0.37	−0.36	.72
EWM Score	0.15	0.42	0.37	.72
PWM Score	0.95	0.41	2.32	.02*
Gesture Condition * VWM Score	0.07	0.14	0.51	.61
Gesture Condition * EWM Score	−0.06	0.16	−.36	.72
Gesture Condition * PWM Score	0.02	0.16	0.15	.88
Random Effects				
Intercept by Word	0.08	0.29		
Intercept by Participant	1.92	1.39		
Fit Statistics				
Maximum Likelihood Deviance	1576.0			
Akaike Information Criterion	1598.0			
Bayesian Information Criterion	1658.5			

Note.

* = $p < .05$.

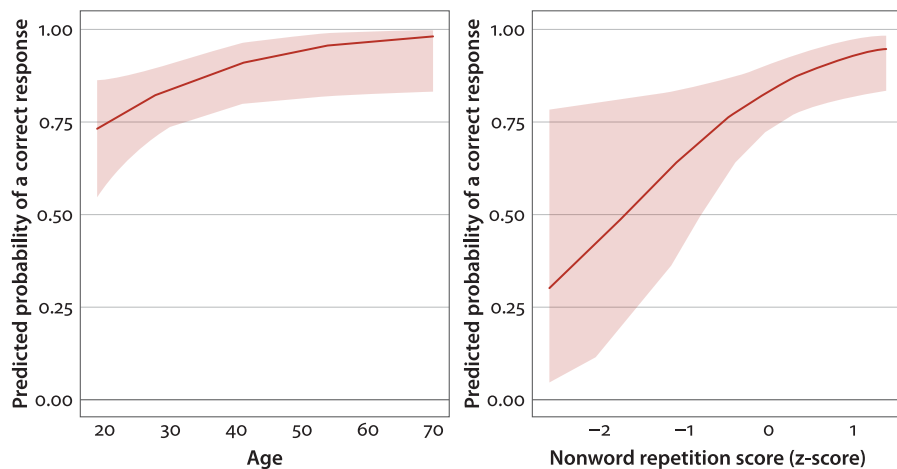


Figure 4. Predicted Probabilities of correct responses as a function of Age and PWM Score (Nonword Repetition), with 95% confidence interval

Reaction times

Within this model, the effects of two control variables were significant: RTs were slower when the third syllable was stressed ($b = 0.07$, $t_{(1331)} = 3.09$) and RTs increased with increasing trial number ($b = 0.002$, $t_{(1331)} = 2.09$). No other effects or interaction effects of predictors were significant, meaning that RTs did not differ between the gesture conditions, nor were they related to participants' performance in any of the WM tests. The model's total explanatory power was moderate (conditional $R^2 = .26$) and the part related to the fixed effects was small (marginal $R^2 = .03$). All parameter estimates for the final model of RTs are reported in Table 5.

Table 5. Parameter estimates of the final Mixed-Effects model for RTs

Parameter	<i>b</i> Est	SE	<i>t</i>
Fixed Effects			
Intercept	-1.11	0.05	-21.96
Stress on 3rd Syllable	0.07	0.02	3.09*
Trial	0.002	0.001	2.09*
Gesture Condition	0.01	0.02	0.27
VWM Score	-0.02	0.06	-0.33
EWM Score	0.05	0.06	0.85
PWM Score	-0.06	0.06	-1.00
Gesture Condition * VWM Score	-0.0004	0.03	-0.02

Table 5. (continued)

Parameter	<i>b</i> Est	<i>SE</i>	<i>t</i>
Gesture Condition * EWM Score	0.02	0.03	0.55
Gesture Condition * PWM Score	−0.02	0.03	−0.58
Random Effects			
Intercept by Participant	0.05	0.21	
Residual	0.15	0.38	
Fit Statistics			
Maximum Likelihood Deviance	1289.2		
Akaike Information Criterion	1313.2		
Bayesian Information Criterion	1375.5		

Note. *lmer()* does not provide *p*-values. Therefore, significant effects identified through likelihood ratio testing are highlighted with *.

Discussion

Acquiring L2 prosody is a challenge for many learners on their road to good L2 perception and production. Beat gestures have attracted attention as potentially useful tools in prosody teaching and learning, but previous findings have been inconclusive. The present study focused on L2 prosody perception and related the potential effects of beat gestures in this context to individual differences in WM components.

Beat gestures did not enhance L2 Spanish stress perception

In this study, participants did not correctly identify more stressed syllables in the GC compared to the NGC. The RTs in gesture trials were also not significantly different from those in the no-gesture trials. These results were unexpected, given that, in their L1, listeners are biased to perceiving word stress on a syllable that co-occurs with a beat gesture (Bosker & Peeters, 2021; Bujok et al., 2024; Rohrer et al., 2024).

It can be argued that the L2 stress perception task, which was administered in a language with which the participants were not familiar, was rather difficult. Assuming that most participants concluded that the correct syllable was always either the second or the third one, guessing alone would account for a 50% success rate in correct identification. This would suggest that the observed accuracy rates of 74% should be considered low, and indicative of a high cognitive load. The

high cognitive load might have prevented the multimodal integration of speech and gesture information, preventing a gesture benefit. Hoetjes and van Maastricht (2020) found a similar connection in their study on gestures and phoneme acquisition, where gestures did not facilitate acquisition in the task that was more difficult, in their case because a phoneme was unfamiliar. This is also in line with work by Kelly and Lee (2012) who state that learning new speech sounds is taxing, and that gestures can in those cases even be a burden. They suggest this may be because of the semantic information encoded in, in their study, iconic, gestures, and wonder whether similar effects would occur with gestures that carry no clear semantic meaning, such as beat gestures. The results of the current study suggest that this might be the case. Regarding the cognitive load of the L2 stress perception Task, future studies could investigate if advanced learners, for whom the L2 should be less taxing, do benefit from beat gestures.

There are also other potential explanations for the lack of effect of perceiving beat gestures. An open question is whether the beat gestures in the current study were noticeable enough. It therefore seems worth investigating whether the use of other types of gestures—e.g., more pronounced and larger beat gestures—would give different results. The current design could also be extended by giving participants explicit instructions to pay attention to the gestures and to test whether this affects the perceptual outcomes using an eye-tracking design.

The present study is the first to investigate the role of beat gestures in L2 stress perception, whereas previous L2 research on beat gestures was focused on speech production. Accordingly, there is little comparative material for the present findings. Contrary to the studies that report an effect of beat gestures and hand-clapping in pronunciation training on L2 production (Baills et al., 2018; Baills & Prieto, 2021; Gluhareva & Prieto, 2017; Llanes-Coromina et al., 2018; Zhang et al., 2020), the present study reveals that beat gestures may not facilitate L2 word stress perception. This contrast with previous work might be due to differences between L2 production and perception processes. Production and perception have traditionally been considered two separate processes in L2 research (Thomson, 2022), and although perception may facilitate production, one cannot simply be explained by the other. Learners can produce target-like prosody despite poor perception, or poor prosody production can occur despite successful perception (Altmann, 2006). Similarly, it is conceivable that beat gestures support L2 production even if they do not benefit L2 stress perception. This claim, however, can only be substantiated with more research that compares the effects of beat gestures in L2 stress perception and production using similar designs.

Together with previous studies, the present findings could also indicate that the effectiveness of beat gestures for L2 stress perception and production may depend on the learners' L1 and L2. Van Maastricht et al. (2019a) and van

Maastricht et al. (2022) did not find significant effects of gesture in L2 Spanish stress production training for Dutch natives (the same population tested in the current study), while Baills et al. (2018), Baills and Prieto (2021), Gluhareva and Prieto (2017), Llanes-Coromina et al. (2018), and Zhang et al. (2020) found benefits of gestures and hand-clapping for Catalan natives on L2 English and French pronunciation, and on Chinese speakers' acoustic measure of French word stress. Future investigations that systematically investigate different language pairs may provide more insights.

No interaction between gesture condition and working memory

The analyses did not show any significant interaction effects of the Gesture Condition and the WM measures. This, too, was an unexpected finding as research into differential effects of WM components in L1 studies with iconic gestures showed that VWM and EWM predicted to what extent participants benefited from visual input (Özer & Göksun, 2020; Wu & Coulson, 2014; Wu et al., 2022). It is, however, in line with Rohrer et al. (2025), who did not find PWM or VWM to modulate the beat gesture effect on L1 stress perception in Spanish. It might be that the multimodal integration of iconic gestures differs from beat gestures in general. One could speculate that VWM is more important with iconic gestures because they carry a semantic meaning that needs to be interpreted, while beat gestures are visually less complex and carry less, if any, semantic meaning. Future studies may directly compare the use of iconic and beat gestures in this context, as well as L1 and L2 to better understand the current variation in results.

Stress perception benefited from PWM

Our analyses revealed a main effect of PWM on stress perception accuracy, showing that participants with better PWM identified more stressed syllables correctly. This is in line with the *P/E Model* of L2 acquisition (Wen, 2019, 2023) where PWM is described as the system for short-term storage and articulatory rehearsal of sound information and as the *language acquisition device* with a particularly important role in the early stages of L2 learning. Both PWM processes were essential for completing the L2 stress perception Task in either gesture condition. The effect of PWM is also in line with studies where PWM has been found to determine speech input sensitivity in multimodal speech processing (Özer & Göksun, 2020; Wu & Coulson, 2014). Therefore, the present study adds to the evidence that learners with better PWM have an advantage in L2 learning.

Limitations

One limitation of this study is related to the instruction of the L2 stress perception Task and could also explain why no VWM and EWM effects were found: Participants were asked to indicate which of the three syllables they thought was stressed after completing the distraction task (cf. Figure 2). Because the task and response options were always the same, it is possible that participants simply remembered the number of the stressed syllable after viewing the video, rather than rehearsing speech and gesture in WM, reducing the chance of observing WM effects. Future research could make use of forced-choice tasks with minimal pairs as in Bosker and Peeters (2021) and Rohrer et al. (2024). Participants would then not be asked which syllable was stressed, but to decide between two written options whose position on the screen are randomized, for example, ‘aCEPto’ or ‘acepTÓ’ with capital letters indicating stress. Future studies could also test more advanced learners and present them with these verb forms in a full sentence. This would enable a design including both experimental items in which the subject of the verb is asked for (e.g., *Who accepted the invitation, he or I?*) and fillers in which other aspects of the sentence are questioned (e.g., *When did he accept the invitation?*). Such a set-up would require listeners to remember the utterance as a whole, rather than only its prominent syllable.

Even though our sample size was similar to those used in previous comparable studies (e.g., Bosker & Peeters, 2021; Rohrer et al., 2024; Gluhareva & Prieto, 2017), it cannot be ruled out that our design lacked the statistical power needed to detect a significant effect of gesture on lexical stress perception. The mixed-effects models could not always include all possible random slopes and intercepts to avoid divergence. Including random slopes for Participant and Item could have allowed the effects of the investigated factors to emerge. Future studies should put a large focus on a-priori power and sample size calculation in order to gain better insights into their results. Furthermore, the mixed-effects models had moderate to substantial overall explanatory power, with only small parts related to the fixed effects. This means that most of the variance could not be explained by Gesture Condition, WM scores, and control variables. Future research should investigate which other factors contribute to L2 stress perception, such as musical ability to L2 stress production in van Maastricht et al. (2022).

Finally, a limitation from an experimental point of view is that there was no condition in which participants saw a gesture that was not a beat gesture. Although the choice for a beat gesture was made because of its natural link to prosody, we cannot rule out that similar effects (or lack thereof) would apply to other types of gestures.

Conclusion

This study aimed to examine the effect of beat gestures on L2 stress perception and the role of WM components in this process. The presence of beat gestures did not significantly improve the performance of Dutch natives in the identification of word stress in Spanish verb forms. Except for a positive relation between PWM and accuracy in the L2 stress perception Task, no main effects of WM components nor interactions of WM components with gesture condition were found. These findings emphasize the importance of including different WM measures in future research to avoid overlooking its potentially larger-than-expected impact. In our study, there was an effect of PWM, but not of gesture or of VWM or EWM, which suggests that taking specific individual aspects of WM into account in the classroom may be more important for learning than including gestures.

Funding

Open Access publication of this article was funded through a Transformative Agreement with Radboud University Nijmegen.

Acknowledgements

We would like to thank Núria Domínguez Rodríguez for help in creating the video materials and Sweekriti Narayanan for sharing the PWM task with us.

References

- Altmann, H. (2006). The perception and production of second language stress: A cross-linguistic experimental study [Ph.D. Dissertation, University of Delaware]. Retrieved January 24, 2024, from https://www.ling.uni-stuttgart.de/institut/ifla/PDF_Upload/Heidi_Altmann/altmann-dissertation.pdf
-  Arvaniti, A. (2020). The phonetics of prosody. In A. Arvaniti (Ed.), *Oxford research encyclopedia of linguistics*. Oxford University Press.
-  Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of learning and motivation* (pp. 47–89, Vol. 8). Academic Press.
-  Baills, F., & Prieto, P. (2021). Embodying rhythmic properties of a foreign language through hand-clapping helps children to better pronounce words. *Language Teaching Research*, 136216882098671.
-  Baills, F., Zhang, Y., & Prieto, P. (2018). Hand-clapping to the rhythm of newly learned words improves L2 pronunciation: Evidence from Catalan and Chinese learners of French. *Speech Prosody* 2018, 498–502.

- doi Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- doi Bosker, H. R., & Peeters, D. (2021). Beat gestures influence which speech sounds you hear. *Proceedings of the Royal Society B*, 288(1943), 20202419.
- doi Bujok, R., Maran, M., Meyer, A. S., & Bosker, H. R. (2025). Beat gestures facilitate lexical access in constraining sentence contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- doi Bujok, R., Meyer, A. S., & Bosker, H. R. (2024). Audiovisual perception of lexical stress: Beat gestures and articulatory cues. *Language and Speech*, 68(1), 181–203.
- doi Cocks, N., Morgan, G., & Kita, S. (2011). Iconic gesture and speech integration in younger and older adults. *Gesture*, 11(1), 24–39.
- doi Darcy, I., Park, H., & Yang, C.-L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, 40, 63–72.
- de Bree, E. H. (2007). Dyslexia and phonology: A study of the phonological abilities of Dutch children at-risk of dyslexia. *LOT*, 155. <https://dspace.library.uu.nl/handle/1874/21522>
- doi Dörnyei, Z., & Skehan, P. (2008). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Wiley.
- doi Gluhareva, D., & Prieto, P. (2017). Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations. *Language Teaching Research*, 21(5), 609–631.
- doi Hoetjes, M., & van Maastricht, L. (2020). Using gesture to facilitate L2 phoneme acquisition: The importance of gesture and phoneme complexity. *Frontiers in Psychology*, 11, 575032.
- doi Honbolygó, F., Kóbor, A., & Csépe, V. (2019). Cognitive components of foreign word stress processing difficulty in speakers of a native language with non-contrastive stress. *International Journal of Bilingualism*, 23(2), 366–380.
- doi Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics*, 30(2), 225–249.
- doi Jackson, D. O. (2020). Working memory and second language development: A complex, dynamic future? *Studies in Second Language Learning and Teaching*, 10(1), 89–109.
- doi Jesse, A., & McQueen, J. M. (2014). Suprasegmental lexical stress cues in visual speech can guide spoken-word recognition. *Quarterly Journal of Experimental Psychology* (2006), 67(4), 793–808.
- doi Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, 44(2), 137–166.
- doi Kelly, S. D., & Lee, A. L. (2012). When actions speak too much louder than words: Hand gestures disrupt word learning when phonetic demands are high. *Language and Cognitive Processes*, 27(6), 793–807.
- doi Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57(3), 396–414.
- doi Leonard, T., & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, 26(10), 1457–1471.

- doi Li, P., Baills, F., & Prieto, P. (2020). Observing and producing durational hand gestures facilitates the pronunciation of novel vowel-length contrasts. *Studies in Second Language Acquisition*, 42(5), 1015–1039.
- doi Li, S. (2019). Six decades of language aptitude research. In Z. Wen, P. Skehan, A. Biedroń, S. Li & R. L. Sparks (Eds.), *Language aptitude* (pp. 78–96). Routledge.
- doi Linck, J.A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: a meta-analysis. *Psychonomic Bulletin & Review*, 21(4), 861–883.
- doi Llanes-Coromina, J., Prieto, P., & Rohrer, P.L. (2018). Brief training with rhythmic beat gestures helps L2 pronunciation in a reading aloud task. *Speech Prosody 2018*, 498–502.
- Makowski, D., Ben-Shachar, M.S., Patil, I., & Lüdtke, D. (2021). Automated results reporting as a practical tool to improve reproducibility and methodological best practices adoption. CRAN. <https://github.com/easystats/report>
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
- doi Mora, J.C., & Darcy, I. (2023). Individual differences in attention control and the processing of phonological contrasts in a second language. *Phonetica*, 80(3–4), 153–184.
- Neurobehavioral Systems, Inc. (2024a). Presentation Cognitive Experiments III v3. Retrieved May 10, 2024, from https://www.neurobs.com/menu_presentation/menu_teaching/exp_pack?pack_id=8
- Neurobehavioral Systems, Inc. (2024b). Presentation Cognitive Experiments IV v2. Retrieved May 10, 2024, from https://www.neurobs.com/menu_presentation/menu_teaching/exp_pack?pack_id=9
- doi Özer, D., & Göksun, T. (2020). Visual-spatial and verbal abilities differentially affect processing of gestural vs. spoken expressions. *Language, Cognition and Neuroscience*, 35(7), 896–914.
- doi Ping, R., & Goldin-Meadow, S. (2010). Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive Science*, 34(4), 602–619.
- Posit Team. (2024). *RStudio: Integrated development environment for R*. Posit Software, PBC. Boston, MA. <http://www.posit.co/>
- doi Pouw, W., de Jonge-Hoekstra, L., Harrison, S.J., Paxton, A., & Dixon, J.A. (2021). Gesture–speech physics in fluent speech and rhythmic upper limb movements. *Annals of the New York Academy of Sciences*, 1491(1), 89–105.
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- doi Reinisch, E., Jesse, A., & McQueen, J.M. (2010). Early use of phonetic information in spoken word recognition: Lexical stress drives eye movements immediately. *Quarterly Journal of Experimental Psychology* (2006), 63(4), 772–783.
- doi Rohrer, P.L., Bujok, R., van Maastricht, L., & Bosker, H.R. (2024). The timing of beat gestures affects lexical stress perception in Spanish. *Proceedings of Speech Prosody 2024*.
- doi Rohrer, P.L., Bujok, R., Van Maastricht, L., & Bosker, H.R. (2025). From “I dance” to “she danced” with a flick of the hands: Audiovisual stress perception in Spanish. *Psychonomic Bulletin & Review*.

- Sabourin, L., Leclerc, J.-C., Lapierre, M., Burkholder, M., & Brien, C. (2016). The language background questionnaires in L2 research: Teasing apart the variables. In L. Hracs (Ed.), *Proceedings of the 2016 Annual Conference of the Canadian Linguistics Association*. Retrieved April 1, 2024, from <https://erplinguottawa.weebly.com/lbq.html>
-  Schubotz, L., Holler, J., Drijvers, L., & Özyürek, A. (2021). Aging and working memory modulate the ability to benefit from visible speech and iconic gestures during speech-in-noise comprehension. *Psychological Research*, 85(5), 1997–2011.
-  Schwab, S., Giroud, N., Meyer, M., & Dellwo, V. (2020). Working memory and not acoustic sensitivity is related to stress processing ability in a foreign language: An ERP study. *Journal of Neurolinguistics*, 55, 100897.
- Skehan, P. (1989). *Individual differences in second language learning*. Routledge. Retrieved December 20, 2020, from <http://gbv.eblib.com/patron/FullRecord.aspx?p=1665702>
-  Stepanov, A., Kodrič, K. B., & Stateva, P. (2020). The role of working memory in children's ability for prosodic discrimination. *PloS One*, 15(3), e0229857.
-  Thompson, L.A. (1995). Encoding and memory for visible speech and gestures: A comparison between young and older adults. *Psychology and Aging*, 10(2), 215–228.
-  Thompson, L.A., & Guzman, F.A. (1999). Some limits on encoding visible speech and gestures using a dichotic shadowing task. *The Journals of Gerontology: Series B*, 54(6), P347–9.
-  Thomson, R. (2022). The relationship between L2 speech perception and production. In T.M. Derwing, M.J. Munro & R. Thomson (Eds.), *The Routledge handbook of second language acquisition and speaking* (pp. 370–385). Routledge Taylor & Francis Group.
-  van Maastricht, L., Hoetjes, M., & van Drie, E. (2019a). Do gestures during training facilitate L2 lexical stress acquisition by Dutch learners of Spanish? *Proceedings of the 15th International Conference on Auditory-Visual Speech Processing (AVSP2019)*.
-  van Maastricht, L., Hoetjes, M., & van der Heijden, L. (2022). Learning L2 Prosody using Gestures: The role of individual differences related to musicality. *Speech Prosody 2022*, 718–722.
-  van Maastricht, L., Krahmer, E., & Swerts, M. (2016a). Native speaker perceptions of (non)native prominence patterns: Effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness. *Speech Communication*, 83, 21–33.
-  van Maastricht, L., Krahmer, E., & Swerts, M. (2016b). Prominence patterns in a second language: Intonational transfer from Dutch to Spanish and vice versa. *Language Learning*, 66(1), 124–158.
-  van Maastricht, L., Krahmer, E., Swerts, M., & Prieto, P. (2019b). Learning direction matters: A study on L2 rhythm acquisition by Dutch learners of Spanish and Spanish learners of Dutch. *Studies in Second Language Acquisition*, 41(1), 87–121.
-  van Maastricht, L., Zee, T., Krahmer, E., & Swerts, M. (2021). The interplay of prosodic cues in the L2: How intonation, rhythm, and speech rate in speech by Spanish learners of Dutch contribute to L1 Dutch perceptions of accentedness and comprehensibility. *Speech Communication*, 133, 81–90.
-  Wen, Z. (2019). Working memory as language aptitude: The Phonological/Executive Model. In Z. Wen, P. Skehan, A. Biedroń, S. Li & R.L. Sparks (Eds.), *Language aptitude* (pp. 187–214). Routledge.

-  Wen, Z. (2023). Phonological and executive working memory. In Z. Wen, A. Biedroń & R. L. Sparks (Eds.), *Cognitive individual differences in second language acquisition* (pp. 97–124). De Gruyter Mouton.
-  Wen, Z., Juffs, A., & Winke, P. (2020). Measuring working memory. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 167–176). Routledge.
-  Wen, Z., & Li, S. (2019). Working memory in L2 learning and processing. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 365–389). Cambridge University Press.
-  Wen, Z., Skehan, P., Biedroń, A., Li, S., & Sparks, R. L. (Eds.). (2019). *Language aptitude*. Routledge.
-  Woods, D. L., Kishiyama, M. M., Lund, E. W., Herron, T. J., Edwards, B., Poliva, O., Hink, R. F., & Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal of Clinical and Experimental Neuropsychology*, 33(1), 101–111.
-  Wu, Y. C., & Coulson, S. (2014). Co-speech iconic gestures and visuo-spatial working memory. *Acta Psychologica*, 153, 39–50.
-  Wu, Y. C., Müller, H. M., & Coulson, S. (2022). Visuospatial working memory and understanding co-speech iconic gestures: Do gestures help to paint a mental picture? *Discourse Processes*, 59(4), 275–297.
-  Zhang, Y., Baills, F., & Prieto, P. (2020). Hand-clapping to the rhythm of newly learned words improves L2 pronunciation: Evidence from training Chinese adolescents with French words. *Language Teaching Research*, 24(5), 666–689.
-  Zhu, Y., & Mok, P. (2022). The role of prosody across languages. In T. M. Derwing, M. J. Munro & R. Thomson (Eds.), *The Routledge handbook of second language acquisition and speaking*. Routledge Taylor & Francis Group.

Address for correspondence

Lieke van Maastricht
 Centre for Language Studies
 Donders Institute for Brain, Cognition, and Behaviour
 Radboud University
 Erasmusplein 1
 Office E8.20
 6525HT Nijmegen
 The Netherlands
 lieke.vanmaastricht@ru.nl
<https://orcid.org/0000-0001-9968-4331>

Co-author information

Luise Böttcher
Radboud University
luise.boettcher@outlook.com
 <https://orcid.org/0000-0002-4023-6670>

Marieke Hoetjes
Centre for Language Studies
Radboud University
marieke.hoetjes@ru.nl
 <https://orcid.org/0000-0002-9760-7727>

Publication history

Date received: 12 January 2026
Date accepted: 28 August 2026
Published online: 9 October 2026