



University of Dundee

Explaining flexible continuous speech comprehension from individual motor rhythms

Lubinus, Christina; Keitel, Anne; Obleser, Jonas; Poeppel, David; Rimmele, Johanna M

DOI:

[10.1101/2022.04.01.486685](https://doi.org/10.1101/2022.04.01.486685)

Publication date:

2022

Licence:

CC BY-NC-ND

Document Version

Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Lubinus, C., Keitel, A., Obleser, J., Poeppel, D., & Rimmele, J. M. (2022). *Explaining flexible continuous speech comprehension from individual motor rhythms*. BioRxiv. <https://doi.org/10.1101/2022.04.01.486685>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Explaining flexible continuous speech comprehension from individual motor rhythms

Christina Lubinus¹, Anne Keitel², Jonas Obleser^{3,4}, David Poeppel^{5,6,7}, and Johanna M. Rimmele^{1,6}

¹Department of Neuroscience and Department of Cognitive Neuropsychology
Max-Planck-Institute for Empirical Aesthetics
60322 Frankfurt am Main, Germany,

²Psychology
University of Dundee,
Dundee DD1 4HN, UK,

³Department of Psychology
University of Lübeck,
Lübeck, Germany,

⁴Center for Brain, Behavior, and Metabolism
University of Lübeck,
Lübeck, Germany,

⁵Department of Psychology
New York University,
New York, NY, USA

⁶Max Planck NYU Center for Language, Music, and Emotion
Frankfurt am Main, Germany, New York, NY, USA

⁷Ernst Strüngmann Institute for Neuroscience (in Cooperation with Max Planck Society)
Frankfurt am Main, Germany

Corresponding author:

Christina Lubinus
Email: christina.lubinus@ae.mpg.de
Phone: +49 69 8300 479-348

Abstract

When speech is too fast, the tracking of the acoustic signal along the auditory pathway deteriorates, leading to suboptimal speech segmentation and decoding of speech information. Thus, speech comprehension is limited by the temporal constraints of the auditory system. Here we ask whether individual differences in auditory-motor coupling strength in part shape these temporal constraints. In two behavioral experiments, we characterize individual differences in the comprehension of naturalistic speech as function of the individual synchronization between the auditory and motor systems and the preferred frequencies of the systems. Obviously, speech comprehension declined at higher speech rates. Importantly, however, both higher auditory-motor synchronization and higher spontaneous speech motor production rates were predictive of better speech-comprehension performance. Furthermore, performance increased with higher working memory capacity (Digit Span) and higher linguistic, model-based sentence predictability – particularly so at higher speech rates and for individuals with high auditory-motor synchronization. These findings support the notion of an individual preferred auditory-motor regime that allows for optimal speech processing. The data provide evidence for a model that assigns a central role to motor-system-dependent individual flexibility in continuous speech comprehension.

Keywords: speech perception; speech production; auditory-motor synchronization; oscillations; audiomotor.

1 1. Introduction

2 Speech comprehension relies on temporal processing, as speech and other naturalistic signals
3 have a complex temporal structure with information at different timescales(1). The temporal
4 constraints of the auditory system limit our ability to understand speech at fast rates(2,3). In-
5 terestingly, the motor system can under certain conditions provide temporal predictions that
6 aid auditory perception(4,5). Accordingly, current oscillatory models of speech comprehension
7 propose that properties of the auditory but also the motor system affect the quality of auditory
8 processing(6,7). In two behavioral experiments, we investigate how the auditory, the motor
9 system, and their synchronization shape individual flexibility of comprehending fast continuous
10 speech.

11 Auditory temporal constraints have been observed as preferred rates of auditory speech(8,9)
12 processing (but also of tones(10,11), and amplitude modulated sounds(11–14)) and explained
13 in the context of neurocognitive models of speech perception. According to such proposals,
14 humans capitalize on temporal information by dynamically aligning ongoing brain activity in
15 auditory cortex to the temporal patterns inherent to the acoustic speech signal(15–18). By hy-
16 pothesis, endogenous theta brain rhythms in auditory cortex partition the continuous auditory
17 stream into smaller chunks at roughly the syllabic scale by tracking quasi-rhythmic temporal
18 fluctuations in the speech envelope. This chunking mechanism allows for the decoding of seg-
19 mental phonology – and ultimately linguistic meaning(15,18–20). The decoding of the speech
20 signal is accomplished seemingly effortlessly within an optimal range centered in the traditional
21 theta band(18), whereas comprehension deteriorates strongly for speech presented beyond
22 ~9 Hz(2,3). While much research has focused on the apparent *stability* of the average acoustic
23 modulation rate at the syllabic scale(8,9), the *flexibility* in speech comprehension(9,21), that
24 is, what constitutes individual differences in understanding fast speech rates, is poorly under-
25 stood.

26 The motor system, and neural auditory-motor coupling in particular, is a plausible candidate to
27 facilitate individual differences in auditory speech processing abilities. Two arguments sup-
28 porting this notion are the motor systems' modulatory effect on auditory perception(22–24) and
29 its susceptibility to training(25–27). While there is evidence suggesting that the auditory and
30 speech motor brain areas are intertwined during speech comprehension(28–32), the extent to
31 which speech motor processing modulates auditory processing is debated(5,33,34). Specifi-
32 cally, endogenous brain rhythms in both auditory(20,35) and motor(35,36) cortex have been

33 observed to track the acoustic speech signal, and are characterized by preferred frequen-
34 cies(19,37,38). In contrast to neural measures of preferred frequencies(37–39), here we used
35 a behavioral estimate termed “preferred” or “spontaneous” rate. Furthermore, neural coupling
36 between auditory and motor brain areas during speech processing(35,36,40,41) has been hy-
37 pothesized to provide temporal predictions about upcoming sensory events to the auditory
38 cortex(4,41–43). The precision of these predictions may be proportional to the strength of au-
39 ditory-motor cortex coupling.

40 Auditory-motor cortex coupling strength varies across the population, as shown by recent
41 work(6,10,40,44,45). Assaneo et al.(40) developed a behavioral protocol (spontaneous
42 speech synchronization test; SSS-test) which quantifies the strength of auditory-to-motor syn-
43 chronization during speech production in individuals. The authors reported that auditory-motor
44 synchronization is characterized by a bimodal distribution in the population, classifying individ-
45 uals into high versus low synchronizers. (The rejection of unimodality has been previously
46 shown with large sample sizes(40, see also: ,46).) Importantly, in addition to superior behav-
47 ioral synchronization, high synchronizers have stronger structural and functional connectivity
48 between auditory and speech motor cortices (see 40, Figure 3A and B). Thus, the SSS-test
49 provides not only a behavioral measure but also approximates individual differences in neu-
50 ronal auditory-motor coupling strength. We propose that the individual variability in auditory-
51 motor synchronization, previously observed to predict differences in word learning(40), syllable
52 detection(6), and rate discrimination(10), as well as the individual variability in preferred audi-
53 tory and motor rate, predicts differences in an individuals’ ability to comprehend continuous
54 speech at fast syllabic rates.

55 The influence of individual auditory-motor coupling strength on behavioral performance has so
56 far been established for behavioral paradigms using rather basic auditory and speech stimuli,
57 e.g. tones or syllables(6,10,40). The current study assesses its importance in a more natural-
58 istic context: during the comprehension of continuous speech. This adds several layers of
59 complexity. First, as speech unfolds over time, processing of continuous, i.e. longer and more
60 complex, speech naturally demands more working memory capacity for maintenance and ac-
61 cess to linguistic and context information(47). Second, rich linguistic context is used to derive
62 linguistic predictions about upcoming words and sentences(48–51). When linguistic predicta-
63 bility of a sentence is high(52), speech comprehension is improved, even in adverse listening
64 situations(53,54). Thus, similar to auditory-motor synchronization, linguistic predictability offers
65 a compensatory mechanism when comprehension is difficult.

66 In summary, we investigate the role of auditory-motor synchronization with the SSS-test and
67 the role of preferred rhythms of the auditory and motor systems for the individual flexibility of
68 the comprehension of continuous speech. First, based on an established literature(3,18,55–
69 57), we expected a decline in comprehension performance at syllabic rates beyond the theta
70 range. Second, as a facilitatory effect of auditory-motor coupling on auditory processing has
71 been observed(6,10,40), we hypothesized that individual differences in comprehension perfor-
72 mance could be predicted by individual auditory-motor synchronization, with superior speech
73 comprehension for high synchronizers. Such a facilitatory effect might be strongest in demand-
74 ing listening situations, such as at fast syllabic rates(5,10). Third, while the consequences of
75 potential individual variation in the preferred rates of the motor and auditory systems are not
76 clearly understood, based on previous findings(35) we expected a systematic relation of both
77 preferred auditory and motor rates with individual speech comprehension performance. Fi-
78 nally, we hypothesized that linguistic predictability and working memory span should positively
79 affect speech comprehension. Similar to auditory-motor synchronization, we expected linguis-
80 tic predictability to interact with syllabic rate, such that both systems would become stronger
81 predictors for speech comprehension as syllabic rate increases.

82 2. Methods

83 Two behavioral experiments and a control experiment were conducted: Experiment 1 was per-
84 formed in the laboratory and investigated the influence of the spontaneous speech motor pro-
85 duction rate on speech comprehension performance. In Experiment 2 we aimed to understand
86 the complex interplay of multiple variables during speech comprehension beyond the sponta-
87 neous speech motor production rate. To this end, we additionally measured participants' pre-
88 ferred auditory rate, auditory-motor synchronization, and working memory capacity. Experi-
89 ment 2 and the control experiment were online studies. All studies were approved by the local
90 ethics committees (Experiment 1: committee of the School of Social Sciences, University of
91 Dundee, UK (No. UoD-SoSS-PSY-UG-2019-88), Experiment 2 and control experiment: pro-
92 cedures were approved by the Ethics Council of the Max Planck Society (2017_12)).

93 2.1. Participants

94 Participants were English native speakers with normal hearing and no neurological or psycho-
95 logical disorders (Exp 1: N = 34, Exp 2: N = 82, Control: N = 39). Participation was voluntary.
96 For a detailed description of participants, stimuli, exclusion criteria, and tasks please refer to
97 Supplementary Methods, Figures 1-2, and Tables 1-2.

98 2.2. Design and materials

99 *Speech comprehension task*

100 In two speech comprehension tasks, we measured participants ability to comprehend sen-
101 tences at various syllabic rates. Sentences were presented at 7 (Exp 1: [8.2, 9.0, 9.8, 11.0,
102 12.1, 14.0, 16.4]) or 6 (Exp 2: [5.00, 10.69, 12.48, 13.58, 14.38, 15.00]) rates. In Experiment
103 1, participants performed a classic intelligibility task, i.e. also termed “word identification task”
104 (58,59, for review: ,60). On each trial (N = 70), a sentence was presented through headphones
105 and participants verbally repeated the sentence as accurately as possible (Fig. 1A). Re-
106 sponses were recorded.

107 In Experiment 2, speech comprehension was measured by a word-order task. Participants
108 listened to one sentence per trial (N = 240), followed by the presentation of two words from the
109 sentence on screen. Participants indicated via button press which word they heard first (Fig.
110 2A).

111 *Speech production task*

112 In the speech production tasks we estimated participants individual spontaneous speech motor
113 production rate. In Experiment 1, the speech production task was operationalized by partici-
114 pants reading a text excerpt (216 words) from a printout. Participants were instructed to read
115 the text excerpt out loud at a comfortable and natural pace while their speech was recorded
116 (Fig. 1B).

117 In Experiment 2, participants were asked to produce continuous, “natural” speech. To facilitate
118 fluent production, they were prompted by a question/statement belonging to six thematic cat-
119 egories (6 trials; own life, preferences, people, culture/traditions, society/politics, general
120 knowledge, see Supplementary Table 2). Each response period lasted 30 seconds and trials
121 were separated by self-paced breaks (Fig. 2C). While speaking, participants simultaneously
122 listened to white noise. The white noise was introduced to measure the preferred rate of the

123 motor system, without potential interference from auditory feedback. A second reason was to
124 be consistent with the protocol from the SSS-test(40,61; also see below). Note that this proce-
125 dure was not applied in Experiment 1.

126 *Auditory rate task (only Exp 2)*

127 To measure participants preferred auditory rate, we implemented a two-interval forced choice
128 (2IFC) task, presenting a reference and a comparison stimulus at each trial. Participants indi-
129 cated via button press which stimulus they preferred (Fig. 2B). Stimuli were presented at syl-
130 labic rates from 3.00 to 8.50 syllables/s (3.00, 3.92, 4.83, 5.75, 6.67, 7.58, 8.50). A reference
131 rate, e.g. 3.00 syllables/s, was compared to all syllabic rates, including itself. For each refer-
132 ence/comparison pair the same sentence was presented – the stimuli only differed in their
133 syllabic rates. Additionally, the task included catch trials to measure participant’s engagement
134 (see Supplementary Methods for details).

135 *Spontaneous speech synchronization (SSS) test (only Exp 2)*

136 We measured participant’s auditory-motor synchronization using the SSS-test (for details: ,40).
137 In the main task, participants listened to a random syllable train and whispered along for a
138 duration of 80s. They were instructed to synchronize their own syllable production to the stim-
139 ulus presented through their headphones (Fig. 2D). The syllable rate in the auditory stimulus
140 progressively increased in frequency from 4.3 to 4.7 syllables/s in increments of 0.1 syllables/s,
141 every 60 syllables. Participants completed two trials, while the whispering was recorded.

142 Participants’ syllable production was masked by the simultaneously presented auditory sylla-
143 ble train. The masking procedure suppresses auditory feedback, allowing us better to isolate
144 the synchronization of motor production to the auditory input, without interference of auditory
145 feedback(44).

146 *Digit span test (only Exp 2)*

147 Working memory capacity was quantified using the forward and backward(62) digit span test.
148 As for the backward test data is missing for N = 21 participants, only the forward span is re-
149 ported. Digit spans were presented auditorily and participants typed in their responses(63).

150 *Control Experiment*

151 We designed a control experiment to test if the correct word order from the word order task of
152 Experiment 2 could be guessed from the target words alone, that is, without understanding the
153 sentence. The task consisted in judging which of two words would be more likely to occur first
154 in a *hypothetical sentence*. On each trial, two words were presented on screen and participants
155 indicated their choice via button press. Importantly, 1) participants did not listen to a full sen-
156 tence at any time and 2) the target words were taken from the stimulus materials actually
157 presented in Experiment 2.

158 2.3. Analysis

159 *Spontaneous speech motor production rate (Exp 1 + 2)*

160 The individual *spontaneous speech motor production rate*, i.e. articulation rate(64), was com-
161 puted using Praat software(65) by automatically detecting syllable nuclei. The number of syl-
162 lable nuclei was divided by the duration of the utterance, disregarding silent pauses. For Ex-
163 periment 1, the production rate was computed across the entire reading paragraph. For Ex-
164 periment 2, it was first calculated for each trial (30 s) separately. The motor rate was then
165 averaged across all trials.

166 *Preferred auditory rate (Exp 2)*

167 First, participants with low performance in the catch trials of the preferred auditory rate task
168 (below 75% correct) were excluded; amongst the remaining participants ($N = 82$) catch trial
169 performance was very high ($M = 98.48\%$, $SD = 3.71$). To compute the preferred auditory rate,
170 a distribution of preferred frequencies was derived from all trials -except catch trials- by aggre-
171 gating the frequency of each trials' preferred item. Then a gaussian function was fitted to each
172 participants' distribution and two parameters were extracted: the peak as index for the pre-
173 ferred frequency and the full-width-at-half-maximum (FWHM) as index for the specificity of the
174 response (lower FWHM equals stronger preference for one frequency).

175 *Auditory-motor synchronization (Exp 2)*

176 From the SSS-test(40) we derived the participant's auditory-motor synchronization by calcu-
177 lating the phase-locking value (PLV)(66) between the (cochlea) envelopes of the auditory and
178 the speech signals.

179

$$180 \quad PLV = \frac{1}{T} \left| \sum_{t=1}^T e^{i(\theta_1(t) - \theta_2(t))} \right| \quad (1)$$

181 where T is the total number of time points, t denotes the discretized time, and θ_1 and θ_2 are
182 the phase of the first and the second signals, respectively.

183 To obtain the cochlear envelope of the syllable train (auditory channels: 180–7,246 Hz), we
184 used the Chimera Software toolbox(67). For the recorded speech signal the amplitude enve-
185 lope was quantified as the absolute value of the Hilbert transform. Both envelopes were
186 downsampled to 100 Hz and bandpass filtered (3.5-5.5 Hz) before their phase was extracted
187 by means of the Hilbert transform. The PLV was first estimated for each trial of the SSS-test
188 (time windows 5s, overlap 2s) and then averaged across runs, resulting in a mean PLV. The
189 distribution of mean PLV values was subjected to a k-means algorithm(68) ($k = 2$) to split
190 participants into a high- and a low-synchronizer group. Speech auditory-motor synchronization
191 (PLV) was treated as bimodal variable based on previous research that rejected unimodality
192 based on larger samples(40, see also: ,46).

193 *Linguistic predictability – Recurrent neural network (Exp 2)*

194 Linguistic predictability of all stimulus sentences was measured by deriving single-sentence
195 perplexity from a recurrent neural network language model. A language model, such as a re-
196 current neural network, assigns probabilities to all words in a sequence of words. From the
197 single-word probabilities, we derived one value per sentence, quantifying its predictabil-
198 ity(69,70). This so-called perplexity is the most common intrinsic evaluation metric of language
199 models(71–73). It is computed as the inverse of the mean probability of a sentence weighted
200 by sentence length(69), i.e. lower perplexity values equal higher sentence predictability (see
201 Supplementary Methods for full details on RNN and perplexity).

202 *Mixed effects models*

203 For both experiments we performed mixed effects analyses to quantify how speech compre-
204 hension was affected by all variables of interest. Mixed models were computed using the R
205 packages *lme4* (v1.1-29) and *mgcv* (v1.8-39), as set up in Rstudio (version 2022.2.1.461).
206 Mixed-effects, rather than fixed-effects models were chosen to account for idiosyncratic varia-
207 tion within variables, i.e. repeated measures and therefrom resulting interdependencies

208 between data points(74,75). Thus, both models included random intercepts for *participant* and
209 *items*.

210 In Experiment 1, we computed a generalized additive mixed-effects model (GAMM) using the
211 *mgcv:gam* function. For the dependent variable *speech comprehension*, we calculated the
212 percentage of correctly repeated words for each sentence and subject from the speech com-
213 prehension task. The number of correct words was counted manually and transformed into a
214 percentage. Then the dependent variable (single-trial data) was modelled as a function of the
215 fixed effects *syllabic rate* and *spontaneous speech motor production rate*. A random slope for
216 *syllabic rate* could not be included because the model failed to converge, thus the model in-
217 cluded only random intercepts. Overall, the model explained ~77% of the variance.

218 In Experiment 2, the dependent variable *speech comprehension* was binary (*correct vs incor-*
219 *rect* word order judgment). Thus, we employed a generalized linear mixed-effects model
220 (GLMM; *lme4:glmer* function) with a binomial logit link function. In terms of fixed effects, the
221 model included all variables of interest: *syllabic rate*, *preferred motor rate*, *preferred auditory*
222 *rate*, *auditory-motor synchronization*, *working memory*, *sentence predictability*. Additionally,
223 we introduced several linguistic and other covariates for nuisance control(76): *predictability*
224 *target 1*, *predictability target 2*, *sentence length* (# of words), *target distance* (i.e., distance in
225 words between the target words), *compression/dilation* of audio file. In addition to random in-
226 tercepts, the model contained a by-participant random slope for *syllabic rate*, allowing the
227 strength of the effect of the rate manipulation on the dependent variable to vary between par-
228 ticipants(74,75). Continuous predictor variables were z-transformed to facilitate the interpreta-
229 tion and comparison of the strength of the different predictors(77). Thus, the coefficients of all
230 continuous predictors reflect log changes in comprehension for each unit (*SD*) increase in a
231 given predictor. We observed no problems with (multi-)collinearity, all variance inflation factors
232 were < 1.2 (package *car* version 3.0-10(78)). Overall, the model explained ~38% of the vari-
233 ance.

234 *Control experiment*

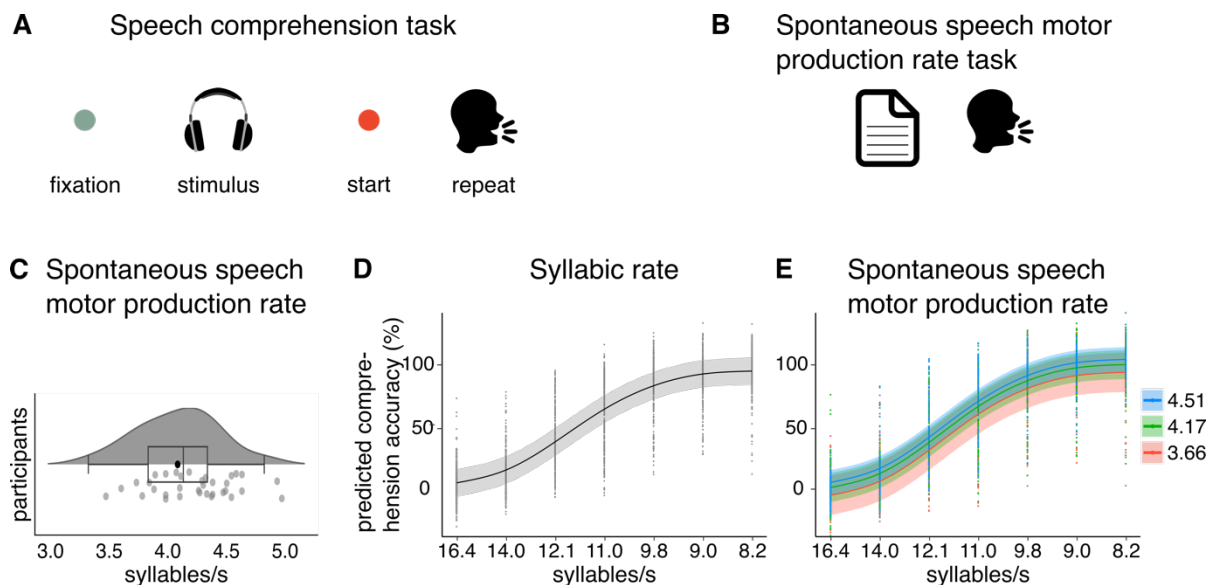
235 For each trial, we computed how many participants correctly guessed the word order (in per-
236 cent, “*word order index*”). In a new GLMM analysis, this *word order index* was added as co-
237 variate into the model from the main analysis while all other parameters remained the same.

238 3. Results

239 3.1. Experiment 1

240 In Experiment 1, we asked the question: to what extent is speech comprehension affected by
241 one's spontaneous speech motor production rate? *Speech comprehension* was measured as
242 the percentage of correctly repeated words in an intelligibility task (2.75% to 93.70% on aver-
243 age across participants). We observed a mean *spontaneous speech motor production rate* of
244 4.11 syllables per second ($SD = 0.35$, min = 3.35, max = 4.85) across participants (Fig. 1C).

245 As expected, the GMM revealed a main effect of *syllabic rate*: slower speech stimuli were
246 associated with better speech comprehension (edf = 4.61, $F = 1260.90$, $p < .001$, Fig. 1D, see
247 Supplementary Table 3). Importantly, we observed that the *spontaneous speech motor pro-*
248 *duction rate* influenced speech comprehension: the higher the individual *spontaneous speech*
249 *motor production rate*, the better the speech comprehension performance (edf = 1.00, $F = 4.37$,
250 $p = .036$, Fig. 1E).



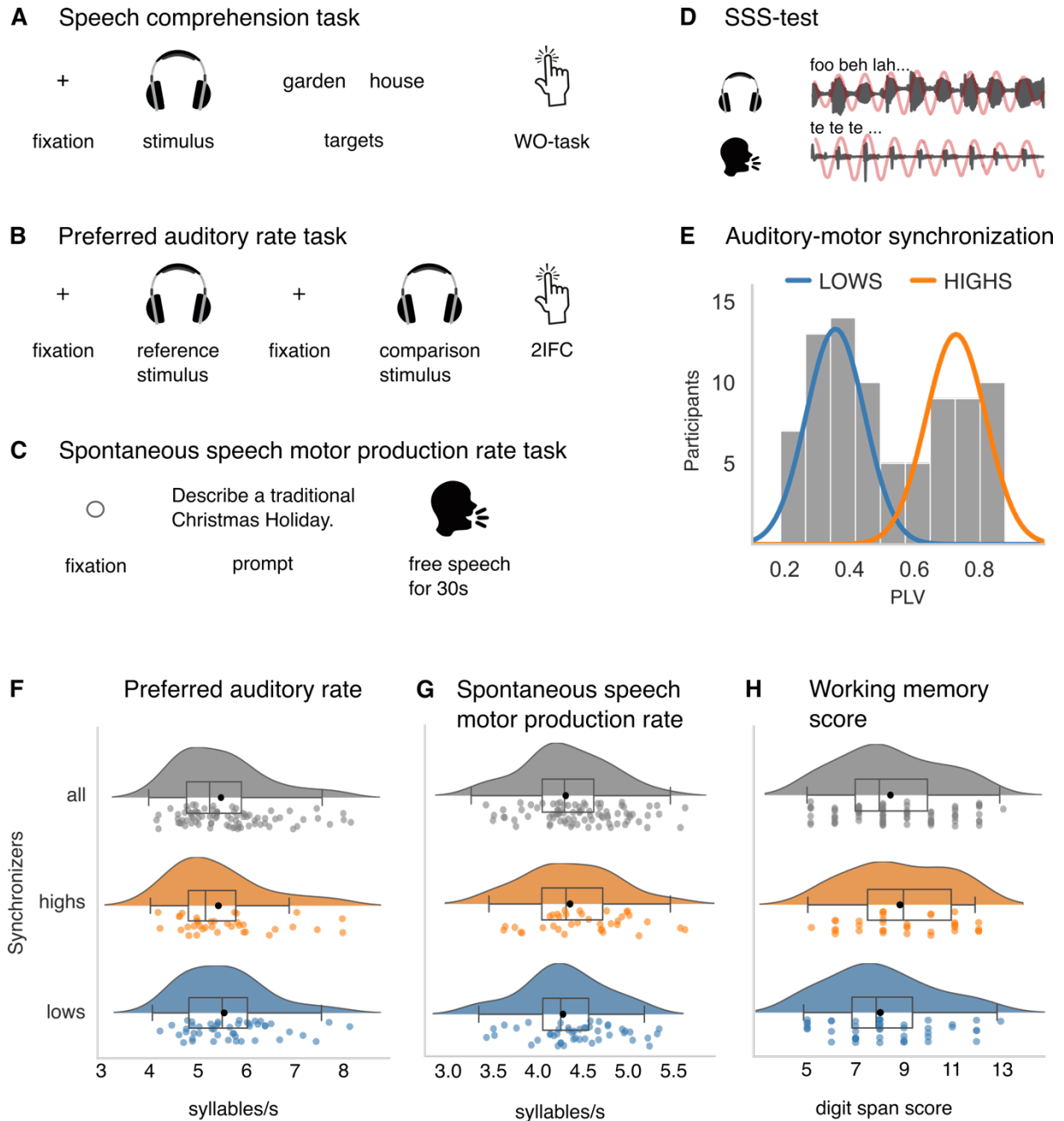
251

252 **Figure 1. A.** Example trial for the speech comprehension task. Participants fixated on a green fixation
253 dot while presented auditorily with a sentence. On stimulus offset the fixation dot turned red, indicating
254 to commence recall, i.e., reporting the sentence back. **B.** Spontaneous speech motor production rate
255 task. Participants read a stimulus paragraph from a paper. **C.** Spontaneous speech motor production
256 rate. We observed spontaneous speech motor production rates between 3.35 and 4.85 syllables/s ($M =$
257 4.11 syllables/s, left). The violin and boxplot show summary statistics and density: the median center
258 line, 25th to 75th percentile hinges, whiskers indicate minimum and maximum within $1.5 \times$ interquartile
259 range. Grey dots represent participants individual speech motor productions rates, averaged across 6
260 trials. **D.** Main effect of syllabic rate. Plot shows the predicted main effect of syllabic rate from the gen-
261 eralized additive mixed model (GMM). Black line indicates the predicted effect with 95% confidence
262 interval in grey. Black dots show trial-level speech comprehension performance per subject and rate

263 condition. **E.** Main effect of spontaneous speech motor production rate. Plot shows the predicted main
264 effect of spontaneous speech motor production rate from the GAMM. Colored lines indicate the pre-
265 dicted effect with 95% confidence interval in the corresponding color. Colored dots show trial-level
266 speech comprehension performance per subject and rate condition.

267 3.2. Experiment 2

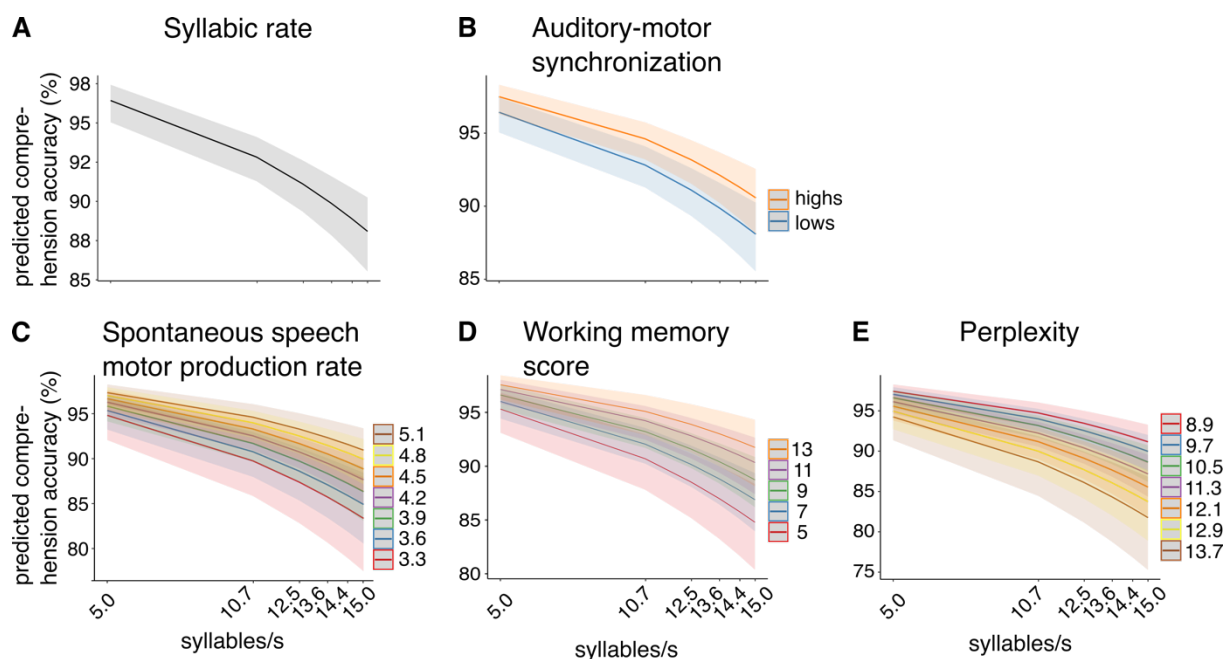
268 First, in line with the first experiment, we observed a mean *spontaneous speech motor pro-*
269 *duction rate* of 4.32 syllables per second across participants ($SD = 0.45$, Min = 3.36, Max =
270 5.38 syllables per second, Fig. 2G). Within-subject variance was low (Supplementary Fig. 3),
271 suggesting that participants' articulation rate was stable across trials. Second, participants
272 showed a *preferred auditory rate* of ~ 5.57 syllables per second (peak: $M = 5.57$, $SD = 0.86$,
273 Min = 4.16, Max = 7.92; FWHM, $M = 4.89$, $SD = 0.50$, Min = 3.23, Max = 5.50; Fig. 2F). Single-
274 subject raw data can be inspected in Supplementary Fig. 4. Third, *auditory-to-motor speech*
275 *synchronization* was quantified using the SSS-test(40), classifying participants as HIGH or
276 LOW synchronizers (mean PLV HIGHs = 0.73, $SD = 0.09$, mean PLV LOWs = 0.36, $SD = 0.09$,
277 Fig. 2E). Fourth, *working memory* was measured by means of the digit span test (62) which
278 revealed a mean forward digit score of $M = 8.46$ ($SD = 2.12$, Min = 5.00, Max = 13.00, Fig.
279 2H).



280

281 **Figure 2.** Panels **A**, **B**, and **C** visualize example trials for the speech comprehension task, the preferred
 282 auditory rate task, and the spontaneous speech motor production task, respectively. **D**. Schematic rep-
 283 resentation of the SSS-test, used to measure auditory-motor synchronization. Participants whisper a
 284 syllable (here /te/). **E**. Histogram of auditory-motor synchronization strength, obtained with the SSS-test.
 285 Participants were classified into high and low synchronizers (highs, lows) based on their PLV using k-
 286 means clustering. Group affiliation is overlaid by colored lines representing fitted normal distributions.
 287 **F**. Participants showed a mean preferred auditory rate of 5.57 syllables/s ($SD = 0.86$), with no differ-
 288 ences between high and low synchronizers ($U = 897.5$, $p = .48$). **G**. We observed a spontaneous speech
 289 motor production rate between 3.36 and 5.38 syllables/s ($M = 4.32$, $SD = 0.45$) and no group difference
 290 between high and low synchronizers ($U = 767.0$, $p = .60$). **H**. Working memory capacity was indicated
 291 by a mean digit-span forward score of 8.46 ($SD = 2.12$) and the score did not differ between high and
 292 low synchronizers ($U = 666.5$, $p = .14$).

293 The GLMM revealed that *syllabic rate* significantly influenced participants' comprehension ac-
294 curacy: for each increase of syllabic rate by one syllable/s, the odds of a correct word order
295 judgment decreased (*odds ratio (OR)* = 0.65, *std. error (SE)* = 0.04, $p < .001$, Fig. 3A). This
296 main effect of syllabic rate is consistent with a decline of speech comprehension performance
297 at higher syllabic rates (3). In line with our hypothesis, we observed main effects for *spontane-*
298 *ous speech motor production rate* and *auditory-motor synchronization*. The higher a partici-
299 pant's *spontaneous speech motor production rate*, the better the performance in the word order
300 task ($OR = 1.19$, $SE = 0.09$, $p = .014$, Fig. 3C), replicating our finding from the first experiment.
301 For *auditory-motor synchronization*, being a dichotomous variable (i.e., HIGH vs. LOW)(40),
302 performance in the word order judgment task was higher for high compared to low synchroniz-
303 ers ($OR = 1.34$, $SE = 0.20$, $p = .048$, Fig. 3B). That is, across all trials, high synchronizers were
304 more likely to correctly perform the task. Additionally, the model revealed a positive effect for
305 *working memory score* ($OR = 1.20$, $SE = 0.09$, $p = .012$, Fig. 3D). This main effect suggests
306 that better working memory performance enabled participants to better perform on the speech
307 comprehension task. We did not observe a reliable effect of *preferred auditory rate* on speech
308 comprehension ($OR = 1.14$, $SE = 0.08$, $p = .072$). In contrast to our hypothesis, we observed
309 no interaction effect of *syllabic rate* and *auditory-motor synchronization* on speech compre-
310 hension ($OR = 0.97$, $SE = 0.07$, $p = .602$).



311

312 **Figure 3.** Significant main effects predicting speech comprehension performance. The generalized lin-
313 ear mixed effects model revealed a negative main effect of syllabic rate (A) and positive main effects of
314 auditory-motor synchronization (B), spontaneous speech motor production rate (C), and working
315 memory score (D). For stimulus perplexity we observed a negative main effect (E). In all panels, error
316 shades indicate 95% confidence intervals. Note that the predictors are shown as a function of syllabic
317 rate for visualization purposes only.

318 *Linguistic predictability and further linguistic variables*

319 To account for the effect of linguistic attributes, we expanded the GLMM by adding several
320 (information-theoretic) linguistic variables: *perplexity*, *probability of target words*, *target dis-*
321 *tance*, and *stimulus length*. Adding these variables (with linguistic variables, AIC: 12675) im-
322 proved model fit (without linguistic variables, AIC: 12848), as measured by a likelihood ratio
323 test ($\chi^2(6) = 184.24$, $p < .001$, see Supplementary Table 4).

324 The full GLMM revealed that *perplexity* had a statistically reliable, negative effect on speech
325 comprehension ($OR = 0.84$, $SE = 0.04$, $p = .001$, Fig. 3E) such that sentences with lower
326 perplexity (which is equal to higher sentence predictability) lead to better speech comprehen-
327 sion performance. Additionally, we observed significant negative effects for *probability of target*
328 *word 1* ($OR = 0.93$, $SE = 0.03$, $p = .026$) and *target word 2* ($OR = 0.92$, $SE: 0.03$, $p = .021$).
329 Contrary to the perplexity effect, this suggests that task performance in the comprehension
330 task was increased for unexpected target words.

331 Furthermore, the model revealed a positive effect for *target distance* ($OR = 1.48$, $SE: 0.05$, p
332 $< .001$), suggesting that larger distance between targets was associated with better speech
333 comprehension performance. In contrast, suggesting the opposite relation, for *stimulus length*
334 we observed a negative effect ($OR = 0.61$, $SE: 0.03$, $p < .001$), i.e., shorter sentences resulted
335 in higher comprehension performance. Due to the large number of variables introduced for
336 nuisance control, we applied a control for multiple comparisons (i.e. false discovery rate; for
337 full results see Supplementary Table 5). All effects remained robust after FDR correction: syl-
338 labic rate: $p < .001$; spontaneous speech motor production rate: $p = .023$; preferred auditory
339 rate: $p = .078$; working memory score: $p = .022$; perplexity: $p = .003$; probability target 1: $p =$
340 $.034$; probability target 2: $p = .030$; compression: $p < .001$; sentence length: $p < .001$; target
341 distance: $p < .001$. Only auditory-motor synchronization changed from a significant effect to a
342 trend ($p = 0.057$) (Note that this was a planned comparison and therefore is discussed).

343 Finally, we explored interaction effects between *syllabic rate*, *auditory-motor synchronization*,
344 and *perplexity*. Adding the interaction term improved model fit ($\chi^2(3) = 13.84$, $p = .004$ (AIC
345 without interaction term: 12675, AIC with interaction term: 12668)). The model revealed two

346 significant 2-way interaction effects: *syllabic rate* × *perplexity* ($OR = 0.88$, $SE = 0.05$, $p = .015$)
347 and *auditory-motor synchronization* × *perplexity* ($OR = 0.86$, $SE = 0.04$, $p = .003$; see Supple-
348 mentary Fig. 5 and Supplementary Table 6). The interaction effect between *syllabic rate* and
349 *perplexity* indicates that particularly comprehension of sentences at fast syllabic rates im-
350 proves when *perplexity* is low. Furthermore, the *auditory-motor synchronization* × *perplexity*
351 interaction effect suggests that while having better overall speech comprehension, high syn-
352 chronizers show a stronger effect of *perplexity* compared to low synchronizers, with even better
353 speech comprehension for more predictable sentences. The *syllabic rate* × *auditory-motor*
354 *synchronization* effect ($OR = 0.94$, $SE = 0.07$, $p = .392$), as tested before, and the three-way
355 interaction effect of *syllabic rate* × *auditory-motor interaction* × *perplexity* ($OR = 1.09$, $SE =$
356 0.06 , $p = .106$) did not show a statistically reliable effect on speech comprehension.

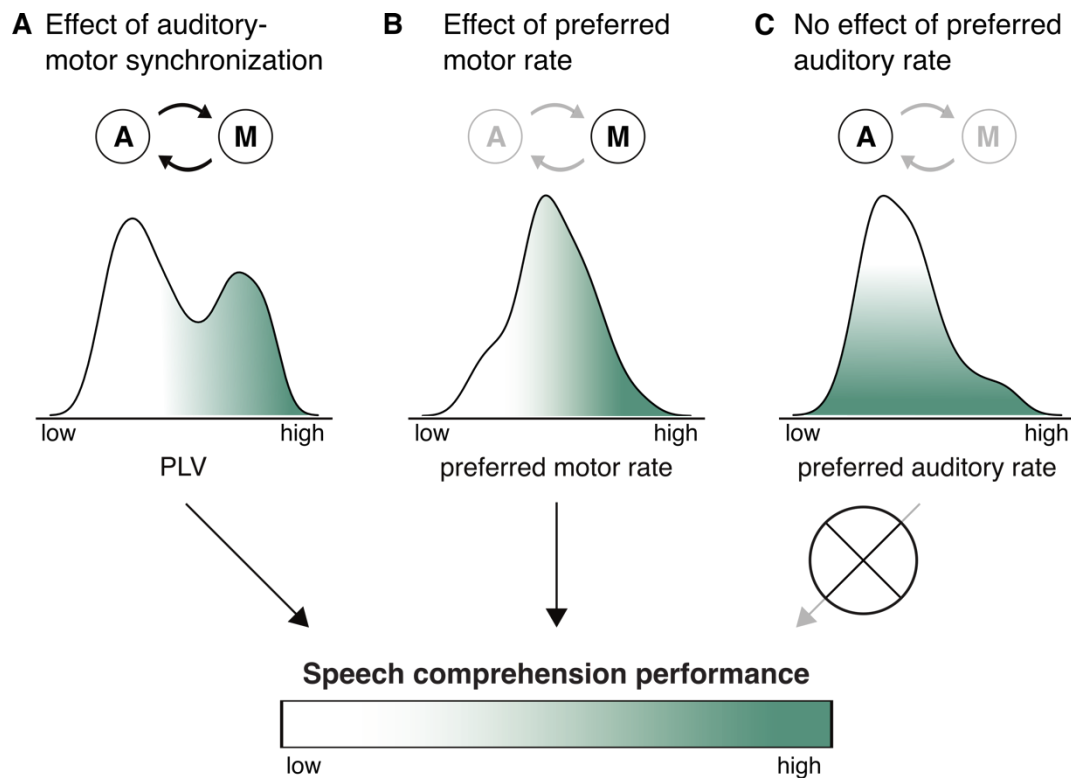
357 *Control experiment*

358 In Experiment 2, speech comprehension performance was exceptionally good, even at high
359 syllabic rates. To ensure the high performance was not an artifact of the task or stimuli, we
360 conducted a control experiment. The analysis revealed that *word order index* did not influence
361 speech comprehension in a statistically meaningful way ($OR = 0.96$, $SE = 0.07$, $p = .219$, see
362 Supplementary Table 7).

363 4. Discussion

364 In two behavioral experiments, we show clear effects of *syllabic rate* on the comprehension of
365 continuous speech. This finding is in line with proposals of speech comprehension being tem-
366 porally constrained such that it is optimal for speech at lower syllabic rates. Crucially, in both
367 protocols we observed that speech comprehension across a wide range of frequencies (5-15
368 syllables/s) was affected by participants' *spontaneous speech motor production rate*, with
369 higher rates predicting better speech comprehension. In the second experiment we showed
370 that, beyond the spontaneous rate of the speech motor system, the individual strength of
371 speech auditory-motor synchronization also affected comprehension. In contrast, the preferred
372 speech perception rate was not related to speech comprehension performance. Together,
373 these findings suggest that while speech comprehension is limited by general processing char-
374 acteristics of the auditory system, interindividual differences in comprehension flexibility are
375 moderated by the motor system and interactions between the auditory and motor systems (Fig.
376 4). Our findings furthermore allow us to generalize the effects of individual differences in the

377 motor system on auditory perception, which have been previously shown for simpler stim-
378 ulti(6,10,40,79), to more natural continuous speech.



379

380 **Figure 4.** Schematic illustrating the relationship between speech comprehension performance and au-
381 ditory-motor synchronization (A), the preferred rates of the motor (B) and of the auditory systems (C).
382 All three predictor variables are represented by the corresponding distribution generated from our ex-
383 perimental data. We propose that better speech comprehension at demanding rates - and by hypothesis
384 auditory behavior more generally - is accompanied by a higher preferred rate of the motor system as
385 well as stronger auditory-motor synchronization. In contrast, the preferred rate of the auditory system
386 seems not to determine auditory behavior. Circled A and M illustrate the auditory and motor systems.
387 The arrows connecting them express the relevance of synchronization between the systems for the
388 variable in question.

389 As expected(2,18,55–57), we observed that speech comprehension accuracy declined as syl-
390 labic rate increased. Although speech comprehension dropped at higher rates in both para-
391 digms, the overall level of comprehension accuracy was much higher in Experiment 2, with
392 accuracy remaining very high (~85%), even for speech as fast as 15 syllables/s. In contrast,
393 in Experiment 1 the increase in syllabic rate resulted in a dramatic drop of comprehension
394 performance. This is in line with our expectations, as the nature of the word-order task is likely
395 to yield overall better performance than the classic intelligibility task. Additionally, our control
396 experiment rules out a potential confound by demonstrating that the high performance in

397 Experiment 2 is not due to simple guessing of the correct word order (see Results section and
398 Supplementary Table 7). Interestingly, however, in both experiments performance decreased
399 later than previously observed, that is, beyond rates of 9 syllables/s(56,80). However, in line
400 with our findings, several other studies, also observed shallower decreases in speech compre-
401 hension, with relatively high comprehension at higher syllable rates (~12 sylla-
402 bles/s)(3,56,81,82). We consider several possible explanations for these discrepancies. One
403 explanation for the different and higher speech-rate decline in comprehension performance is
404 that naturally produced fast speech (with matched degrees of compression across syllabic
405 rates, as used in Experiment 2), in contrast to linearly compressed speech, results in more
406 variance of the speech rate and thus allows for part of the sentences to be understood. How-
407 ever, this explanation does not account for Experiment 1, in which all stimuli were synthesized
408 at the same rate (varying in degrees of compression). Furthermore, the high performance level
409 might be related to different complexity between more naturalistic sentences, providing
410 stronger context information to compensate loss of information, as compared to the words(18),
411 digits(83), or simple sentences(55) used in previous work. Finally, it is notable that while some
412 studies conceptualized the syllabic rate based on the ‘theta-syllable’ (an information unit defi-
413 ned by cortical function(84)), we define syllabic rate as linguistically defined syllables per sec-
414 ond, following other studies(36).

415 Auditory-motor speech synchronization, a behavioral estimate of auditory-motor cortex cou-
416 pling strength(40), had a modulatory -albeit small- effect on speech comprehension. We ob-
417 served that high compared to low synchronizers exhibited better speech comprehension per-
418 formance. These results expand on findings which showed superior statistical word learn-
419 ing(40) or syllable discrimination(6) for individuals with stronger auditory-motor coupling by
420 showing a similar effect for comprehending more naturalistic, continuous speech. Note that
421 this effect requires further validation as it did not survive control for multiple comparisons (Sup-
422 plementary Table 5). Additionally, we expected an interaction of syllabic rate and auditory-
423 motor synchronization, as reported for rate discrimination in tone sequences(10). However,
424 the modulation observed here occurred across all syllabic rates, suggesting that an interaction
425 effect may be masked and compensated for by context and linguistic information in continuous
426 speech comprehension. Alternatively, it is possible -although unlikely- that the interaction of
427 syllabic rate and auditory-motor synchronization was not observed here due to the different
428 frequency resolution at low frequencies. The difference between HIGHS and LOWS in Kern et

429 al.(10) manifested between 7.14 and 10.29 Hz. In contrast, in the present experiment, there
430 was no frequency condition between 5 and 10.69 syllables/s.

431 Importantly, the spontaneous motor production rate affected speech comprehension, suggest-
432 ing that individuals with higher spontaneous motor production rate have increased speech
433 comprehension abilities (at the higher range). We replicated this finding in the second experi-
434 ment. The finding might reflect a complex interplay of auditory and motor cortex during speech
435 comprehension wherein not only the coupling strength, but also the preferred rates of the motor
436 cortex affect speech perception. A possible role of the preferred speech motor rate for speech
437 processing has been previously discussed(35). Furthermore, our findings are in line with an
438 oscillatory model of speech comprehension(6). An alternative interpretation of our findings
439 might be that general processes such as vigilance and fatigue are equally reflected in the
440 spontaneous speech motor production rate and the speech comprehension performance. This
441 could be, because speech comprehension naturally is tightly intertwined with production, and
442 vigilance effects on production for example might similarly affect comprehension. The behav-
443 ioral protocol does not allow to completely discard such an alternative interpretation. However,
444 given that no correlation of a demanding cognitive task (Digit span) with the spontaneous
445 speech motor production rates was observed (see Supplementary Material), we consider this
446 unlikely. Furthermore, for the effects of speech auditory-motor synchronization on syllable dis-
447 crimination others have ruled out such an interpretation(6).

448 Interestingly, the preferred auditory rate (~5.55 syllables/s) had no effect on speech compre-
449 hension in our study. A possible explanation is that preferred rates in auditory cortex are less
450 flexible compared to preferred rates in motor cortex and thus less prone to individual difference
451 related improvements of speech comprehension. However, comparing the variances of the
452 distribution of preferred auditory ($s^2 = 0.74$) and motor ($s^2 = 0.20$) rates revealed bigger vari-
453 ance in the auditory rate ($F(1,162) = 22.39, p < .001$). Another possibility is that the behavioral
454 estimation of preferred auditory cortex rates were not optimally operationalized. This might
455 also explain the lack of correlation between preferred auditory and spontaneous speech pro-
456 duction rates (see Supplementary Material), which we expected to be correlated. Generally,
457 our behavioral protocol only allows for an indirect assessment of preferred neural rates. Nev-
458 ertheless, behavioral measures have been regarded as proxy for underlying intrinsic brain
459 rhythms(45). Finally, the rates at which speech comprehension decreases are much higher
460 than the preferred auditory and spontaneous speech motor production rates. While the pre-
461 ferred rates were well within the expected range(7,8), the mismatch between maximal

462 comprehension rates and preferred rates was due to the high speech comprehension ability
463 of participants even at high rates.

464 We show that continuous speech comprehension is additionally affected by other higher cog-
465 nitive and linguistic factors. The relevance of linguistic predictability and working memory ca-
466 pacity have been shown in multiple studies(53,54). In agreement with these studies, such cog-
467 nitive variables explained a large amount of variance in speech comprehension. Interestingly,
468 our findings suggest that the facilitatory effect of linguistic predictability is particularly effective
469 at fast rates. Second, we tentatively interpret that facilitation due to linguistic predictability may
470 be used more efficiently from individuals with stronger auditory-motor synchronization. A rele-
471 vant question arising from this is: under what conditions is the impact of the motor system on
472 speech comprehension the strongest? Previous work observed an impact of the motor system
473 on speech comprehension in demanding listening conditions, such as listening to speech in
474 noise (5,33). Our data suggests that this view might extend towards conditions of fast speech
475 (which requires more experiments) or might interact with linguistic predictability.

476 Speech comprehension is a highly predictive process which is affected by different sources of
477 predictions. Here we show that, while speech comprehension is optimal in a preferred auditory
478 temporal regime, the motor-system provides a role for individual flexibility in continuous speech
479 comprehension. Additionally, we report that the well-known facilitatory effects of linguistic pre-
480 dictability on speech comprehension interact with individual differences in the motor system.
481 This sets the stage for future assessments of how predictions from these systems interact and
482 under what circumstances the human brain relies more on one over the other.

483 *Ethics*

484 Experiment 1 was approved by the ethics committee of the School of Social Sciences, Univer-
485 sity of Dundee, UK (No. UoD-SoSS-PSY-UG-2019-88). Procedures for Experiment 2 and the
486 Control Experiment were approved by the Max Planck Society (No. 2017_12).

487 *Data accessibility*

488 Preprocessed data and analysis scripts will be available via OSF upon publication.

489 *Competing interests*

490 We declare we have no competing interests.

491 *Funding*

492 This work was supported by the Max Planck Society.

493 *Acknowledgments*

494 We thank Anna Broggi and Harry Watt for help with data acquisition (Exp. 1), Dr. Klaus Frieler
495 for methodological advice, and Dr. Gregory Hickok for valuable comments on a previous ver-
496 sion of the manuscript.

497

References

- 498 1. Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects.
499 Philosophical Transactions of the Royal Society of London Series B: Biological Sciences. 1992
500 Jun 29;336(1278):367–73.
- 501 2. Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, et al. Temporal En-
502 velope of Time-Compressed Speech Represented in the Human Auditory Cortex. *Journal of*
503 *Neuroscience*. 2009 Dec 9;29(49):15564–74.
- 504 3. Brungart DS, van Wassenhove V, Brandewie E, Romigh G. The effects of temporal
505 acceleration and deceleration on AV speech perception. *AVSP*. 2007;27–34.
- 506 4. Morillon B, Schroeder CE, Wyart V. Motor contributions to the temporal precision of
507 auditory attention. *Nature Communications*. 2014 Dec;5(1).
- 508 5. Stokes RC, Venezia JH, Hickok G. The motor system’s [modest] contribution to speech
509 perception. *Psychon Bull Rev*. 2019 Aug;26(4):1354–66.
- 510 6. Assaneo MF, Rimmele JM, Sanz Perl Y, Poeppel D. Speaking rhythmically can shape
511 hearing. *Nat Hum Behav*. 2021 Jan;5(1):71–82.
- 512 7. Poeppel D, Assaneo MF. Speech rhythms and their neural foundations. *Nat Rev Neu-*
513 *rosci*. 2020 May 6;
- 514 8. Ding N, Patel AD, Chen L, Butler H, Luo C, Poeppel D. Temporal modulations in speech
515 and music. *Neuroscience & Biobehavioral Reviews*. 2017 Oct;81:181–7.
- 516 9. Pellegrino F, Coupé C, Marsico E. Across-Language Perspective on Speech Infor-
517 mation Rate. *Language*. 2011;87(3):539–58.
- 518 10. Kern P, Assaneo MF, Endres D, Poeppel D, Rimmele JM. Preferred auditory temporal
519 processing regimes and auditory-motor synchronization. *Psychon Bull Rev*. 2021 Jun 7;
- 520 11. Teng X, Tian X, Rowland J, Poeppel D. Concurrent temporal channels for auditory pro-
521 cessing: Oscillatory neural entrainment reveals segregation of function at different scales. Jen-
522 sen O, editor. *PLOS Biology*. 2017 Nov 2;15(11):e2000812.
- 523 12. Drake C, Botte MC. Tempo sensitivity in auditory sequences: Evidence for a multiple-
524 look model. *Perception & Psychophysics*. 1993 May;54(3):277–86.
- 525 13. Teng X, Poeppel D. Theta and Gamma Bands Encode Acoustic Dynamics over Wide-
526 Ranging Timescales. *Cerebral Cortex*. 2020 Apr 14;30(4):2600–14.
- 527 14. Viemeister NF. Temporal modulation transfer functions based upon modulation thresh-
528 olds. *The Journal of the Acoustical Society of America*. 1979 Nov;66(5):1364–80.
- 529 15. Giraud AL, Poeppel D. Cortical oscillations and speech processing: emerging compu-
530 tational principles and operations. *Nat Neurosci*. 2012 Apr;15(4):511–7.
- 531 16. Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, et al. Speech rhythms
532 and multiplexed oscillatory sensory coding in the human brain. Poeppel D, editor. *PLoS Biol*.
533 2013 Dec;11(12):e1001752.
- 534 17. Peelle JE, Davis MH. Neural Oscillations Carry Speech Rhythm through to Compre-
535 hension. *Frontiers in Psychology*. 2012;3.
- 536 18. Ghitza O, Greenberg S. On the possible role of brain rhythms in speech perception:
537 Intelligibility of time-compressed speech with periodic and aperiodic insertions of silence.

- 538 Phonetica. 2009;66(1–2):113–26.
- 539 19. Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RSJ, Laufs H. Endoge-
540 nous cortical rhythms determine cerebral specialization for speech perception and production.
541 Neuron. 2007 Dec;56(6):1127–34.
- 542 20. Luo H, Poeppel D. Phase patterns of neuronal responses reliably discriminate speech
543 in human auditory cortex. Neuron. 2007 Jun;54(6):1001–10.
- 544 21. Greenberg S, Carvey H, Hitchcock L, Chang S. Temporal properties of spontaneous
545 speech—a syllable-centric perspective. Journal of Phonetics. 2003 Jul;31(3–4):465–85.
- 546 22. Morillon B, Baillet S. Motor origin of temporal predictions in auditory attention. Proceed-
547 ings of the National Academy of Sciences. 2017 Oct 17;114(42):E8913–21.
- 548 23. Schubotz RI. Prediction of external events with our motor system: towards a new frame-
549 work. Trends in Cognitive Sciences. 2007 May;11(5):211–8.
- 550 24. Grahn JA, Brett M. Rhythm and Beat Perception in Motor Areas of the Brain. Journal
551 of Cognitive Neuroscience. 2007 May 1;19(5):893–906.
- 552 25. Cason N, Astésano C, Schön D. Bridging music and speech rhythm: Rhythmic priming
553 and audio–motor training affect speech perception. Acta Psychologica. 2015 Feb;155:43–50.
- 554 26. Du Y, Zatorre RJ. Musical training sharpens and bonds ears and tongue to hear speech
555 better. Proc Natl Acad Sci USA. 2017 Dec 19;114(51):13579–84.
- 556 27. Grahn JA, Rowe JB. Feeling the Beat: Premotor and Striatal Interactions in Musicians
557 and Nonmusicians during Beat Perception. Journal of Neuroscience. 2009 Jun
558 10;29(23):7540–8.
- 559 28. Cheung C, Hamilton LS, Johnson K, Chang EF. The auditory representation of speech
560 sounds in human motor cortex. eLife. 2016 Mar 4;5:e12577.
- 561 29. Evans S, Davis MH. Hierarchical Organization of Auditory and Motor Representations
562 in Speech Perception: Evidence from Searchlight Similarity Analysis. Cereb Cortex. 2015
563 Dec;25(12):4772–88.
- 564 30. Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neuro-
565 sci. 2007 May;8(5):393–402.
- 566 31. Morillon B, Arnal LH, Schroeder CE, Keitel A. Prominence of delta oscillatory rhythms
567 in the motor cortex and their relevance for auditory and speech perception. Neuroscience &
568 Biobehavioral Reviews. 2019 Dec;107:136–42.
- 569 32. Scott SK, McGettigan C, Eisner F. A little more conversation, a little less action — can-
570 didate roles for the motor cortex in speech perception. Nat Rev Neurosci. 2009 Apr;10(4):295–
571 302.
- 572 33. Wu ZM, Chen ML, Wu XH, Li L. Interaction between auditory and motor systems in
573 speech perception. Neurosci Bull. 2014 Jun;30(3):490–6.
- 574 34. Rogalsky C, Basilakos A, Rorden C, Pillay S, LaCroix AN, Keator L, et al. The Neuro-
575 anatomy of Speech Processing: A Large-scale Lesion Study. Journal of Cognitive Neurosci-
576 ence. 2022 Jun 10;1–21.
- 577 35. Assaneo MF, Poeppel D. The coupling between auditory and motor cortices is rate-
578 restricted: Evidence for an intrinsic speech-motor rhythm. Sci Adv. 2018 Feb;4(2):eaao3842.

- 579 36. Keitel A, Gross J, Kayser C. Perceptually relevant speech tracking in auditory and mo-
580 tor cortex reflects distinct linguistic features. Bizley J, editor. *PLOS Biology*. 2018
581 Mar;16(3):e2004473.
- 582 37. Keitel A, Gross J. Individual human brain areas can be identified from their character-
583 istic spectral activation fingerprints. Engel AK, editor. *PLOS Biology*. 2016
584 Jun;14(6):e1002498.
- 585 38. Lubinus C, Orpella J, Keitel A, Gudi-Mindermann H, Engel AK, Roeder B, et al. Data-
586 Driven Classification of Spectral Profiles Reveals Brain Region-Specific Plasticity in Blindness.
587 *Cerebral Cortex*. 2021 Mar 31;31(5):2505–22.
- 588 39. Rosanova M, Casali A, Bellina V, Resta F, Mariotti M, Massimini M. Natural Frequen-
589 cies of Human Corticothalamic Circuits. *Journal of Neuroscience*. 2009 Jun 17;29(24):7679–
590 85.
- 591 40. Assaneo MF, Ripollés P, Orpella J, Lin WM, de Diego-Balaguer R, Poeppel D. Spon-
592 taneous synchronization to speech reveals neural mechanisms facilitating language learning.
593 *Nat Neurosci*. 2019 Apr;22(4):627–32.
- 594 41. Park H, Ince RAA, Schyns PG, Thut G, Gross J. Frontal Top-Down Signals Increase
595 Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners.
596 *Current Biology*. 2015 Jun;25(12):1649–53.
- 597 42. Haegens S, Zion Golombic E. Rhythmic facilitation of sensory processing: A critical
598 review. *Neuroscience & Biobehavioral Reviews*. 2018 Mar;86:150–65.
- 599 43. Rimmele JM, Morillon B, Poeppel D, Arnal LH. Proactive sensing of periodic and aperi-
600 odic auditory patterns. *Trends in Cognitive Sciences*. 2018 Oct;22(10):870–82.
- 601 44. Assaneo MF, Rimmele JM, Orpella J, Ripollés P, de Diego-Balaguer R, Poeppel D.
602 The lateralization of speech-brain coupling is differentially modulated by intrinsic auditory and
603 top-down mechanisms. *Front Integr Neurosci*. 2019 Jul 17;13:28.
- 604 45. McPherson T, Berger D, Alagapan S, Fröhlich F. Intrinsic Rhythmicity Predicts Syn-
605 chronization-Continuation Entrainment Performance. *Scientific Reports*. 2018 Dec;8(1).
- 606 46. Rimmele JM, Kern P, Lubinus C, Frieler K, Poeppel D, Assaneo MF. Musical Sophisti-
607 cation and Speech Auditory-Motor Coupling: Easy Tests for Quick Answers. *Front Neurosci*.
608 2022 Jan 4;15:764342.
- 609 47. Emmorey K, Giezen MR, Petrich JAF, Spurgeon E, O’Grady Farnady L. The relation
610 between working memory and language comprehension in signers and speakers. *Acta Psy-
611 chologica*. 2017 Jun;177:69–77.
- 612 48. Arnal LH, Wyart V, Giraud AL. Transitions in neural oscillations reflect prediction errors
613 generated in audiovisual speech. *Nat Neurosci*. 2011 Jun;14(6):797–801.
- 614 49. Zhang Y, Frassinelli D, Tuomainen J, Skipper JI, Vigliocco G. More than words: word
615 predictability, prosody, gesture and mouth movements in natural language comprehension.
616 *Proc R Soc B*. 2021 Jul 28;288(1955):20210500.
- 617 50. Lewis AG, Bastiaansen M. A predictive coding framework for rapid neural dynamics
618 during sentence-level language comprehension. *Cortex*. 2015 Jul;68:155–68.
- 619 51. Kuperberg GR, Jaeger TF. What do we mean by prediction in language comprehen-
620 sion? *Language, Cognition and Neuroscience*. 2016 Jan 2;31(1):32–59.

- 621 52. Grant KW, Seitz PF. The recognition of isolated words and words in sentences: Indi-
622 vidual variability in the use of sentence context. *The Journal of the Acoustical Society of Amer-*
623 *ica*. 107.(2):1000–11.
- 624 53. Obleser J, Wise RJS, Alex Dresner M, Scott SK. Functional Integration across Brain
625 Regions Improves Speech Perception under Adverse Listening Conditions. *Journal of Neuro-*
626 *science*. 2007 Feb 28;27(9):2283–9.
- 627 54. Obleser J, Kotz SA. Expectancy Constraints in Degraded Speech Modulate the Lan-
628 guage Comprehension Network. *Cerebral Cortex*. 2010 Mar 1;20(3):633–40.
- 629 55. Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM. Speech
630 comprehension is correlated with temporal response patterns recorded from auditory cortex.
631 *Proceedings of the National Academy of Sciences*. 2001 Nov;98(23):6.
- 632 56. Dupoux E, Green K. Perceptual adjustment to highly compressed speech. *Journal of*
633 *Experimental Psychology: Human Perception and Performance*. 1997;23(3):914–27.
- 634 57. Pefkou M, Arnal LH, Fontolan L, Giraud AL. θ -Band and β -Band Neural Activity Re-
635 flects Independent Syllable Tracking and Comprehension of Time-Compressed Speech. *The*
636 *Journal of Neuroscience*. 2017 Aug 16;37(33):7930–8.
- 637 58. Beukelman DR, Yorkston KM. The relationship between information transfer and
638 speech intelligibility of dysarthric speakers. *Journal of Communication Disorders*. 1979
639 Jan;12(3):189–96.
- 640 59. Schiavetti N, Sitler RW, Metz DE, Houde RA. Prediction of Contextual Speech Intelligi-
641 bility from Isolated Word Intelligibility Measures. *J Speech Lang Hear Res*. 1984
642 Dec;27(4):623–6.
- 643 60. Schiavetti N. 1. Scaling procedures for the measurement of speech intelligibility. In:
644 Kent RD, editor. *Studies in Speech Pathology and Clinical Linguistics*. Amsterdam: John Ben-
645 jamins Publishing Company; 1992. p. 11.
- 646 61. Lizcano-Cortés F, Gómez-Varela I, Mares C, Wallisch P, Orpella J, Poeppel D, et al.
647 Speech-to-Speech Synchronization protocol to classify human participants as high or low au-
648 ditory-motor synchronizers. *STAR Protocols*. 2022 Jun;3(2):101248.
- 649 62. Richardson JTE. Measures of Short-Term Memory: A Historical Review. *Cortex*. 2007
650 Jan;43(5):635–50.
- 651 63. Olsthoorn NM, Andringa S, Hulstijn JH. Visual and auditory digit-span performance in
652 native and non-native speakers. *International Journal of Bilingualism*. 2014 Dec;18(6):663–73.
- 653 64. de Jong NH, Wempe T. Praat script to detect syllable nuclei and measure speech rate
654 automatically. *Behavior Research Methods*. 2009 May;41(2):385–90.
- 655 65. Boersma P, Weenik D. Praat: doing phonetics by computer [Computer program]. 2020
656 Aug;Version 6.0.40(retrieved from <http://www.praat.org/>).
- 657 66. Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. Measuring phase synchrony in brain
658 signals. *Human brain mapping*. 1999;8(4):194–208.
- 659 67. Smith ZM, Delgutte B, Oxenham AJ. Chimaeric sounds reveal dichotomies in auditory
660 perception. *Nature*. 2002 Mar;416(6876):87–90.
- 661 68. MacQueen J. Some methods for classification and analysis of multivariate observa-
662 tions. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*.

- 663 1967 Jun;1(14):281–97.
- 664 69. Jurafsky D, Martin JH. Speech and language processing: an introduction to natural
665 language processing, computational linguistics, and speech recognition. 2. ed. [Nachdr.]. Up-
666 per Saddle River, NJ: Prentice Hall; 2009. 988 p.
- 667 70. Chien JT, Ku YC. Bayesian Recurrent Neural Network for Language Modeling. IEEE
668 Trans Neural Netw Learning Syst. 2016 Feb;27(2):361–74.
- 669 71. Mikolov T. Recurrent neural network based language model. Eleventh annual confer-
670 ence of the international speech communication association. 2010;24.
- 671 72. Merity S, Keskar NS, Socher R. Regularizing and Optimizing LSTM Language Models
672 [Internet]. arXiv; 2017 [cited 2022 Oct 25]. Available from: <http://arxiv.org/abs/1708.02182>
- 673 73. Fernandez J, Downey D. Sampling Informative Training Data for RNN Language Mod-
674 els. In: Proceedings of ACL 2018, Student Research Workshop [Internet]. Melbourne, Aus-
675 tralia: Association for Computational Linguistics; 2018 [cited 2022 Oct 25]. p. 9–13. Available
676 from: <http://aclweb.org/anthology/P18-3002>
- 677 74. Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random ef-
678 fects for subjects and items. Journal of Memory and Language. 2008 Nov;59(4):390–412.
- 679 75. Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hy-
680 pothesis testing: Keep it maximal. Journal of Memory and Language. 2013 Apr;68(3):255–78.
- 681 76. Sassenhagen J, Alday PM. A common misapplication of statistical inference: Nuisance
682 control with null-hypothesis significance tests. Brain and Language. 2016 Nov;162:42–5.
- 683 77. Schielzeth H. Simple means to improve the interpretability of regression coefficients:
684 *Interpretation of regression coefficients*. Methods in Ecology and Evolution. 2010
685 Jun;1(2):103–13.
- 686 78. Fox J, Weisberg S. An R companion to applied regression. 3rd ed. Thousand Oaks CA:
687 Sage Publications; 2019.
- 688 79. Assaneo MF, Orpella J, Ripollés P, Noejovich L, López-Barroso D, de Diego-Balaguer
689 R, et al. Population-level differences in the neural substrates supporting Statistical Learning.
690 bioRxiv; 2020 Jul.
- 691 80. Ghitza O. Behavioral evidence for the role of cortical theta oscillations in determining
692 auditory channel capacity for speech. Front Psychol. 2014 Jul 4;5.
- 693 81. Giroud J, Lrousseau JP, Pellegrino F, Morillon B. The channel capacity of multilevel
694 linguistic features constrains speech comprehension [Internet]. bioRxiv; 2021 Dec. Available
695 from: <http://biorxiv.org/lookup/doi/10.1101/2021.12.08.471750>
- 696 82. Verschuere E, Gillis M, Decruy L, Vanthornhout J, Francart T. Speech Understanding
697 Oppositely Affects Acoustic and Linguistic Neural Tracking in a Speech Rate Manipulation
698 Paradigm. J Neurosci. 2022 Sep 28;42(39):7442–53.
- 699 83. Doelling KB, Arnal LH, Ghitza O, Poeppel D. Acoustic landmarks drive delta–theta os-
700 cillations to enable speech comprehension by facilitating perceptual parsing. NeuroImage.
701 2014 Jan;85:761–8.
- 702 84. Ghitza O. The theta-syllable: a unit of speech information defined by cortical function.
703 Front Psychol. 2013;4.