1   **Title:** Lexical influences on categorical speech perception are driven by a temporoparietal circuit

2

3

4

5

6   **Authors**: Gavin M. **Bidelman**[1,2,3*]**,** Claire **Pearson** [2,a], & Ashleigh **Harrison** [2,a]

7
8
9   [1]Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

10   [2]School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA

11   [3]University of Tennessee Health Sciences Center, Department of Anatomy and Neurobiology,
12   Memphis, TN, USA

13   [a] These authors contributed equally to this work

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31   ***Correspondence:**
32   Gavin M. Bidelman, PhD
33   School of Communication Sciences & Disorders
34   University of Memphis
35   4055 North Park Loop
36   Memphis, TN, 38152
37   TEL: (901) 678-5826
38   FAX: (901) 678-2472
39   EMAIL: gmbdlman@memphis.edu
40
41

Abstract

Categorical judgments of otherwise identical phonemes are biased toward hearing words (i.e., "Ganong effect") suggesting lexical context influences perception of even basic speech primitives. Lexical biasing could manifest via late stage post-perceptual mechanisms related to decision or alternatively, top-down linguistic inference which acts on early perceptual coding. Here, we exploited the temporal sensitivity of EEG to resolve the spatiotemporal dynamics of these context-related influences on speech categorization. Listeners rapidly classified sounds from a /gi/ - /ki/ gradient presented in opposing word-nonword contexts (*GIFT-kift* vs. *giss-KISS*), designed to bias perception toward lexical items. Phonetic perception shifted toward the direction of words, establishing a robust Ganong effect behaviorally. ERPs revealed a neural analog of lexical biasing emerging within ~200 ms. Source analyses uncovered a distributed neural network supporting the Ganong including middle temporal gyrus (MTG), inferior parietal lobe (IPL), and middle frontal cortex. Yet, among Ganong-sensitive regions, only left MTG and IPL predicted behavioral susceptibility to lexical influence. Our findings confirm lexical status rapidly constrains sub-lexical categorical representations for speech within several hundred milliseconds but likely does so outside the purview of canonical "auditory-linguistic" brain areas.

Keywords: Auditory event-related potentials (ERPs); categorical perception; Ganong effect; lexical bias; context-dependent coding

1. Introduction

An important building block for language is the ability to transform sensory information into abstract linguistic representations (Goldstone and Hendrickson, 2010). Speech sounds vary continuously across time, environments, speaker identities, and stimulus contexts, and yet, listeners easily parse the speech stream into discrete phonemes (Lotto and Holt, 2016; Phillips, 2001; Pisoni and Luce, 1987). The categorical perception (CP) of speech maps infinitely variable acoustic signals into discrete phonetic-linguistic representations on which the speech-language system can operate (Liberman et al., 1967; Pisoni, 1973; Pisoni and Luce, 1987). CP is indicated when gradually morphed speech sounds along a continuum are heard as belonging to one of a few discrete phonetic classes. Tokens labeled with different identities are said to cross the categorical boundary, a psychological border where listeners' responses abruptly flips due to a perceptual warping of the stimulus space (i.e., compression of within-category sounds) (Best and Goldstone, 2019; Goldstone et al., 2000; Livingston et al., 1998).

One nebulous issue in speech perception concerns whether higher-level activation of lexical representations directly affects sub-lexical components (e.g., phoneme categories). On one extreme is the ridged view that once established, internalized speech prototypes (i.e., equivalence classes or category members) are invariant to superficial stimulus manipulation or lexical context (Liberman et al., 1957). Under this model, categories are impervious to influences from surrounding information

79   and sound elements that precede or follow an isolated stimulus cannot influence its categorization or

80   location of the perceptual boundary. On the contrary, acoustic-phonetic categories—traditionally

81   considered early or lower-level constructs of the speech signal—are in fact highly malleable to

82   contextual variations (Elman and McClelland, 1988; Francis and Ciocca, 2003; Ganong, 1980; Holt

83   and Lotto, 2010; Myers and Blumstein, 2008; Norris et al., 2003; Pisoni, 1975). Moreover, the

84   degree to which context influences the category identity of speech varies with language experience

85   (Bidelman and Lee, 2015; Kuhl et al., 1992; Lively et al., 1993). Consequently, it is now well-

86   established that phonetic categories are flexible and perception of even individual speech features

87   depends critically on the surrounding signal (Repp and Liberman, 1987).

88   　　　　Context-dependent effects in CP are best illustrated by the so-called "Ganong effect"

89   (Ganong, 1980). The Ganong phenomenon occurs when listeners' perceived category boundary of

90   a word-nonword continuum of phonemes shifts (is biased) towards the lexical item. When perceiving

91   a "da-ta" continuum, for example, English-speaking listeners show a stark shift in their perceptual

92   category boundary towards lexical items when one of the gradient's endpoints contains a real word

93   (e.g., "DASH-tash") (Ganong, 1980; Ganong and Zatorre, 1980). Similar interpretive biasing can be

94   induced via learning when listeners are exposed to new contexts that shapes their perception of

95   otherwise isolated sounds (Norris et al., 2003). Collectively, behavioral studies suggest that stimulus

96   context expands the mental category for expected or behaviorally relevant stimuli (McMurray et al.,

97   2008).

98   　　　　One interpretation of lexical effects is that they reflect direct linguistic influence on perceptual

99   processes. Alternatively, another school of thought argues lexical context effects are post-perceptual

100   and are therefore related to executive mechanisms (i.e., response selection, decision). Fox (1984)

101   tested the interaction between lexical knowledge and phonetic categorization during speech

102   perception using Ganong-like stimuli. Lexical status did not influence phonetic categorization at

103   shorter response latencies or when participants were given a response deadline, suggesting lexical

104   context influences later stimulus selection rather than perceptual encoding, *per se*. This notion is

105   supported by results from Pitt and Samuel (1993), who found the strength of lexical influences on

106   perception of ambiguous sound tokens depended on their position in a word—lexical effects were

107   weaker when tokens occurred toward the beginning compared to the end of words. These data

108   support "late stage" or "selection-based" models whereby the very formation of categories

109   themselves only emerge at the late decision-stage of the processing hierarchy (e.g., MERGE model;

110   Norris et al., 2000).

111   　　　　Rather than acting at late stages, lexical biasing could instead manifest via top-down (and

112   perhaps bi-directional) modulations of early perceptual processing with the lexical interface. Indeed,

113   growing evidence from neuroimaging studies (Gow et al., 2008; Myers and Blumstein, 2008; Noe

114   and Fischer-Baum, 2020; van Linden et al., 2007) reaffirms such interactive, connectionist views of

3

115   categorization (e.g., TRACE; McClelland and Elman, 1986). Employing fMRI with a Ganong task,

116   Myers and Blumstein (2008) found that the placement of the phonetic boundary modulated activity

117   both in perceptual [e.g., superior temporal gyrus, STG), inferior parietal lobule (IPL)] and frontal

118   executive brain areas (IFG, ACC), with greater activity for ambiguous items near the boundary. The

119   mere involvement of the STG strongly suggested lexical shifts not due solely to executive decision

120   processes but at minimum, includes a perceptual component that either itself has direct access to

121   lexical properties or is interactively reactivated to integrate phonetic and extra-phonetic factors in

122   placing the phonetic boundary (Gow et al., 2008; Myers and Blumstein, 2008; Noe and Fischer-

123   Baum, 2020). While fMRI offers excellent spatial characterization of potential lexical effects, it lacks

124   the temporal precision necessary to resolve the underlying brain dynamics of category formation

125   (Bidelman et al., 2013) and related lexical influences (Gow et al., 2008), both of which unfold within

126   a few hundred milliseconds after speech onset (e.g., Mahmud et al., 2020).

127   Extending prior neuroimaging work (Gow et al., 2008; Myers and Blumstein, 2008), the aim

128   of the present study was to characterize the spatiotemporal dynamics of context-dependent lexical

129   influences on CP with the goal of establishing where and when speech categories are prone to

130   Ganong-like biasing. We used EEG coupled with source reconstruction to assess the underlying

131   neural bases of phoneme categorization and its lexical modulation. Our task included word-nonword

132   (*GIFT-kift*) and nonword-to-word (*giss-KISS*) acoustic gradients of an otherwise identical /gi/-/ki/

133   acoustic-phonetic continuum designed to bias listeners' perception toward the lexical item and shift

134   their perceptual category boundary (Ganong, 1980; Myers and Blumstein, 2008). Our findings

135   confirm that lexical status rapidly (~200-300 ms) constrains sub-lexical category speech

136   representations but further suggests this interactivity occurs outside canonical auditory-linguistic

137   brain structures. Instead, among Ganong-sensitive brain regions, we find engagement of a

138   temporoparietal circuit (i.e., inferior parietal, middle temporal gyrus) is critical to describing listeners'

139   susceptibility to contextual biasing during category judgements.

2. Materials & Methods

2.1 Participants

142   Sixteen young adults (3 male, 13 female; age: M = 24.5, SD = 12.9 years) were recruited from the

143   University of Memphis student body[1]. Sample size was based on several previous neuroimaging

144   studies on context-effects in CP (e.g., Gow et al., 2008; Myers and Blumstein, 2008). All exhibited

145   normal hearing sensitivity confirmed via audiometric screening (i.e., < 25 dB HL, octave frequencies

146   250 - 8000 Hz). Each participant was strongly right-handed (74.8 ± 27.0% laterality index; Oldfield,

---

[1] EEG was not recorded from one participant due to technical error resulting in a final sample size of n=15 for the neural data (behavioral data were unaffected).

147   1971), had obtained a collegiate level of education (18.8 ± 2.7 years formal schooling), and was a

148   native speaker of American English. Participants were considered nonmusicians (e.g., Mankel and

149   Bidelman, 2018), having on average 3.25 ± 3.3 years of music training. All were paid for their time

150   and gave informed consent in compliance with a protocol approved by the Institutional Review

151   Board at the University of Memphis.

152   **2.2 Speech stimulus continua**

153   Stimuli were adapted from Myers and Blumstein (2008). Speech tokens consisted of a /gi/ to /ki/

154   stop-consonant continuum presented in two word/nonword contexts. Each continuum was

155   constructed using 8 equally spaced voice-onset times (VOTs) incrementing from 18 ms (/g/ percept)

156   to 70 ms (/k/ percept) (**Fig. 1**). This otherwise identical VOT continuum was used to create word-to-

157   nonword (*GIFT-kift*) and nonword-to-word (*giss-KISS*) gradients designed to bias listeners'

158   phonemic perception toward the lexical item (**Fig. 1b**). This was achieved by splicing the appropriate

159   aspiration (i.e., "-ft" for *GIFT-kift;* "-ss" for *giss-KISS*) to the end of the otherwise identical /gi/-/ki/

160   sounds (for details, see Myers and Blumstein, 2008). All tokens were 200 ms in duration and RMS

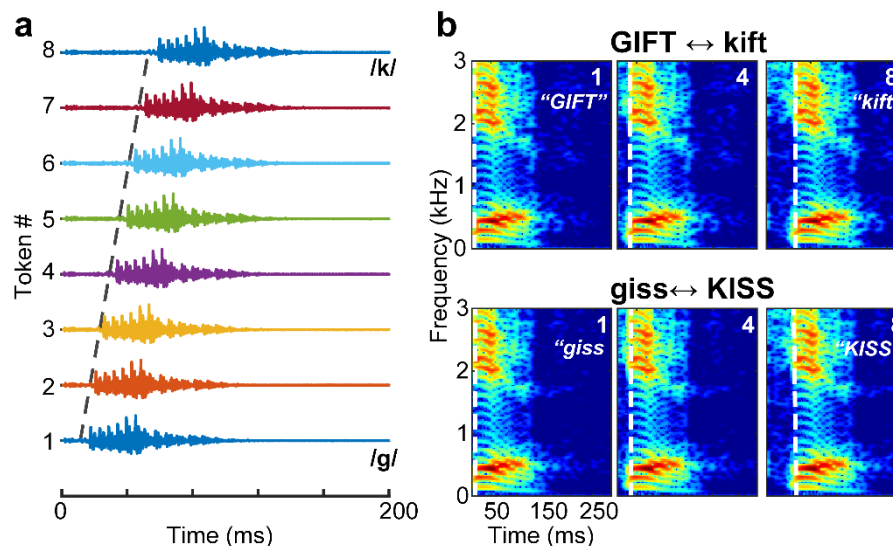161   amplitude normalized.

162

164   **Figure 1:** Speech stimuli used to probe the neural basis of lexical effects on categorical speech processing. (a) Acoustic
165   waveforms of the continuum. Stimuli varied continuously in equidistant VOT steps **to** yield a morphed gradient from /gi/ to
166   /ki/. (b) Spectrograms. The /gi/ to /*ki*/ continuum was presented in one of two word-nonword contexts (GIFT-kift and giss-
167   KISS) such that at any point along the acoustic gradient, the same stop consonant could be perceived more as a word (or
168   nonword) depending on lexical bias from the continuum's endpoint. Dotted lines, onset of voicing demarcating VOT
169   duration.

170

171   During EEG recording, listeners heard 120 trials of each individual token (per context) in

172   which they labelled the sound with a binary response ("g" or "k") as quickly and accurately as

5

173   possible. Following, the interstimulus interval (ISI) was jittered randomly between 800 and 1000 ms

174   (20 ms steps, uniform distribution) to avoid rhythmic entrainment of the EEG and anticipating

175   subsequent stimuli. Block order for the *GIFT-kift* vs. *giss-KISS* continua were randomized within and

176   between participants. The auditory stimuli were delivered binaurally at 79 dB SPL through shielded

177   insert earphones (ER-2; Etymotic Research) controlled by a TDT RP2 signal processor (Tucker

178   Davis Technologies).

2.3 EEG recordings

179

180   EEGs were recorded from 64 sintered Ag/AgCl electrodes at standard 10-10 scalp locations

181   (Oostenveld and Praamstra, 2001). Continuous data were digitized at 500 Hz (SynAmps RT

182   amplifiers; Compumedics Neuroscan) using an online passband of DC-200 Hz. Electrodes placed

183   on the outer canthi of the eyes and the superior and inferior orbit monitored ocular movements.

184   Contact impedances were maintained < 10 kΩ. During acquisition, electrodes were referenced to an

185   additional sensor placed ~ 1 cm posterior to Cz. Data were re-referenced offline to the common

186   average for analysis. Pre-processing was performed in BESA® Research (v7.1) (BESA, GmbH).

187   Ocular artifacts (saccades and blinks) were corrected in the continuous EEG using principal

188   component analysis (PCA) (Picton et al., 2000). Cleaned EEGs were then filtered (1-20 Hz),

189   epoched (-200-800 ms), baselined to the pre-stimulus interval, and ensemble averaged resulting in

190   16 ERP waveforms per participant (8 tokens*2 contexts).

2.4 Behavioral data analysis

191

192   Identification scores were fit with a sigmoid function $P = 1/[1+e^{-\beta 1(x - \beta 0)}]$, where P is the proportion of

193   trials identified as a given vowel, x is the step number along the stimulus continuum, and β_0 and β_1

194   the location and slope of the logistic fit estimated using nonlinear least-squares regression.

195   Comparing parameters between speech contexts revealed possible differences in the "steepness"

196   (i.e., rate of change) and more critically, the location of the categorical boundary as a function of

197   speech context. A lexical bias (i.e., Ganong effect) is indicated when the location of the perceptual

198   boundary (β_0) in phoneme identification shifts dependent on the anchoring speech context (Ganong,

199   1980; Myers and Blumstein, 2008). Behavioral labeling speeds (i.e., reaction times [RTs]) were

200   computed as listeners' median response latency across trials for a given condition. RTs outside 250-

201   2500 ms were deemed outliers (e.g., fast guesses, lapses of attention) and were excluded from the

202   analysis (Bidelman et al., 2013; Bidelman and Walker, 2017).

2.5 EEG data analysis

203

204   ***ERP sensor responses.*** From channel-level waveforms, we measured lexical bias effects in

205   the speech ERPs by comparing scalp topographies at the ambiguous midpoint token (Tk4) evoked

206   in the two different speech contexts (i.e., GIFT4 vs. KISS4). This token step is where lexical bias

207   effects were most prominent behaviorally (see Fig. 2). Topographic t-tests were conducted in

208   EEGLAB (Delorme and Makeig, 2004).

209   ***Source analysis***. To estimate the underlying sources contributing to the lexical effect, we

210   used Classical Low Resolution Electromagnetic Tomography Analysis Recursively Applied (CLARA)

211   [BESA® (v7)] (Iordanov et al., 2014) to estimate the neuronal current density underlying the scalp

212   ERPs (e.g., Alain et al., 2017; Bidelman, 2018). CLARA models the inverse solution as a large

213   collection of elementary dipoles distributed over nodes on a mesh of the cortical volume. The

214   algorithm estimates the total variance of the scalp data and applies a smoothness constraint to

215   ensure current changes minimally between adjacent brain regions (Michel et al., 2004; Picton et al.,

216   1999). CLARA renders more focal source images by iteratively reducing the source space during

217   repeated estimations. On each iteration (x2), a spatially smoothed LORETA solution (Pascual-

218   Marqui et al., 2002) was recomputed and voxels below a 1% max amplitude threshold were

219   removed. This provided a spatial weighting term for each voxel on the subsequent step. Two

220   iterations were used with a voxel size of 7 mm in Talairach space and regularization (parameter

221   accounting for noise) set at 0.01% singular value decomposition. Source activations were visualized

222   on BESA's adult brain template (Richards et al., 2016).

223   ***Dipole source modeling.*** To quantify the time course of source activations, we seeded

224   discrete dipoles within the activation centroids identified in the CLARA volume images at a latency of

225   286 ms, where scalp data showed maximally lexical effects (see Fig. 4a) (cf. Bidelman and Walker,

226   2019). CLARA localized activity to five major foci including middle temporal gyrus (MTG), inferior

227   parietal lobe (IPL), and middle frontal gyrus (MFG) in left hemisphere, and precentral gyrus (PCG)

228   and insular (INS) cortex of right hemisphere (see Fig. 4d). Dipole time courses represent the

229   estimated current within each regional source. We then used this 5 dipole model to create a virtual

230   source montage to transform each participant's scalp potentials (sensor-level recordings) into

231   source space (Scherg et al., 2019; Scherg et al., 2002). This digital re-montaging applied a spatial

232   filter to all electrodes (defined by the foci of our dipole configuration) to transform the electrode

233   recordings to a reduced set of source signals reflecting the neuronal current (in units nAm) as seen

234   *within* each anatomical region of interest (ROI) (Bidelman, 2018; Bidelman et al., 2018). Critically,

235   we fit individual dipole orientations to each participant's own data (anatomical locations remained

236   fixed) to maximize the explained variance of the model at the individual subject level. The model

237   provided a good fit to the grand averaged scalp data (Goodness of fit, entire epoch window = 75%),

238   confirming the ERPs could be described by a restricted number of sources.

239   ***Brain-behavior correspondence***. From the source waveform time courses, we measured

240   peak amplitudes within the 200-300 ms time window, where lexical effects were prominent in raw

241   EEG data (see Fig 4a,b). We then regressed source amplitudes (for each ROI) with listeners'

242   behavioral Ganong effect, computed as the magnitude of shift in their perceptual boundary between

243   speech contexts (i.e., data in Fig. 2c). The allowed us to assess the behavioral relevance of each

244   brain ROI and how context-dependent changes in neural activity (i.e., "neural Ganong" effect) relate

245   to lexical biases in CP measured behaviorally.

2.6 Statistics

247   We analyzed the data using mixed-model ANOVAs in R (R Core team, 2018; *lmer4* package) with

248   fixed effects of token (8 levels) and speech context (2 levels). Subjects served as a random effect.

249   Multiple comparisons were corrected using Tukey-Kramer adjustments. Brain-behavior relations

250   were assessed using robust regression (bisquare weighting) performed using the 'fitlm' function in

251   MATLAB® 2020a (The MathWorks, Inc.).

3. Results

3.1 Behavioral data

254   Behavioral identification functions are shown for the two speech contexts in **Figure 2a**. Listeners

255   more frequently reported /g/ responses in the *GIFT-kift* continuum and more /k/ responses for the

256   *giss-KISS* context*,* confirming that perception for otherwise identical stop consonants is biased

257   toward hearing words. The perceptual boundary location depended strongly on context [$t_{15} = 2.37$, p

258   < 0.0317] (**Fig. 2c** and **2e**). Consistent with prior studies (Ganong, 1980; Myers and Blumstein,

259   2008; Noe and Fischer-Baum, 2020), context-dependent effects in CP where most evident near the

260   ambiguous midpoint of the continuum (Tk 4), where listeners identification abruptly shifted phoneme

261   categories [$t_{15} = 6.00$, $p < 0.0001$] (**Fig. 2d**). Ganong shifts also varied across individuals (e.g., Lam

262   et al., 2017), with some listeners showing strong influence to lexical bias and others showing little to

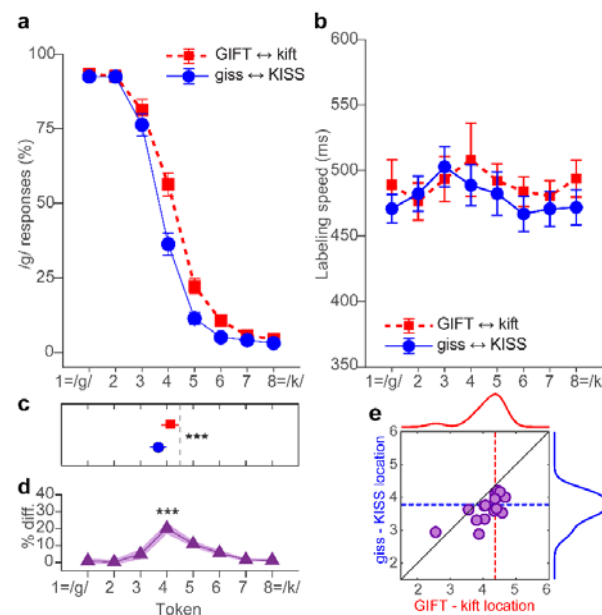263   no changes in perception with speech context (**Fig. 3**).

264

265 **Figure 2:** Lexical context biases the perceptual categorization of speech. (a) Psychometric identification functions show a
266 shift in the perceptual boundary towards lexical items. Listeners more frequently reported /g/ responses in the GIFT-kift
267 continuum and more /k/ responses *for* the giss-KISS context, confirming perception for otherwise identical stop consonants
268 is biased toward hearing words. (b) Reaction times. Labeling speeds are faster for endpoint vs. midpoint tokens of the
269 continuum consistent with category ambiguity near the midpoint of the continuum (Pisoni and Tash, 1974). (c) Critically,
270 the location of the perceptual boundary (i.e., β0) shifts depending on the lexical context. (**d**) Identification performance
271 differs maximally between contexts near the midpoint of the continua (i.e., Tk 4). (e) Comparison of boundary locations
272 (β0) for the GIFT-kift *vs. giss-KISS* continua. The *diagonal* represents the case of an identical perceptual boundary
273 between contexts. Boundaries shift leftward for giss-KISS compared to GIFT-kift, reflecting a *higher precedence of* /k/
274 responses in that context (vice versa for the other context). errorbars = ±1 s.e.m; ***p<0.0001.
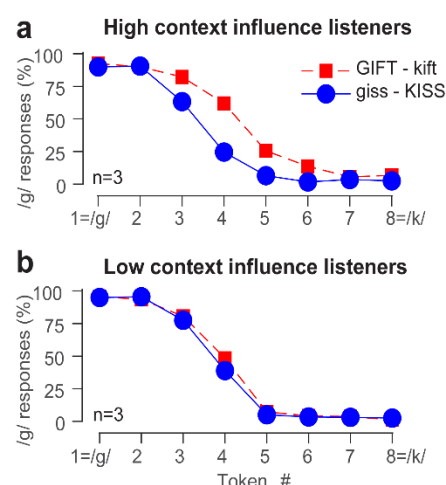
275

276
277

278

279 **Figure 3**: Lexical influences on CP are subject to individual differences. Identification functions for representative listeners
280 (n=3) who showed the strongest (a) and weakest (b) influence of lexical context on speech categorization. High influence
281 listeners perceptual boundary shifts dramatically with context whereas low influence listeners show little change in
282 perception with lexical context.

283

284   Speech labeling speeds were strongly modulated by context [$F_{1,225}$ = 5.15, p=0.024] and

285 token [$F_{7,225}$= 2.14, p=0.0408] (**Fig. 2b**). Identification was faster overall when categorizing tokens in

286 the *giss-KISS* context (p=0.024). *GIFT-kift* responses further showed the hallmark slowing near the

287 midpoint of the continuum relative to prototypical endpoints [Tk 1,2,7,8 vs. Tk 4,5: p =0.024]

288 (Bidelman and Walker, 2017; Pisoni and Tash, 1974), attributable to more ambiguity in decision

289 nearer the perpetual boundary. This categorical RT pattern was not observed for *giss-KISS* (p =

290 0.14). Collectively, these behavioral results suggest that lexical information (words) biases listeners'

291 categorization of otherwise identical phonetic features; even basic phoneme perception is latticed by

292 the surrounding lexical context of the speech signal.

293 **3.2 Electrophysiological data**

294 Scalp ERPs are shown in **Figure 4.** To quantify the "neural Ganong" effect we contrasted ERPs to

295 tokens at the perceptual boundary (i.e., Tk 4) (e.g., Myers and Blumstein, 2008), where lexical bias

9

296   was strongest behaviorally (see Fig 2). Difference waves computed between midpoint tokens

297   evoked during *giss-KISS* vs. *GIFT-kift* continua revealed context-dependent modulations in the time

298   window between 200-300 ms [t_{14} = 3.03, *p*=0.009] (**Fig. 4a, b**). That is, despite identical acoustic

299   information, stop consonants were perceived differentially depending on the word context they

300   carried. The topography of the neural Ganong was broadly distributed over the scalp, spanning

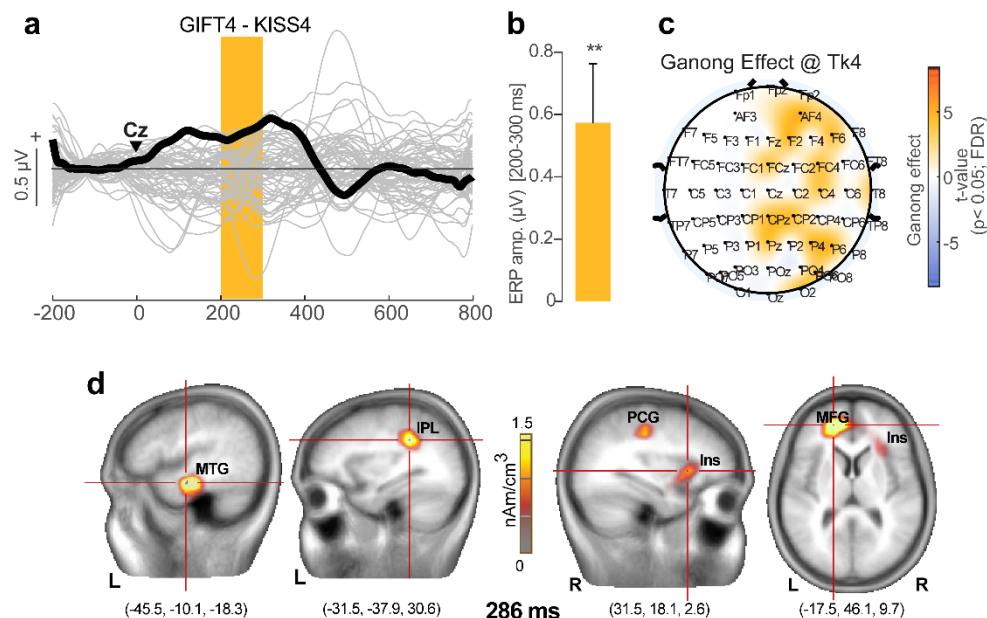301   frontal, temporal, and parietal electrodes (**Fig. 4c**).

302

303   **Figure 4**: Neuroelectric brain activity reveals evidence of lexical biasing on speech categories. (a) Butterfly plot of
304   ERP time courses (bold = Cz electrode) reflecting difference waves between *Tk* 4 responses when presented in GIFT-
305   kift vs. giss-KISS contexts. ▼ = speech stimulus onset. A running t-test (Guthrie and Buchwald, 1991) reveals lexical
306   biasing between 200-300 ms (p<0.05; shaded segment). (b) Mean ERP amplitude at Cz (200-300 ms) differs from 0
307   indicating differentiation of identical speech tokens dependent on lexical context. (c) Topographic distribution of the
308   Ganong effect across the scalp. Statistical maps (paired t-test, p<0.01, FDR-corrected; Benjamini and Hochberg,
309   1995*)* contrasting *Tk4* responses in the two contexts. (d) Brain volumes show *CLARA* (Iordanov et al., 2014)
310   distributed source activation maps underlying lexical bias during speech catherization. Maps were rendered at latency
311   of 286 ms, where the effect was most prominent at the scalp (e.g., Fig 4a). Functional data are overlaid on an adult
312   template brain (Richards et al., 2016). MTG, middle temporal gyrus, IPL, inferior parietal lobe, (MFG) middle frontal
313   gyrus, (PCG) precentral gyrus, (INS) insular cortex. errorbars = ±1 s.e.m.; **p<0.01.
314

315   Source analysis of the ERPs exposed neural activations coding lexical bias in CP within five

316   major foci among the auditory-linguistic-motor loop (e.g., Hickok and Poeppel, 2007; Rauschecker

317   and Scott, 2009), including MTG, IPL (proximal to supramarginal gyrus, SMG), and MFG in left

318   hemisphere, and PCG and INS in right hemisphere (**Fig. 4d**). For each participant, we extracted the

319   time course of source activity from dipoles seeded at the centroids of these ROIs. We then

320   measured and regressed the peak activation within each ROI—reflecting the magnitude of "neural

321   Ganong"—against listeners' behavioral Ganong (i.e., perceptual boundary shift; Fig. 2c). These

322   brain-behavior correlations revealed strong associations between left MTG and left IPL activity and

323   behavioral bias. These findings suggest that context-dependent modulations within a restricted

10

324 temporo-parietal circuit were most inducive to describing the degree to which listeners' CP was

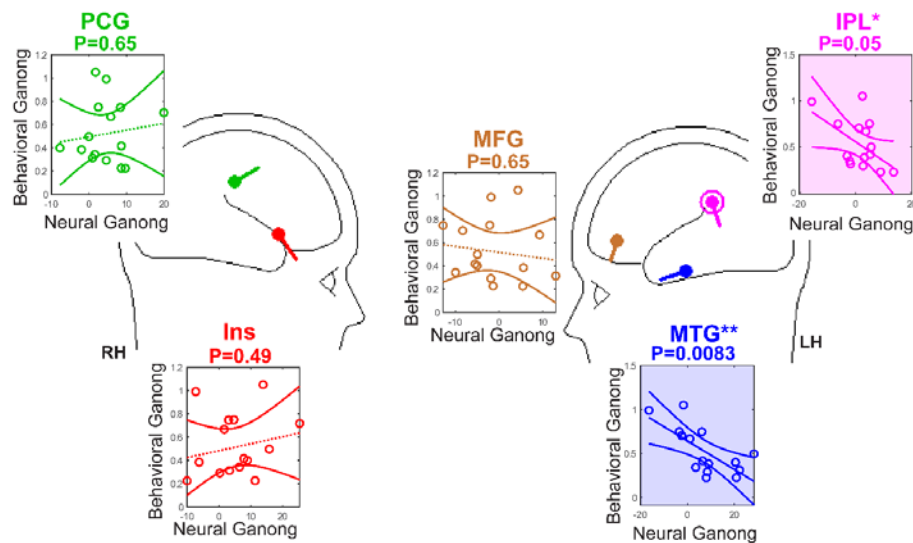325 susceptible to lexical influences.

326

327 **Figure 5:** Lexical bias in CP is driven by engagement of MTG and IPL in left hemisphere. Cartoon heads illustrate the
328 location of the dipole sources underlying the neural Ganong effect. Individual scatters show the relation between neural
329 and behavioral Ganong effect measured from each ROI (shading, p <0.05). Solid regression lines, significant brain-
330 behavior relation; Dotted lines, n.s. Flanking curved lines reflect 95% CIs. Of the active regions, only left MTG and IPL
331 correspond with listeners' behavioral bias. MTG, middle temporal gyrus, IPL, inferior parietal lobe, (MFG) middle frontal
332 gyrus, (PCG) precentral gyrus, (INS) insular cortex, LH/RH, left/right hemisphere. *p<0.05; **p<0.01.

333

4. Discussion

335 By measuring neuroelectric brain activity during rapid speech categorization tasks, our data reveal

336 strong lexical bias in phonetic processing; perception for otherwise identical speech phonemes is

337 attracted toward the direction of words, shifting listeners' categorical boundary dependent on

338 surrounding speech context. We show a neural analog of lexical biasing emerging within ~200 ms

339 within brain activity localized to a distributed, bilateral temporoparietal network including MTG and

340 IPL. Our findings confirm that when perceiving speech, lexical status rapidly constrains sub-lexical

341 representations to their category membership within several hundred milliseconds, establishing a

342 direct linguistic influence on early speech processing.

343   Decoding speech and lexical biasing could be realized via phonetic "feature detectors"

344 (Eimas and Corbit, 1973) that occupy and are differentially sensitive to different segments of the

345 acoustic-linguistic space. Tunable detectors would tend to create quasi "acoustic foveae" that

346 naturally build categories via overrepresentation of the stimulus space near stimulus protypes

347 (Rozsypal et al., 1985). Adaptation studies—in which continuum sounds are presented repetitively

348 and/or in serial order (Eimas and Corbit, 1973; Miller, 1975)—suggest that movement of the

349 category boundary is explained by one detector becoming more desensitized from fatigue, thereby

11

350   causing a boundary shift in the direction toward the un-adapted detector at the polar end of the

351   continuum (Rozsypal et al., 1985). As confirmed empirically, larger boundary shifts would be

352   expected for less strongly categorized continua (Rozsypal et al., 1985), e.g., vowels vs. stop

353   consonants (Altmann et al., 2014), acoustically degraded speech (Bidelman et al., 2020; Bidelman

354   et al., 2019), and for ambiguous speech tokens as shown here and previously (Ganong, 1980; Gow

355   et al., 2008; Lam et al., 2017; Myers and Blumstein, 2008; Noe and Fischer-Baum, 2020).

356   Alternatively, the Ganong-like displacements in perception we observe could occur if linguistic status

357   moves the category boundary toward the most likely lexical candidate. Similarly, nonlinear

358   dynamical models of perception posit that lexical items more strongly activate perceptual "attractor

359   states" which pull auditory percepts toward word items (Tuller et al., 1994). Under this interpretation,

360   the brain would presumably process even non-speech sounds through the lens of a "speech mode,"

361   continually monitoring the auditory stream for lexical units (Liberman et al., 1981; Remez et al.,

362   1981).

363   Considerable debate persists as to whether lexical effects in spoken word recognition result

364   from feedback or feedforward processes (Gow et al., 2008; Myers and Blumstein, 2008; Norris et al.,

365   2000; Pitt, 1995; Samuel and Pitt, 2003). Ganong shifts could occur if lexical knowledge exerts top-

366   down influences to directly affect perceptual processes. Under these frameworks, lexical-based

367   modulation of auditory-sensory brain areas (i.e., STG; Myers and Blumstein, 2008; van Linden et al.,

368   2007) could result from top-input from higher levels associated with wordforms (e.g., SMG, MTG).

369   Alternatively, a purely feedforward architecture (Norris et al., 2000) posits that lexical and phonetic

370   outputs combine and interact at later post-perceptual stages of processing that are intrinsic to overt

371   perceptual tasks (for illustration of these diametric models, see Fig. 1 of Gow et al., 2008; Fig 7:

372   Myers and Blumstein, 2008). In attempts to resolve these conflicting models, Gow et al. (2008) used

373   functional connectivity analyses applied to MEG data and showed that causal neural signaling

374   directed from left SMG to "lower-level" areas (e.g., STG) modulates sensory representations for

375   speech within a latency of 280–480 ms (Gow et al., 2008). The top-down nature of their effects

376   strongly favored a feedback, perceptual account of the Ganong whereby lexical representations

377   influence the earlier encoding of sublexical speech features (e.g., Myers and Blumstein, 2008; Noe

378   and Fischer-Baum, 2020; van Linden et al., 2007).

379   Our EEG findings closely agree with MEG data by demonstrating a neural analog of Ganong

380   biasing that unfolds very early in the chronometry of speech perception. We observed lexical

381   modulation of speech ERPs beginning ~200 ms after sound onset and no later than 300 ms. The

382   early time window (200 ms) of these effects aligns roughly with the P2 wave of the auditory ERPs, a

383   component that is highly sensitive to perceptual object formation, category structure (Bidelman et

384   al., 2020; Bidelman et al., 2013; Bidelman and Walker, 2017; Liebenthal et al., 2010), and context

385   effects in speech identification (Bidelman and Lee, 2015). Source analysis uncovered a Ganong

386 neural circuit spanning five nodes including MTG, IPL, and MFG in left hemisphere and PCG, INS in
387 right hemisphere. The engagement of frontal brain areas (MFG, INS) is consistent with the notion
388 that lexical effects partly evoke post-perceptual, executive processes (Norris et al., 2000). The
389 involvement of insular cortex is perhaps also expected in light of prior imaging work; bilateral inferior
390 frontal activation is particularly evident for speech contrasts that are acoustically ambiguous
391 (Bidelman and Dexter, 2015; Feng et al., 2018; Guediche et al., 2013) and under conditions of
392 increased lexical uncertainty (Bidelman and Walker, 2019; Luthra et al., 2019) that place higher
393 demands on attention (Bouton et al., 2018). Indeed, resolving phoneme ambiguity (as in the
394 Ganong) may be one of the first processes to come online before the decoding of specific lexical
395 features (Gwilliams et al., 2018). This may account for the early time course of our neural effects.
396 　　　　　Notable among the Ganong circuit were nodes in left SMG and MTG. Critically, these
397 regions were the only two areas associated with behavior illustrating their important role in the
398 lexical effect. MTG forms a major component of the ventral speech-language pathway that performs
399 sound-to-meaning inference and acts as a lexical interface linking phonological and semantic
400 information (Hickok and Poeppel, 2007, 2004). MTG is also associated with accessing word
401 meaning (Acheson and Hagoort, 2013), a likely operation in our Ganong task when ambiguous
402 phonemes are perceptually (re)interpreted as words. Relatedly, left inferior parietal lobe an adjacent
403 SMG are strongly recruited during auditory phoneme sound categorization (Desai et al., 2008; Gow
404 et al., 2008; Luthra et al., in press), suggesting their role in phonological coding (Sliwinska et al.,
405 2012). Parietal engagement is especially prominent when speech items are more perceptually
406 confusable (Feng et al., 2018) or require added lexical readout as in Ganong paradigms (Oberfeld
407 and Klöckner-Nowotny, 2016) and may serve as the sensory-motor interface for speech (Hickok et
408 al., 2009; Hickok and Poeppel, 2000). Moreover, using machine learning to decode full brain EEG,
409 we have recently shown that left SMG and related outputs from parietal cortex are among the most
410 salient brain areas that code for category decisions (Al-Fahad et al., 2020; Mahmud et al., 2020).
411 Similar results were obtained in a multivariate pattern decoding analysis of Luthra et al. (in press),
412 who showed left parietal (SMG) and right temporal (MTG) regions were among the most informative
413 for describing moment-to-moment variability in categorization. Additionally, the link between MTG
414 and PCG implied in our data points to a pathway between the neural substrates that map sounds to
415 meaning and sensorimotor regions that execute motor commands (Al-Fahad et al., 2020; Du et al.,
416 2014). Still, the early time course of these neural effects (~250 ms) occurs well before listeners'
417 behavioral RTs (cf. Fig. 2b vs. Fig. 4), suggesting these mechanisms operate at an early perceptual
418 level. These findings lead us to infer that rapid (200-300 ms) context-dependent modulations within
419 a restricted temporo-parietal circuit are most inducive to describing the degree to which listeners are
420 susceptible to lexical influences during speech labeling.

13

421   Notably absent from our Ganong circuit—identified via differences waves—was canonical

422   auditory-linguistic brain regions (e.g., STG). While somewhat unexpected, these data agree with

423   previous fMRI results using a nearly identical Giss-Kiss continuum (Myers and Blumstein, 2008).

424   Indeed, Myers and Blumstein (2008) reported strong IPL but no Ganong-related differences in

425   several brain areas previously shown to be sensitive to phonetic category structure including STG

426   and IFG. STG activity is greater when stimuli are maximally shifted from their VOT-matched

427   counterparts (Myers and Blumstein, 2008). Although we observe a measurable Ganong effect, it is

428   possible that stronger STG differentiation would have been observed in our EEG data with more

429   salient lexical biasing stimuli. Still, the fact that correlations between neural and behavioral Ganong

430   occurred in areas beyond canonical auditory-sensory cortex (e.g., STG) suggests that high order,

431   top-down mechanisms drive or at least dominate lexical biasing (Gow et al., 2008) rather than

432   auditory temporal cortex, *per se.* Though they do so rapidly. Alternatively, rather than a binary

433   feedforward or feedback model of the lexical effect (Gow et al., 2008), it is possible the formation of

434   speech categories operates in near parallel within lower-order (sensory) and higher-order (cognitive-

435   control) brain structures (Mahmud et al., 2020; Toscano et al., 2018). Our data are broadly

436   consistent with such notions. Category representations also need not be isomorphic across the

437   brain. Category formation might reflect a cascade of events where speech units are reinforced and

438   further discretized by a recontact of acoustic-phonetic with lexical representations (Mahmud et al.,

439   2020; Myers and Blumstein, 2008).

440   Our data are best cast in terms of interactive rather than serial frameworks of speech

441   perception as in the TRACE model of spoken word recognition (McClelland and Elman, 1986). As

442   confirmed empirically (Ganong, 1980; Gow et al., 2008; Lam et al., 2017; Myers and Blumstein,

443   2008; Noe and Fischer-Baum, 2020), these models predict stronger lexical biasing when speech

444   sounds carry ambiguity. Indeed, neural correlates of the Ganong effect were most evident at the

445   midpoint of our speech continua, where word influences exert their strongest effect. The very nature

446   of TRACE is that activation traverses from one level to the next before computations at any one

447   stage are complete (McClelland and Elman, 1986). Indeed, available evidence coupled with present

448   results suggest that word recognition could involve simultaneous activation of both continuous

449   acoustic cues and phonological categories (Toscano et al., 2018). It is also possible that the

450   acoustic–phonetic conversion and post-perceptual phonetic decision both localize so the same brain

451   areas (Gow et al., 2008, p.621). Nevertheless, our data show that the acoustic-phonetic encoding of

452   speech is rapidly subject to linguistic influences within several hundred milliseconds. While the early

453   time-course implies a stage of perceptual processing, we find that lexical effects occur strongest

454   outside the purview of canonical auditory-linguistic brain areas via a restricted temporoparietal

455   circuit.

456

Acknowledgments

461 **Figure captions**

462 **Figure 1:** Speech stimuli used to probe the neural basis of lexical effects on categorical speech

463 processing. (**a**) Acoustic waveforms of the continuum. Stimuli varied continuously in equidistant VOT

464 steps to yield a morphed gradient from /gi/ to /ki/. (**b**) Spectrograms. The /gi/ to /ki/ continuum was

465 presented in one of two word-nonword contexts (*GIFT-kift* and *giss-KISS*) such that at any point

466 along the acoustic gradient, the same stop consonant could be perceived more as a word (or

467 nonword) depending on lexical bias from the continuum's endpoint. Dotted lines, onset of voicing

468 demarcating VOT duration.

469 **Figure 2:** Lexical context biases the perceptual categorization of speech. (**a**) Psychometric

470 identification functions show a shift in the perceptual boundary towards lexical items. Listeners more

471 frequently reported /g/ responses in the *GIFT-kift* continuum and more /k/ responses for the *giss-*

472 *KISS* context*,* confirming perception for otherwise identical stop consonants is biased toward

473 hearing words. (**b**) Reaction times. Labeling speeds are faster for endpoint vs. midpoint tokens of

474 the continuum consistent with category ambiguity near the midpoint of the continuum (Pisoni and

475 Tash, 1974). (**c**) Critically, the location of the perceptual boundary (i.e., β_0) shifts depending on the

476 lexical context. (**d**) Identification performance differs maximally between contexts near the midpoint

477 of the continua (i.e., Tk 4). (**e**) Comparison of boundary locations (β_0) for the *GIFT-kift* vs. *giss-KISS*

478 continua. The diagonal represents the case of an identical perceptual boundary between contexts.

479 Boundaries shift leftward for *giss-KISS* compared to *GIFT-kift,* reflecting a higher precedence of /k/

480 responses in that context (vice versa for the other context)*.* errorbars = ±1 s.e.m; ***p<0.0001.

481 **Figure 3**: Lexical influences on CP are subject to individual differences. Identification functions for

482 representative listeners (n=3) who showed the strongest (a) and weakest (b) influence of lexical

483 context on speech categorization. High influence listeners perceptual boundary shifts dramatically

484 with context whereas low influence listeners show little change in perception with lexical context.

485 **Figure 4:** Neuroelectric brain activity reveals evidence of lexical biasing on speech categories. **(a)**

486 Butterfly plot of ERP time courses (bold = Cz electrode) reflecting difference waves between Tk 4

487 responses when presented in *GIFT-kift* vs. *giss-KISS* contexts. ▼ = speech stimulus onset. A

488 running t-test (Guthrie and Buchwald, 1991) reveals lexical biasing between 200-300 ms (*p*<0.05;

489 shaded segment). **(b)** Mean ERP amplitude at Cz (200-300 ms) differs from 0 indicating

490 differentiation of identical speech tokens dependent on lexical context. (**c**) Topographic distribution

491 of the Ganong effect across the scalp. Statistical maps (paired t-test, p<0.01, FDR-corrected;

492 Benjamini and Hochberg, 1995) contrasting Tk4 responses in the two contexts. (**d**) Brain volumes

493 show CLARA (Iordanov et al., 2014) distributed source activation maps underlying lexical bias

16

494   during speech catherization. Maps were rendered at latency of 286 ms, where the effect was most

495   prominent at the scalp (e.g., Fig 4a). Functional data are overlaid on an adult template brain

496   (Richards et al., 2016). MTG, middle temporal gyrus, IPL, inferior parietal lobe, (MFG) middle frontal

497   gyrus, (PCG) precentral gyrus, (INS) insular cortex. errorbars = ±1 s.e.m.; **p<0.01.

498   **Figure 5:** Lexical bias in CP is driven by engagement of MTG and IPL in left hemisphere. Cartoon

499   heads illustrate the location of the dipole sources underlying the neural Ganong effect. Individual

500   scatters show the relation between neural and behavioral Ganong effect measured from each ROI

501   (shading, p <0.05). Solid regression lines, significant brain-behavior relation; Dotted lines, *n.s.*

502   Flanking curved lines reflect 95% CIs. Of the active regions, only left MTG and IPL correspond with

503   listeners' behavioral bias. MTG, middle temporal gyrus, IPL, inferior parietal lobe, (MFG) middle

504   frontal gyrus, (PCG) precentral gyrus, (INS) insular cortex, LH/RH, left/right hemisphere. *p<0.05;

505   **p<0.01.

506

507

508

509 **References**

510 What's in a Word? Morphological Awareness and Vocabulary Knowledge in Three Languages.

511 Al-Fahad, R., Yeasin, M., Bidelman, G.M., 2020. Decoding of single-trial EEG reveals unique states
512 of functional brain connectivity that drive rapid speech categorization decisions. Journal of Neural
513 Engineering 17, 016045.

514 Alain, C., Arsenault, J.S., Garami, L., Bidelman, G.M., Snyder, J.S., 2017. Neural correlates of
515 speech segregation based on formant frequencies of adjacent vowels. Scientific Reports 7, 1-11.

516 Altmann, C.F., Uesaki, M., Ono, K., Matsuhashi, M., Mima, T., Fukuyama, H., 2014. Categorical
517 speech perception during active discrimination of consonants and vowels. Neuropsychologia 64C,
518 13-23.

519 Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful
520 approach to multiple testing. Journal of the Royal Statistical Society, Series B (Methodological) 57,
521 289-300.

522 Best, R.M., Goldstone, R.L., 2019. Bias to (and away from) the extreme: Comparing two models of
523 categorical perception effects. Journal of Experimental Psychology: Learning, Memory, and
524 Cognition 45, 1166-1176.

525 Bidelman, G.M., 2018. Subcortical sources dominate the neuroelectric auditory frequency-following
526 response to speech. Neuroimage 175, 56–69.

527 Bidelman, G.M., Bush, L.C., Boudreaux, A.M., 2020. Effects of noise on the behavioral and neural
528 categorization of speech. Frontiers in Neuroscience 14, 1-13.

529 Bidelman, G.M., Davis, M.K., Pridgen, M.H., 2018. Brainstem-cortical functional connectivity for
530 speech is differentially challenged by noise and reverberation. Hearing Research 367, 149-160.

531 Bidelman, G.M., Dexter, L., 2015. Bilinguals at the "cocktail party": Dissociable neural activity in
532 auditory-linguistic brain regions reveals neurobiological basis for nonnative listeners' speech-in-
533 noise recognition deficits. Brain and Language 143, 32-41.

534 Bidelman, G.M., Lee, C.-C., 2015. Effects of language experience and stimulus context on the
535 neural organization and categorical perception of speech. Neuroimage 120, 191-200.

536 Bidelman, G.M., Moreno, S., Alain, C., 2013. Tracing the emergence of categorical speech
537 perception in the human auditory system. Neuroimage 79, 201-212.

538 Bidelman, G.M., Sigley, L., Lewis, G., 2019. Acoustic noise and vision differentially warp speech
539 categorization. Journal of the Acoustical Society of America 146, 60-70.

540 Bidelman, G.M., Walker, B., 2017. Attentional modulation and domain specificity underlying the
541 neural organization of auditory categorical perception. European Journal of Neuroscience 45, 690-
542 699.

543 Bidelman, G.M., Walker, B.S., 2019. Plasticity in auditory categorization is supported by differential
544 engagement of the auditory-linguistic network. Neuroimage 201, 1-10.

545 Bouton, S., Chambon, V., Tyrand, R., Guggisberg, A.G., Seeck, M., Karkar, S., van de Ville, D.,
546 Giraud, A.L., 2018. Focal versus distributed temporal cortex activity for speech sound category

547 assignment. Proceedings of the National Academy of Sciences of the United States of America 115,
548 E1299-e1308.

549 Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG
550 dynamics including independent component analysis. Journal of Neuroscience Methods 134, 9-21.

551 Desai, R., Liebenthal, E., Waldron, E., Binder, J.R., 2008. Left posterior temporal regions are
552 sensitive to auditory categorization. Journal of Cognitive Neuroscience 20, 1174-1188.

553 Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2014. Noise differentially impacts phoneme
554 representations in the auditory and speech motor systems. Proceedings of the National Academy of
555 Sciences of the United States of America 111, 1-6.

556 Eimas, P.D., Corbit, J.D., 1973. Selective adaptation of linguistic feature detectors. Cognitive
557 Psychology 4, 99-109.

558 Elman, J.L., McClelland, J.L., 1988. Cognitive penetration of the mechanisms of perception:
559 Compensation for coarticulation of lexically restored phonemes. Journal of Memory and Language
560 27, 143-165.

561 Feng, G., Gan, S., Wan, S., Wong, P.C.M., Chandrasekaran, B., 2018. Task-general and acoustic-
562 invariant neural representation of speech categories in the human brain. Cerebral Cortex 28, 3241-
563 3254.

564 Fox, R.A., 1984. Effect of lexical status on phonetic categorization. Journal of Experimental
565 Psychology: Human Perception and Performance 10, 526.

566 Francis, A.L., Ciocca, V., 2003. Stimulus presentation order and the perception of lexical tones in
567 Cantonese. Journal of the Acoustical Society of America 114, 1611-1621.

568 Ganong, W.F., 3rd, 1980. Phonetic categorization in auditory word perception. Journal of
569 Experimental Psychology: Human Perception and Performance 6, 110-125.

570 Ganong, W.F., Zatorre, R.J., 1980. Measuring phoneme boundaries four ways. Journal of the
571 Acoustical Society of America 68, 431-439.

572 Goldstone, R.L., Hendrickson, A.T., 2010. Categorical perception. Wiley Interdisciplinary Reviews:
573 Cognitive Science 1, 69-78.

574 Goldstone, R.L., Steyvers, M., Spencer-Smith, J., Kersten, A., 2000. Interactions between
575 perceptual and conceptual learning. In: Dietrich, E., Markman, A.B. (Eds.), Cognitive dynamics:
576 Conceptual change in humans and machines. Lawrence Erlbaum, Mahwah, NJ, pp. 191–228.

577 Gow, D.W., Jr., Segawa, J.A., Ahlfors, S.P., Lin, F.-H., 2008. Lexical influences on speech
578 perception: a Granger causality analysis of MEG and EEG source estimates. Neuroimage 43, 614-
579 623.

580 Guediche, S., Salvata, C., Blumstein, S.E., 2013. Temporal cortex reflects effects of sentence
581 context on phonetic processing. Journal of Cognitive Neuroscience 25, 706-718.

582 Guthrie, D., Buchwald, J.S., 1991. Significance testing of difference potentials. Psychophysiology
583 28, 240-244.

584 Gwilliams, L., Linzen, T., Poeppel, D., Marantz, A., 2018. In spoken word recognition, the future
585 predicts the past. Journal of Neuroscience 38, 7585-7599.

586 Hickok, G., Okada, K., Serences, J.T., 2009. Area Spt in the human planum temporale supports
587 sensory-motor integration for speech processing. Journal of Neurophysiology 101, 2725-2732.

588 Hickok, G., Poeppel, D., 2000. Towards a functional neuroanatomy of speech perception. Trends in
589 Cognitive Sciences 4, 131-138.

590 Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nature Reviews
591 Neuroscience 8, 393-402.

592 Holt, L.L., Lotto, A.J., 2010. Speech perception as categorization. Attention, Perception, &
593 Psychophysics 72, 1218-1227.

594 Iordanov, T., Hoechstetter, K., Berg, P., Paul-Jordanov, I., Scherg, M., 2014. CLARA: classical
595 LORETA analysis recursively applied. OHBM 2014.

596 Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B., 1992. Linguistic experience
597 alters phonetic perception in infants by 6 months of age. Science 255, 606-608.

598 Lam, B.P.W., Xie, Z., Tessmer, R., Chandrasekaran, B., 2017. The downside of greater lexical
599 influences: Selectively poorer speech perception in noise. Journal of Speech, Language, and
600 Hearing Research 60, 1662-1673.

601 Liberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M., 1967. Perception of the
602 speech code. Psychological Review 74, 431-461.

603 Liberman, A.M., Harris, K.S., Hoffman, H.S., Griffith, B.C., 1957. The discrimination of speech
604 sounds within and across phonemic boundaries. Journal of Experimental Psychology 54, 358-368.

605 Liberman, A.M., Isenberg, D., Rakerd, B., 1981. Duplex perception of cues for stop consonants:
606 Evidence for a phonetic mode. Perception and Psychophysics 30, 133-143.

607 Liebenthal, E., Desai, R., Ellingson, M.M., Ramachandran, B., Desai, A., Binder, J.R., 2010.
608 Specialization along the left superior temporal sulcus for auditory categorization. Cerebral Cortex 20,
609 2958-2970.

610 Lively, S.E., Logan, J.S., Pisoni, D.B., 1993. Training Japanese listeners to identify English /r/ and
611 /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories.
612 Journal of the Acoustical Society of America 94, 1242–1255.

613 Livingston, K.R., Andrews, J.K., Harnad, S., 1998. Categorical perception effects induced by
614 category learning. Journal of Experimental Psychology: Learning, Memory, and Cognition 24, 732-
615 753.

616 Lotto, A.J., Holt, L.L., 2016. Speech perception: The view from the auditory system. Neurobiology of
617 Language. Elsevier, pp. 185-194.

618 Luthra, S., Guediche, S., Blumstein, S.E., Myers, E.B., 2019. Neural substrates of subphonemic
619 variation and lexical competition in spoken word recognition. Language, Cognition and
620 Neuroscience 34, 151-169.

621 Luthra, S.C., Correia, J.M., Kleinschmidt, D.F., Mesite, L., Myers, E.B., in press. Lexical information
622 guides retuning of neural patterns in perceptual learning for speech. Journal of Cognitive
623 Neuroscience https://doi.org/10.1162/jocn_a_01612, 1-12.

624 Mahmud, S., Yeasin, M., Bidelman, G.M., 2020. Data-driven machine learning models for decoding
625 speech categorization from evoked brain responses. bioRxiv [preprint] doi:
626 https://doi.org/10.1101/2020.08.03.234997.

627 Mankel, K., Bidelman, G.M., 2018. Inherent auditory skills rather than formal music training shape
628 the neural encoding of speech. Proceedings of the National Academy of Sciences of the United
629 States of America 115, 13129-13134.

630 McClelland, J.L., Elman, J.L., 1986. The TRACE model of speech perception. Cognitive Psychology
631 18, 1–86.

632 McMurray, B., Dennhardt, J.L., Struck-Marcell, A., 2008. Context effects on musical chord
633 categorization: Different forms of top-down feedback in speech and music? Cognitive Science 32,
634 893-920.

635 Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG
636 source imaging. Clinical Neurophysiology 115, 2195-2222.

637 Miller, J.L., 1975. Properties of feature detectors for speech: Evidence from the effects of selective
638 adaptation on dichotic listening. Perception and Psychophysics 18, 389-397.

639 Myers, E.B., Blumstein, S.E., 2008. The neural bases of the lexical effect: An fMRI investigation.
640 Cerebral Cortex 18, 278-288.

641 Noe, C., Fischer-Baum, S., 2020. Early lexical influences on sublexical processing in speech
642 perception: Evidence from electrophysiology. Cognition 197, 104162.

643 Norris, D., McQueen, J.M., Cutler, A., 2000. Merging information in speech recognition: Feedback is
644 never necessary. Behavioral and Brain Sciences 23, 299-325; discussion 325-270.

645 Norris, D., McQueen, J.M., Cutler, A., 2003. Perceptual learning in speech. Cognitive Psychology
646 47, 204-238.

647 Oberfeld, D., Klöckner-Nowotny, F., 2016. Individual differences in selective attention predict speech
648 identification at a cocktail party. eLife 5.

649 Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh inventory.
650 Neuropsychologia 9, 97-113.

651 Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and
652 ERP measurements. Clinical Neurophysiology 112, 713-719.

653 Pascual-Marqui, R.D., Esslen, M., Kochi, K., Lehmann, D., 2002. Functional imaging with low-
654 resolution brain electromagnetic tomography (LORETA): a review. Methods and Findings in
655 Experimental and Clinical Pharmacology 24 Suppl C, 91-95.

656 Phillips, C., 2001. Levels of representation in the electrophysiology of speech perception. Cognitive
657 Science 25, 711–731.

658   Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J.,
659   Trujillo, N.J., 1999. Intracerebral sources of human auditory-evoked potentials. Audiology and
660   Neuro-Otology 4, 64-79.

661   Picton, T.W., van Roon, P., Armilio, M.L., Berg, P., Ille, N., Scherg, M., 2000. The correction of
662   ocular artifacts: A topographic perspective. Clinical Neurophysiology 111, 53-65.

663   Pisoni, D.B., 1973. Auditory and phonetic memory codes in the discrimination of consonants and
664   vowels. Perception and Psychophysics 13, 253-260.

665   Pisoni, D.B., 1975. Auditory short-term memory and vowel perception. Memory and Cognition 3, 7-
666   18.

667   Pisoni, D.B., Luce, P.A., 1987. Acoustic-phonetic representations in word recognition. Cognition 25,
668   21-52.

669   Pisoni, D.B., Tash, J., 1974. Reaction times to comparisons within and across phonetic categories.
670   Perception and Psychophysics 15, 285-290.

671   Pitt, M.A., 1995. The locus of the lexical shift in phoneme identification. Journal of Experimental
672   Psychology: Learning, Memory, and Cognition 21, 1037-1052.

673   Pitt, M.A., Samuel, A.G., 1993. An empirical and meta-analytic evaluation of the phoneme
674   identification task. Journal of Experimental Psychology: Human Perception and Performance 19,
675   699-725.

676   Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: Nonhuman primates
677   illuminate human speech processing. Nature Neuroscience 12, 718–724.

678   Remez, R.E., Rubin, P.E., Pisoni, D.B., Carrell, T.D., 1981. Speech perception without traditional
679   speech cues. Science 212, 947-949.

680   Repp, B.H., Liberman, A.M., 1987. Phonetic category boundaries are flexible. In: Harnad, S.R. (Ed.),
681   Categorical perception: the groundwork of cognition. Cambridge University Press, New York, pp. 89-
682   112.

683   Richards, J.E., Sanchez, C., Phillips-Meek, M., Xie, W., 2016. A database of age-appropriate
684   average MRI templates. Neuroimage 124, 1254-1259.

685   Rozsypal, A.J., Stevenson, D.C., Hogan, J.T., 1985. Dispersion in models of categorical perception.
686   Journal of Mathematical Psychology 29, 271-288.

687   Samuel, A.G., Pitt, M.A., 2003. Lexical activation (and other factors) can mediate compensation for
688   coarticulation. Journal of Memory and Language 48.

689   Scherg, M., Berg, P., Nakasato, N., Beniczky, S., 2019. Taking the eeg back into the brain: The
690   power of multiple discrete sources. Frontiers in Neurology 10.

691   Scherg, M., Ille, N., Bornfleth, H., Berg, P., 2002. Advanced tools for digital EEG review: Virtual
692   source montages, whole-head mapping, correlation, and phase analysis. Journal of Clinical
693   Neurophysiology 19, 91-112.

694   Sliwinska, M.W., Khadilkar, M., Campbell-Ratcliffe, J., Quevenco, F., Devlin, J.T., 2012. Early and
695   Sustained Supramarginal Gyrus Contributions to Phonological Processing. Frontiers in Psychology
696   3.

697   Toscano, J.C., Anderson, N.D., Fabiani, M., Gratton, G., Garnsey, S.M., 2018. The time-course of
698   cortical responses to speech revealed by fast optical imaging. Brain and Language 184, 32-42.

699   Tuller, B., Case, P., Ding, M., Kelso, J.A.S., 1994. The nonlinear dynamics of speech categorization.
700   Journal of Experimental Psychology: Human Perception and Performance 20, 3-16.

701   van Linden, S., Stekelenburg, J.J., Tuomainen, J., Vroomen, J., 2007. Lexical effects on auditory
702   speech perception: An electrophysiological study. Neuroscience Letters 420, 49-52.

703