1    **Classification:** Biological sciences, Neuroscience

2    **Cortical tracking of speech in noise accounts for reading strategies in children**

3    Florian Destoky[a,*], Julie Bertels[a,b,*], Maxime Niesen[a,c], Vincent Wens[a,d], Marc Vander Ghinst[a],

4    Jacqueline Leybaert[e], Marie Lallier[f], Robin A. A. Ince[g], Joachim Gross[g,h], Xavier De Tiège[a,d],

5    Mathieu Bourguignon[a,e,f]

6    [a]Laboratoire de Cartographie fonctionnelle du Cerveau, UNI – ULB Neuroscience Institute,

7    Université libre de Bruxelles (ULB), Brussels, Belgium.

8    [b]Consciousness, Cognition and Computation group, UNI – ULB Neuroscience Institute,

9    Université libre de Bruxelles (ULB), Brussels, Belgium.

10   [c]Service d'ORL et de chirurgie cervico-faciale, ULB-Hôpital Erasme, Université libre de

11   Bruxelles (ULB), 1070 Brussels, Belgium.

12   [d]Department of Functional Neuroimaging, Service of Nuclear Medicine, CUB Hôpital Erasme,

13   Université libre de Bruxelles (ULB), Brussels, Belgium.

14   [e]Laboratoire Cognition Langage et Développement, UNI – ULB Neuroscience Institute,

15   Université libre de Bruxelles (ULB), Brussels, Belgium.

16   [f]BCBL, Basque Center on Cognition, Brain and Language, 20009 San Sebastian, Spain.

17   [g]Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK.

18   [h]Institute for Biomagnetism and Biosignal analysis, University of Muenster, Muenster,

19   Germany.

20   [*]These authors equally contributed to the article.

21   **Corresponding author**

22   Florian Destoky, Laboratoire de Cartographie fonctionnelle du Cerveau, UNI – ULB

23   Neuroscience Institute, Université libre de Bruxelles, 808 Lennik Street, 1070 Brussels,

24   Belgium. E-mail: florian.destoky@ulb.ac.be. Tel. +32 2 555 3286.

25

26   **Contributions**

27   M.B., F.D., J.B., X.D., designed the study; F.D., J.B., M.N., M.B. collected the data; F.D., J.B.,

28   M.B., R.A.A.I. analyzed the data; F.D. and M.B. wrote the initial version of the manuscript;

29   and all authors discussed the results and commented on the manuscript.

30    **Number of tables:** 5, **number of figures:** 4, **number of pages:** 46, **number of words in the**

31    **Abstract:** 184, **number of words in the main text:** 5620

32    **Conflicts of interest**

33    None of the authors disclose any potential conflict of interest.

34 **Abstract (~150 words; now: 184)**

35    Humans' propensity to acquire literacy relates to several factors, among which, the

36    ability to understand speech in noise (SiN). Still, the nature of the relation between reading and

37    SiN perception abilities remains poorly understood. Here, we dissect the interplay between (i)

38    reading abilities, (ii) classical behavioral predictors of reading (phonological awareness,

39    phonological memory and lexical access), and (iii) electrophysiological markers of SiN

40    perception in 99 elementary school children (26 with dyslexia). We demonstrate that cortical

41    representation of phrasal content of SiN relates to the development of the lexical (but not

42    sublexical) reading strategy. In contrast, classical behavioral predictors of reading abilities and

43    the ability to benefit from visual speech to represent the syllabic content of SiN account for

44    global reading performance (i.e., speed and accuracy of lexical and sublexical reading). Finally,

45    we found that individuals with dyslexia properly integrate visual speech information to

46    optimize processing of syntactic information, but not to sustain acoustic/phonemic processing.

47    These results clarify the nature of the relation between SiN perception and reading abilities in

48    typical and dyslexic child readers, and identified novel electrophysiological markers of

49    emergent literacy.

50

51    **Keywords: Cortical tracking of speech, MEG, Reading abilities, Dyslexia, Speech in noise**

52

53

## Introduction

Acquiring literacy is tremendously important in our societies. Central for reading acquisition are adequate phonological awareness,[1–3] phonological memory,[4,5] and lexical access.[6–8] The adequacy of the learning environment also plays a major role.[9,10] In particular, the presence of recurrent noise in the learning environment can substantially hinder reading acquisition.[11,12] Therefore, the ability to understand speech in noise (SiN)—which is known to differ between individuals [13,14]—should modulate the negative impact of environmental noise on reading acquisition. And indeed, the quality of brainstem responses to syllables in noise predicts reading abilities and its precursors.[15] Moreover, individuals with dyslexia often exhibit a SiN perception deficit,[16,17] that is particularly apparent when the background noise is composed of speech.[18] This deficit has been hypothesized to root in a deficit in phonological awareness,[19,20] but contradictory reports do exist.[21] The question of whether SiN processing abilities relate to reading due to a common dependence on classical behavioral predictors (i.e., phonological awareness, phonological memory and lexical access), or due to other cognitive or neurophysiological processes specific to SiN processing, is thus open. Furthermore, it is also unexplored which aspects of reading and SiN processing abilities are related. Understanding these relations is especially important given that acoustic noise is ubiquitous, and given how adverse dyslexia can be for the cognitive and social development of children.

Reading is a multifaceted process. Hence, it is reasonable to think that SiN processing might relate to a restricted set of aspects of reading. Following the Dual Route Cascaded (DRC) model, reading is supported by two different routes: the *sublexical* and the *lexical* routes.[22,23] The sublexical route implements the grapheme to phoneme conversion. It is used when reading unfamiliar words or pseudowords, but it is not suitable to correctly read irregular words (*i.e., yacht*) and to acquire fluent reading. Skilled reading relies on the lexical route that supports fast recognition of the orthographic word form of familiar words. The lexical route is

79    indispensable to read irregular words, benefits reading of regular words, and does not

80    contribute to reading pseudowords. Remarkably, the brain would implement these two reading

81    strategies in two distinct neural pathways.[24–27]

82        There are also several distinct aspects of SiN processing that could relate to reading,

83    and these can be derived from electrophysiological recordings of brain activity during

84    connected speech listening. When listening to connected speech, human auditory cortical

85    activity tracks the fluctuations of speech temporal envelope at frequencies matching the speech

86    hierarchical linguistic structures, i.e., phrases/sentences (0.2–1.5 Hz) and words/syllables (2–8

87    Hz).[28–38] Such cortical tracking of speech (CTS) is thought to be essential for speech

88    comprehension.[31,33,35,37,39–41] Corresponding brain oscillations would subserve the

89    segmentation or parsing of incoming connected speech to promote speech

90    recognition.[31,32,37,39,42] In SiN conditions, children and adults' brain preferentially tracks the

91    attended speech rather than the global auditory scene, though with reduced fidelity when the

92    noise hinders comprehension.[28,29,38,43–53] Assessing CTS in noise can therefore provide

93    objective measures of the impact of noise on the cortical representation of the different

94    hierarchical linguistic structures of speech. Also relevant is how SiN perception is impacted by

95    noise properties. In essence, the relevant parameters for an acoustic noise in SiN conditions are

96    the degree of energetic and informational masking.[54] The noise is energetic when its spectrum

97    is similar to that of speech, and non-energetic otherwise. The noise is informational when it is

98    made up of other speech signals, and non-informational otherwise. An energetic noise

99    introduces *physical* interferences and an informational noise introduces *perceptual*

100   interferences. Finally, to enhance SiN processing, humans also benefit from visual information

101   of the speaker's articulatory mouth movements.[55,56] All these aspects of SiN perception can be

102   captured by measures of CTS.

103         In this study, we investigated the relations between reading abilities, neural

104     representations of SiN quantified with CTS, and classical behavioral predictors of reading in

105     elementary school children. To fully characterize cortical SiN processing, we measured CTS

106     in several types of background noises introducing different levels of energetic and

107     informational masking and in conditions where the face of the speaker was visible (*lips*) or not

108     (*pics*) while talking. This study was designed to answer three major questions: (i) What aspects

109     of cortical SiN processing and reading abilities are related in typically-developing elementary

110     school children (ii) To what extent are these relations mediated by classical behavioral

111     predictors of reading? (iii) Do these relations translate to alterations in dyslexic children in

112     comparison with typical readers matched for age or reading-level? As preliminary steps to

113     tackle these questions, we identify relevant features of CTS in noise, and assessed in a global

114     analysis the nature of the information about reading brought by all the identified features of

115     CTS in noise and classical behavioral predictors of reading abilities.

116

## Results

118         We first report on 73 children with typical reading abilities. Then, we report on 26

119     children with dyslexia matched with a sub-sample of the 73 typical readers for age (n = 26) or

120     reading level (n = 26). Both control groups were included to tell whether development or

121     reading experience can explain potentially uncovered SiN deficits.[57] Reading performance and

122     its classical behavioral predictors were characterized in a comprehensive cognitive evaluation

123     (Table 1). Children' brain activity was recorded with magnetoencephalography (MEG) while

124     they were attending to 4 videos of ~6 min each (Figure 1). Each video featured 9 conditions: 1

125     noiseless and 8 SiN resulting from the combination of energetic or non-energetic and

126     informational or non-informational noise with *lips* and *pics* visual inputs. For each condition,

127     we regressed the temporal envelope of the attended speech on MEG signals with several time

6

128     lags using ridge regression and cross-validation (see methods section for details).[58] The ensuing

129     regression model was used to reconstruct speech temporal envelope from the recorded MEG

130     signal. CTS was computed as the correlation between the genuine and reconstructed speech

131     temporal envelopes. We did this for MEG and speech envelope signals filtered at 0.2–1.5 Hz

132     (phrasal rate)[28,59] and 2–8 Hz (syllabic rate),[47,51,60,61] and for MEG sensor signals in the left and

133     right hemispheres separately.

134          Table 2 presents the percentage of the 73 typical readers showing statistically

135     significant phrasal and syllabic CTS, for both hemispheres, and each condition. All typical

136     readers showed significant phrasal CTS in noiseless and non-informational (non-speech

137     hereafter) noise conditions, and still most of them in informational (babble hereafter) noise

138     conditions (mean ± SD across conditions, 98.3 ± 2.1 %). Most of the typical readers showed

139     significant syllabic CTS in noiseless and non-speech noise conditions (93.8 ± 3.2 %), and

140     slightly less of them in babble noise conditions (80.1 ± 4.3 %). This result clearly indicates that

141     CTS can be robustly assessed at the subject level.
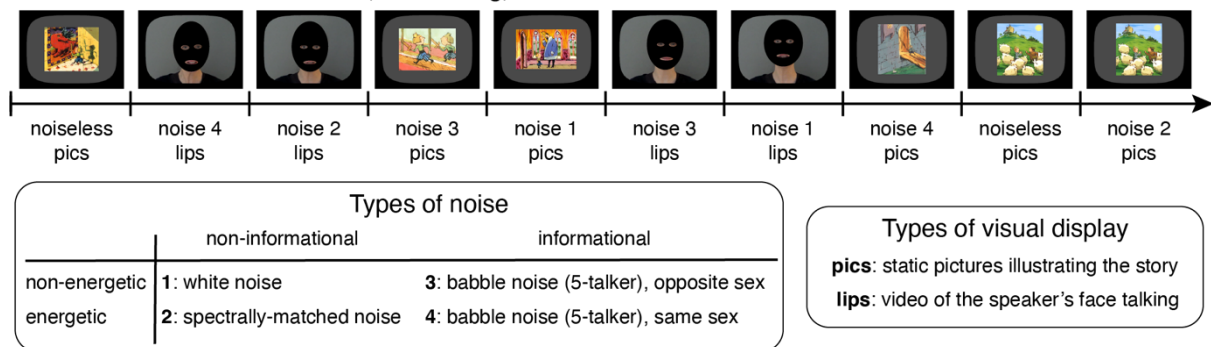


142

143    **Figure 1.** Illustration of the experimental material used in the neuroimaging assessment.

144    Subjects viewed 4 videos of ~6 min duration. Each video was divided into 10 blocks to which

145    experimental conditions were assigned. There were 2 blocks of the *noiseless* condition, and 8

146    blocks of speech-in-noise (SiN) conditions: 1 block for each possible combination of the 4

147    types of noise and two types of visual display. The interference introduced by the noise was

148    either energetic or not and informational or not. The visual display provided visual speech

149    information (*lips*) or not (*pics*).

7

150 **Table 1.** Mean and standard deviation of behavioral scores in each reading group of 26 children

151 and comparisons (*t*-tests) between groups. The number of degrees of freedom (*df*) was 50 for

152 all comparisons except those involving auditory attention (TAP) scores (dyslexic readers vs.

153 controls in age, *df* = 49; dyslexic readers vs. controls in reading level, *df* = 38) and socio-

154 economic status (dyslexic readers vs. controls in age, *df* = 49; dyslexic readers vs. controls in

155 reading level, *df* = 47). IQ, intelligence quotient; RAN, rapid automatized naming; SD, standard

156 deviation.

| | Dyslexic readers | | Age-matched control | | Reading level control | | Dyslexic readers compared with controls | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | in age | | in reading level | |
| | Mean | SD | Mean | SD | Mean | SD | p | t(*df*) | p | t(*df*) |
| Chronological age | 9.81 | 1.26 | 9.52 | 0.96 | 7.21 | 0.53 | 0.38 | 0.89 | **<0.0001** | **10.31** |
| Non-verbal IQ | 111 | 11 | 114 | 10 | 112 | 9 | 0.30 | -1.04 | 0.784 | -0.28 |
| Socio-economic status | 6.12 | 2.44 | 6.96 | 1.45 | 6.96 | 2.47 | 0.14 | -1.50 | 0.17 | -1.40 |
| Alouette reading accuracy | 89.0 | 5.7 | 96.2 | 2.1 | 89.0 | 6.46 | **<0.0001** | **-6.07** | 0.988 | 0.01 |
| Alouette reading speed | 141 | 61 | 292 | 91 | 138 | 64 | **<0.0001** | **-7.04** | 0.867 | 0.17 |
| Irregular words reading [words/s] | 0.54 | 0.33 | 1.16 | 0.44 | 0.40 | 0.35 | **<0.0001** | **-5.82** | 0.15 | 1.47 |
| Regular words reading [words/s] | 0.73 | 0.41 | 1.35 | 0.41 | 0.61 | 0.35 | **<0.0001** | **-5.51** | 0.29 | 1.06 |
| Pseudo-words reading [words/s] | 0.42 | 0.24 | 0.78 | 0.30 | 0.39 | 0.21 | **<0.0001** | **-4.88** | 0.61 | 0.50 |
| Visual attention | 30.3 | 3.74 | 32.0 | 2.69 | 27.4 | 4.43 | 0.070 | -0.95 | **0.014** | **2.53** |
| Phoneme suppression | 7.92 | 2.15 | 9.04 | 1.75 | 8.42 | 1.27 | **0.046** | **-2.05** | 0.313 | -1.02 |
| Phoneme fusion | 7.73 | 1.59 | 9.31 | 0.97 | 8.92 | 1.16 | **<0.0001** | **-4.32** | **0.003** | **-3.09** |
| Forward digit span | 5.08 | 0.84 | 5.8 | 0.69 | 5.15 | 0.78 | **0.001** | **-3.41** | 0.735 | -0.34 |
| Backward digit span | 3.69 | 0.79 | 4.5 | 1.33 | 3.38 | 0.75 | **0.011** | **-2.66** | 0.156 | 1.44 |
| RAN time [s] | 24.4 | 7.84 | 20.1 | 3.02 | 30.6 | 7.51 | **0.013** | **2.59** | **0.005** | **-2.91** |
| TAP mean response time [ms] | 627 | 99.0 | 613 | 75.4 | 667 | 93.4 | 0.59 | 0.53 | **0.07** | **-1.86** |
| TAP SD response time [ms] | 140 | 45.0 | 129 | 30.3 | 171 | 46.7 | 0.33 | 0.98 | **0.02** | **-2.36** |
| TAP correct responses | 15.6 | 0.58 | 15.7 | 0.68 | 15.3 | 1.07 | 0.42 | -0.81 | 0.11 | 1.65 |
| TAP false responses | 2.15 | 2.26 | 0.84 | 1.28 | 1.21 | 0.97 | **0.014** | **2.54** | 0.89 | 0.13 |

157

158 **Table 2.** Percentage of the 73 typical readers showing significant cortical tracking of speech

159 (CTS) at phrasal and syllabic rates in the 9 different conditions, in the left hemisphere (LH), in

160 the right hemisphere (RH), or in at least one hemisphere. The two values provided for the

161 noiseless condition correspond to two arbitrary subdivisions of the *noiseless* data to match the

162 amount of data for the eight noise conditions.

| | Phrasal CTS | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Noiseless | | Non-Informational Noise | | | | Informational Noise | | | |
| | | | Non-energetic | | Energetic | | Non-energetic | | Energetic | |
| | Pics | Pics | Pics | Lips | Pics | Lips | Pics | Lips | Pics | Lips |
| Left hemisphere | 100 | 100 | 100 | 100 | 100 | 100 | 91.8 | 97.3 | 90.4 | 97.3 |
| Right hemisphere | 100 | 100 | 100 | 100 | 100 | 100 | 94.5 | 98.6 | 90.4 | 95.9 |
| At least one hemisphere | 100 | 100 | 100 | 100 | 100 | 100 | 97.3 | 100 | 95.9 | 100 |
| | Syllabic CTS | | | | | | | | | |
| Left hemisphere | 89 | 84.9 | 82.2 | 83.6 | 82.2 | 86.3 | 53.4 | 65.7 | 49.3 | 61.6 |
| Right hemisphere | 91.8 | 89 | 89 | 91.8 | 86.3 | 91.8 | 69.9 | 76.7 | 67.1 | 76.7 |
| At least one hemisphere | 97.3 | 94.5 | 94.5 | 94.5 | 87.7 | 94.5 | 78.1 | 86.3 | 76.7 | 79.4 |

163

8

164 **What aspects of SiN processing modulate the measures of CTS in noise?**

165 First, we identify the main factors modulating CTS in SiN conditions. To that aim, we

166 evaluated with linear mixed-effects modeling how the normalized CTS (nCTS) in SiN

167 conditions depends on hemisphere, noise properties, and visibility of the talker's lips. The

168 nCTS is a contrast between CTS in SiN and noiseless conditions (see Methods) that takes

169 values between –1 and 1, with negative values indicating that the noise reduces CTS. Such

170 contrast presents the advantage of being specific to SiN processing abilities by factoring out

171 the global level of CTS in the noiseless condition. In that analysis, nCTS values were corrected

172 for age, time spent at school and IQ.

173 Table 3 presents the final linear mixed-effects model of phrasal and syllabic nCTS, and

174 Figure 2 illustrates underlying values.

175 The pattern of how nCTS changed with different types of noises was overall similar for

176 phrasal and syllabic nCTS. Non-speech noise did not substantially change CTS (nCTS was

177 close to 0). However, babble noise resulted in a substantial reduction of CTS compared to the

178 noiseless condition for both hemispheres and irrespective of the availability of visual speech

179 information. That is, nCTS in babble noise conditions was roughly between –0.1 and –0.3,

180 indicating that CTS in babble noise was 20–50% (values obtained by inverting the formula of

181 nCTS) lower than CTS in noiseless conditions.

182 Availability of visual speech information (*lips* conditions) increased the level of nCTS

183 only in babble noise conditions for phrasal nCTS, and in all noise conditions for syllabic nCTS.

184 And finally, the noise impacted differently nCTS in the left and right hemispheres. The

185 phrasal nCTS was higher in the left than right hemisphere in babble noise conditions. It was

186 the other way round for syllabic nCTS in all noise conditions.

187 In summary, the CTS is mostly impacted by informational noises, and is also modulated

188 by the availability of visual speech and the hemisphere (only in informational noise conditions

189    for phrasal CTS, and in all noise conditions for syllabic CTS). These observations guided the

190    elaboration of 8 relevant features (contrasts) of nCTS in SiN conditions (see Supplementary

191    Methods): the global level of nCTS and its informational, visual and hemispheric modulations,

192    all for phrasal and syllabic nCTS. In the next sections we unravel the associations between

193    these features, reading abilities, and classical behavioral predictors of reading.

194

195    **Table 3.** Factors included in the final linear mixed-effects model fit to the normalized cortical

196    tracking of speech (nCTS) at phrasal and syllabic rates. Factors are listed in their order of
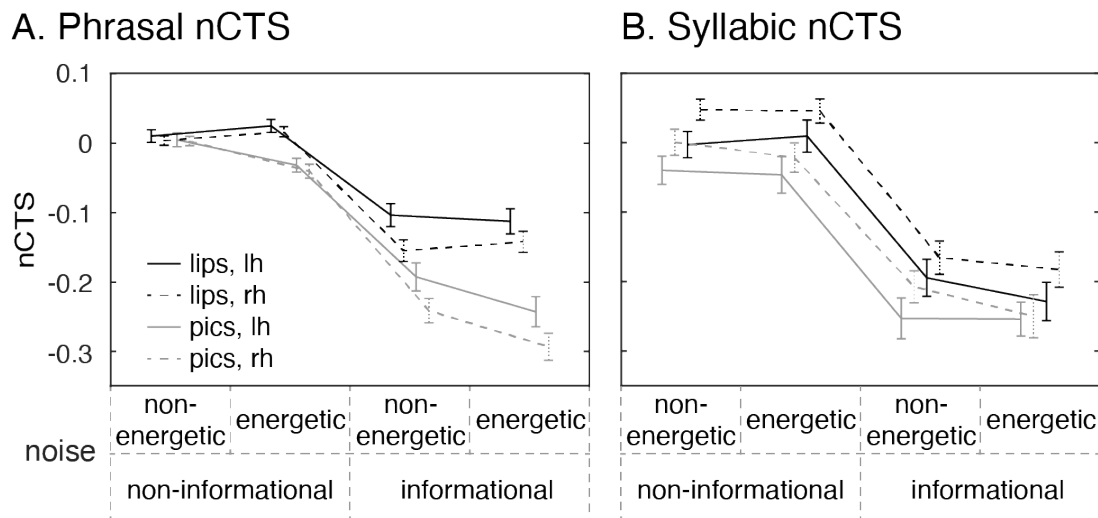
197    inclusion.

| | | $x^2$ | | p |
|---|---|---|---|---|
| | df | value | | |
| phrasal nCTS | | | | |
| noise | 3 | 597.50 | | < 0.0001 |
| visual | 1 | 127.28 | | < 0.0001 |
| hemisphere | 1 | 17.27 | | < 0.0001 |
| noise × visual | 3 | 67.70 | | < 0.0001 |
| noise × hemisphere | 3 | 10.98 | | 0.012 |
| syllabic nCTS | | | | |
| noise | 3 | 340.61 | | < 0.0001 |
| visual | 1 | 21.42 | | < 0.0001 |
| hemisphere | 1 | 10.38 | | 0.0013 |

198

199

200

201

10

## A. Phrasal nCTS

## B. Syllabic nCTS

**Figure 2.** Impact of the main fixed-effects on the normalized cortical tracking of speech (nCTS) at phrasal (**A**) and syllabic rates (**B**). Mean and SEM values are displayed as a function of noise properties. The four traces correspond to visual conditions with the speaker's talking face visible (*lips*; black traces) and with static pictures illustrating the story (*pics*; gray traces), within the left (*lh*; connected traces) and right (*rh*; dashed traces) hemispheres. nCTS values are bounded between –1 and 1, with values below 0 indicating lower CTS in speech-in-noise conditions than in noiseless conditions.

**What is the nature of the information about reading abilities brought by measures of SiN processing and classical behavioral predictors of reading?**

Having identified relevant features of cortical SiN processing, we first evaluated to which extent these features and classical behavioral predictors of reading bring information about reading abilities. More precisely, we used partial information decomposition (PID) to dissect the information about reading abilities (target) brought by behavioral scores (first set of explanatory variables) and features of the nCTS in noise (second set of explanatory variables).[62–64] Generally speaking, PID can reveal to which extent two sets of explanatory variables bring unique information about a target (information present in one set but not in the other), redundant information (information common to the two sets), and synergistic information (information emerging from the interaction of the two sets). Here, the target consisted of 5 reading scores: (i) an accuracy and (ii) a speed score for the reading of a

11

## A. Phrasal nCTS

## B. Syllabic nCTS

nCTS

lips, lh
lips, rh
pics, lh
pics, rh

non-energetic · energetic · non-energetic · energetic

non-informational · informational

noise

non-energetic · energetic · non-energetic · energetic

non-informational · informational

**Figure 2.** Impact of the main fixed-effects on the normalized cortical tracking of speech (nCTS) at phrasal (**A**) and syllabic rates (**B**). Mean and SEM values are displayed as a function of noise properties. The four traces correspond to visual conditions with the speaker's talking face visible (*lips*; black traces) and with static pictures illustrating the story (*pics*; gray traces), within the left (*lh*; connected traces) and right (*rh*; dashed traces) hemispheres. nCTS values are bounded between –1 and 1, with values below 0 indicating lower CTS in speech-in-noise conditions than in noiseless conditions.

**What is the nature of the information about reading abilities brought by measures of SiN processing and classical behavioral predictors of reading?**

Having identified relevant features of cortical SiN processing, we first evaluated to which extent these features and classical behavioral predictors of reading bring information about reading abilities. More precisely, we used partial information decomposition (PID) to dissect the information about reading abilities (target) brought by behavioral scores (first set of explanatory variables) and features of the nCTS in noise (second set of explanatory variables).[62–64] Generally speaking, PID can reveal to which extent two sets of explanatory variables bring unique information about a target (information present in one set but not in the other), redundant information (information common to the two sets), and synergistic information (information emerging from the interaction of the two sets). Here, the target consisted of 5 reading scores: (i) an accuracy and (ii) a speed score for the reading of a

11

223    connected meaningless text (Alouette test), and scores (number of correctly read words per unit

224    of time) for the reading of a list of (iii) irregular words, (iv) regular words and (v) pseudowords.

225    The first set of explanatory variables consisted of a total of 5 measures indexing phonological

226    awareness (scores on phoneme suppression and fusion tasks), phonological memory (scores on

227    forward and backward digit repetition), and lexical access (rapid automatized naming (RAN)

228    score). The second set of explanatory variables was the 8 features of nCTS in SiN conditions

229    identified in the previous subsection. Again, in that analysis, all measures were corrected for

230    age, time spent at school and IQ.

231        As a result, features of nCTS in noise brought significant unique information about

232    reading abilities (unique information = 0.61; $p = 0.016$), while classical behavioral predictors

233    did not (unique information = 0.31; $p = 0.10$). Both sets of explanatory variables brought

234    significant redundant but not synergistic information about reading (redundant information =

235    0.16; $p = 0.0020$; synergistic information = 0.12; $p = 0.26$).

236        These results indicate that the way the CTS is impacted by ambient noise relates to

237    reading abilities in a way that is not fully explained by classical behavioral predictors of

238    reading. Further analyses will therefore strive to identify which aspects of SiN processing and

239    reading are related, and which of these relations are mediated by classical behavioral predictors

240    of reading.

241

242    **Which features of SiN processing relate to reading abilities in a way that is not mediated**

243    **by classical behavioral predictors of reading?**

244        We next identified with linear mixed-effects modeling (i) the set of classical behavioral

245    predictors of reading that best explains reading abilities, and (ii) the set of features of nCTS in

246    noise that brings additional information about reading abilities. Importantly, all measures were

247    corrected for age, time spent at school, and IQ, and further standardized. In that analysis, the

12

248    type of reading score used to assess reading abilities was taken as a factor. Classical behavioral

249    predictors of reading (5 measures) were first entered as regressors, before considering the

250    features of nCTS in noise (8 measures) as additional regressors.

251        Table 4 presents the final linear mixed-effects model fit to reading scores. It shows that

252    lexical access (indexed by RAN score) and phonological memory (indexed by the forward digit

253    span) relate to global reading abilities. It also shows that two aspects of SiN processing, the

254    visual and informational modulations in phrasal nCTS, explain a different part of the variance

255    in reading abilities. Importantly, these two indices relate to reading in a way that depend on the

256    type of reading score. These effects are illustrated with simple Pearson correlations in Table 5.

257    The time necessary to fulfil the *RAN* task was significantly negatively correlated with all

258    reading scores. The *forward digit span* was significantly positively correlated with all reading

259    scores. The visual modulation in phrasal nCTS was overall positively correlated with scores

260    involving reading speed (Alouette speed score and regular-, irregular- and pseudoword reading

261    scores; significantly so for pseudoword reading only) but not with the Alouette accuracy score.

262    The informational modulation in phrasal nCTS was characterized by a significant positive

263    correlation with the score on irregular word reading only. Interestingly, the correlation was not

264    significant—and even negative—with the score on pseudoword reading.

265        We will now attempt to better understand the meaning of this last association (between

266    the informational modulation in phrasal nCTS and irregular- but not pseudoword reading).

267    Given that different routes support reading of irregular words (lexical route) and pseudowords

268    (sublexical route), the contrast between corresponding standardized scores (irregular

269    – pseudowords) indicates reading strategy. We henceforth refer to this index as the reading

270    strategy index. Further strengthening the correlation pattern highlighted above for the

271    informational modulation in phrasal nCTS, this latter index correlated even more strongly with

272    the reading strategy index ($r = 0.44$, $p < 0.0001$; See Fig. 3, left) than with the score on irregular

13

273 word reading. This suggests that irregular and pseudoword reading scores bring synergistic

274 information about the informational modulation in phrasal nCTS. To confirm this, we used

275 partial information decomposition (PID) to dissect the information about the informational

276 modulation in phrasal nCTS (target) brought by irregular reading scores (first explanatory

277 variable) and pseudoword reading scores (second explanatory variable). This analysis revealed

278 that the score on irregular word reading carried significant unique information about the

279 informational modulation in phrasal nCTS (unique information = 0.044, $p = 0.015$), while the

280 score on pseudowords did not (unique information = 0.0001, $p = 0.62$), and most interestingly,

281 that these two reading scores carried significant synergistic but not redundant information

282 about the informational modulation in phrasal nCTS (redundant information = 0.0019, $p = 0.44$;

283 synergistic information = 0.146, $p < 0.0001$).

284 Figure 3 (right panel) further illustrates that the reading strategy index was correlated

285 with phrasal nCTS only in the informational noise conditions.

286 In summary, classical behavioral predictors of reading were informative about global

287 reading abilities (similar correlation with all 5 measures of reading), while two aspects of the

288 CTS in noise (informational and visual modulations in phrasal nCTS) related to specific aspects

289 of reading (correlation with some but not all 5 measures of reading). The extent to which visual

290 speech boosts phrasal CTS in noise was related to reading speed but not accuracy, and the

291 ability to maintain adequate phrasal CTS in babble noise related to reading strategy (dominant

292 reliance on the lexical rather than sublexical route).

293

294

295

296

297

298

14

**Table 4.** Regressors included in the final linear mixed-effects model fit to the 5 reading scores taken as factors. Regressors are listed in their order of inclusion.

| | $x^2$ | | p |
|---|---|---|---|
| | df | value | |
| RAN | 1 | 15.79 | < 0.0001 |
| forward digit span | 1 | 11.10 | 0.0009 |
| visual modulation in phrasal nCTS | 1 | 4.85 | 0.028 |
| informational modulation in phrasal nCTS dependant on reading score | 5 | 15.63 | 0.0080 |
| visual modulation in phrasal nCTS dependant on reading score | 4 | 11.09 | 0.026 |

**Table 5.** Pearson correlation between measures of reading abilities and relevant brain and behavioral measures. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, [#] $p < 0.1$. nCTS, normalized cortical tracking of speech.

| | RAN | forward digit span | Visual modulation in phrasal nCTS | Informational modulation in phrasal nCTS | Visual modulation in syllabic nCTS | Phoneme suppression | Phoneme fusion |
|---|---|---|---|---|---|---|---|
| Alouette accuracy | **−0.37 \*\*** | **0.33 \*\*** | 0.00 | −0.35 | **0.29 \*** | 0.11 | **0.25 \*** |
| Alouette speed | **−0.41\*\*\*** | **0.38 \*\*\*** | 0.21 [#] | 0.08 | **0.30 \*\*** | **0.31 \*\*** | **0.30 \*\*** |
| Irregular words | **−0.35 \*\*** | **0.42 \*\*\*** | 0.18 | **0.26 \*** | **0.37 \*\*** | 0.21 [#] | 0.17 |
| Regular words | **−0.42 \*\*\*** | **0.35 \*\*** | 0.18 | 0.12 | **0.31\*\*** | **0.25 \*** | 0.19 |
| Pseudowords | **−0.34 \*\*** | **0.30 \*** | **0.31 \*\*** | −0.07 | **0.23 \*** | 0.21 [#] | 0.11 |

Relation between the reading strategy index and phrasal nCTS

314
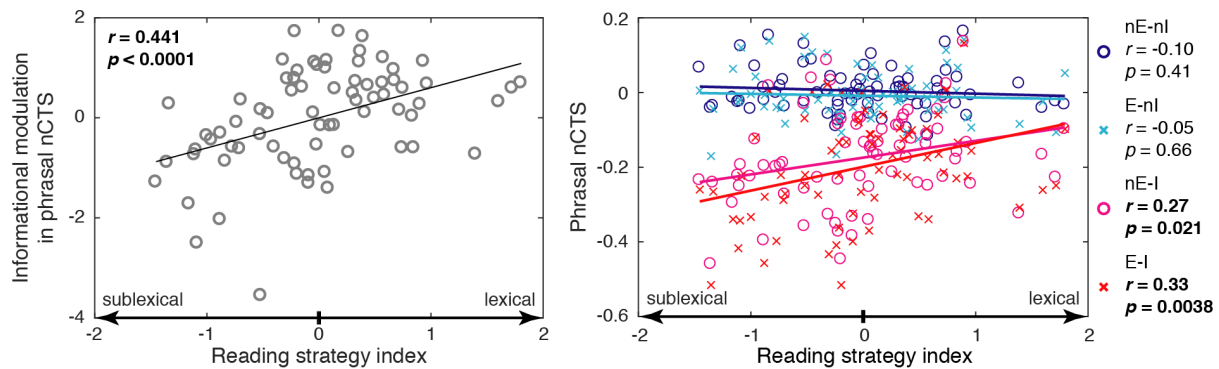
**Figure 3.** Relation between the reading strategy index and the normalized cortical tracking of speech (nCTS) at phrasal rate. **Left** — The informational modulation in phrasal nCTS as function of the reading strategy index. Gray circles depict participants' values and a black trace is the regression line, with correlation and significance values indicated in the bottom right corner. **Right** — The mean nCTS across visual conditions and both hemispheres for the 4 types of noise: energetic (*E*) or not (*nE*) and informational (*I*) or not (*nI*), with *nE-nI* in blue, *E-nI* in turquoise, *nE-I* in red and *E-I* in pink. Circles (*nE* noise conditions) and crosses (*E*) depict participants' values, and full traces are the regression lines. Correlation and significance level for all noise conditions are indicated on the right of each plot.

324

**Do other features of SiN processing or classical behavioral predictors of reading relate to reading abilities?**

Above, we have identified a set of brain and behavioral measures related to reading. Importantly, each measure was included because it explained a new part of the variance in reading abilities. But the first PID analysis revealed that brain and behavioral measures do carry significant redundant information. This means that some measures might have been left aside if they explained some variance that was already explained (i.e., if they provided mainly redundant information). Accordingly, we also ran the linear mixed-effects analysis with nCTS and behavioral regressors that were not included. This analysis identified an overall positive correlation between reading abilities and (i) the visual modulation in syllabic nCTS ($\mathcal{X}^2(1) = 9.74$, $p = 0.0018$), (ii) phoneme suppression ($\mathcal{X}^2(1) = 4.94$, $p = 0.026$) and (iii) phoneme fusion

16

336      ($\mathcal{X}^2(1) = 4.00$, $p = 0.038$). Corresponding Pearson correlation coefficients are presented in

337      Table 5. A detailed PID analysis revealed that these "side" measures were redundant—and

338      synergistic to some extent—with RAN and forward digit span but not with visual and

339      informational modulations in phrasal nCTS (see Supplementary Results).

340      In summary, scores indexing phonological awareness (score on phoneme suppression

341      and phoneme fusion) and the extent to which visual speech boosts syllabic CTS in noise (visual

342      modulation in syllabic nCTS) relate to global reading abilities in a way that is mediated by the

343      main classical behavioral predictors of reading we identified (RAN and forward digit span) but

344      not with visual and informational modulations in phrasal nCTS

345

346      **Does phonological awareness mediate SiN perception capacities?**

347      Having identified three relations between various aspects of cortical SiN processing

348      and reading, we now specifically test the hypothesis that each of these relations is mediated by

349      phonological awareness. For that, we again relied on PID to decompose the information about

350      reading abilities (target) brought by each identified feature of the CTS in noise (first

351      explanatory variable) and the mean of the two scores indexing phonological awareness (second

352      explanatory variable). Ensuing results are provided in Table S1. In summary, phonological

353      awareness mediated one aspect of the relation between reading and cortical SiN processing

354      (relation with the benefit of visual speech to boost syllabic CTS in noise), but not the two others

355      (relations involving phrasal CTS in noise).

356

357      **Do relations between reading and features of nCTS translate to alterations in dyslexia?**

358      We next evaluated whether the relations between features of nCTS and reading abilities

359      translate to alterations in dyslexia. That analysis was conducted on a group of 26 dyslexic

17

360    readers, 26 age-matched and 26 reading-level-matched typically developing children selected

361    among the 73 children included in the first part of the study.

362         Based on the result that reading abilities relate to phrasal nCTS in informational noise

363    and to the boost in nCTS brought by visual speech, we focused the comparison on the phrasal

364    nCTS in *lips* and *pics* averaged across hemispheres and babble noise conditions (see Fig. 4A).

365    As a result, individuals with dyslexia had significantly lower phrasal nCTS than age-matched

366    controls in *pics* ($t(50) = 3.03$, $p = 0.0039$) but not in *lips* ($t(50) = 0.83$, $p = 0.41$). This difference

367    was not present in the comparison with reading-level-matched controls (*pics*, $t(50) = 0.65$, $p =$

368    $0.52$; *lips*, $t(50) = 0.54$, $p = 0.59$).

369         Based on the result that reading abilities relate to the visual modulation in syllabic

370    nCTS, we focused the comparison on this index (see Fig. 4B left part). This revealed that

371    individuals with dyslexia had significantly lower visual modulation in syllabic nCTS than age-

372    matched ($t(50) = 2.62$, $p = 0.012$) but not reading-level-matched controls ($t(50) = 1.59$, $p =$

373    $0.12$). To better understand the nature of this difference, we further compared between groups

374    the syllabic nCTS in *lips* and *pics* averaged across hemispheres and noise conditions (see Fig.

375    4B right part). As a result, individuals with dyslexia had significantly lower syllabic nCTS than

376    age-matched controls in *lips* ($t(50) = 2.89$, $p = 0.0057$) but not in *pics* ($t(50) = 0.88$, $p = 0.38$).

377    This difference was not present in the comparison with reading-level-matched controls (*lips*,

378    $t(50) = 1.19$, $p = 0.24$; *pics*, $t(50) = 0.25$, $p = 0.81$).

379         In summary, one aspect of cortical SiN processing (reliance on visual speech to boost

380    phrasal nCTS) was not altered in dyslexia while two other aspects (phrasal nCTS in babble

381    noise, and reliance on visual speech to boost syllabic nCTS) were altered in dyslexia in

382    comparison with typical readers matched for age but not reading level. This suggests that these

383    two later aspects are altered as a consequence of reduced reading experience.

18

384
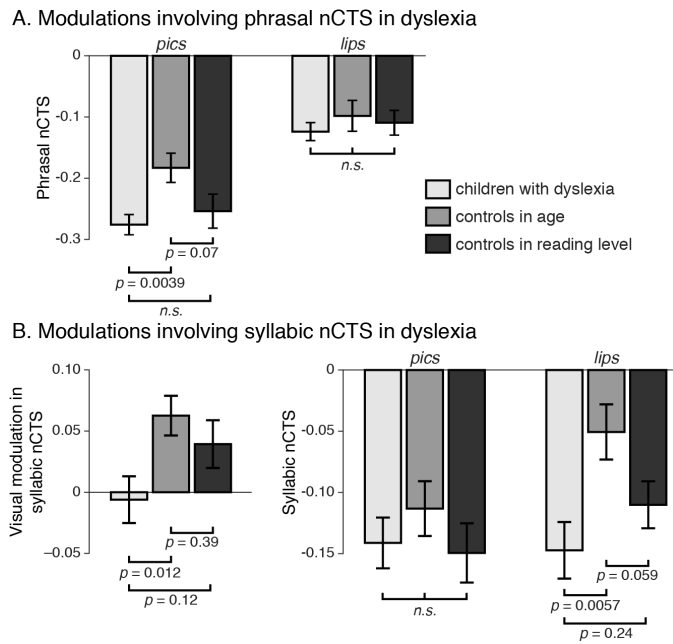
**Figure 4.** Comparison between children with dyslexia and controls in the measures of normalized cortical tracking of speech (nCTS) significantly related to reading abilities. **A** — Modulations involving phrasal nCTS. Displayed are the mean and SEM within groups (dyslexia, control in age, and control in reading level) of phrasal nCTS in the conditions with (*lips*) and without (*pics*) visual speech information. Values of nCTS were averaged across hemispheres and informational noise conditions for phrasal nCTS and across hemispheres and all noise conditions for syllabic nCTS. **B** — Modulations involving syllabic nCTS. On the left is the visual modulation in syllabic nCTS. The right part is as in A.

393

## Discussion

The main objective of this study was to fully characterize the nature of the relation between objective cortical measures of SiN processing and reading abilities in elementary-school children. Results demonstrate that some cortical measures of SiN processing relate to reading performance and reading strategy. First, phrasal nCTS in babble (*i.e.*, informational) noise relates to the ability to read irregular but not pseudowords, which in the DRC model indicates maturation of the lexical route. Second, the ability to leverage visual speech to boost phrasal nCTS in babble noise relates to reading speed (but not accuracy). Third, the ability to leverage visual speech to boost syllabic nCTS in noise relates to global reading abilities. Fourth,

19

403  classical behavioral predictors of reading abilities (RAN, phonological memory and

404  phonological awareness) relate to global reading performance, and not strategy. Importantly,

405  behavioral scores and the two features of phrasal CTS in babble noise explained a different part

406  of the variance in reading abilities. Finally, the first and third relations uncovered in typical

407  readers translated to significant alteration in dyslexia in comparison with aged-matched but not

408  reading-level-matched typically developing children. Two limitations are discussed in

409  Supplementary Discussion.

410  Significant associations were found between reading abilities and some features of

411  phrasal and syllabic nCTS. There is evidence that CTS at phrasal rate (here taken as 0.2–1.5

412  Hz) partly reflects parsing or chunking of words, phrases and sentences.[65] Indeed, the brain

413  tracks phrase and sentence boundaries even when speech is devoid of prosody but only if it is

414  comprehensible;[39] and the phase of below-4-Hz brain oscillations modulates perception of

415  ambiguous sentences.[37] CTS at phrasal/sentential rate would help align neural excitability with

416  syntactic information to optimize language comprehension.[36] In contrast, CTS at syllable rate

417  (here taken as 2–8 Hz) would reflect low level auditory processing.[65] In light of the above, our

418  results highlight that associations between SiN perception and reading abilities build on their

419  shared reliance on both language processing and low-level auditory processing.

420

421  **Robustness of cortical speech representation to babble noise indexes the development of**

422  **the lexical route**

423  Our results indicate that an objective cortical measure of the ability to deal with babble

424  noise relates to the maturation of the lexical route. Technically, the informational modulation

425  in phrasal nCTS correlated significantly positively with the reading score on irregular but not

426  pseudowords. Reading score on irregular words indeed provided unique information about the

427  informational modulation in nCTS. Also, the two reading scores in synergy provided some

428    additional information about the informational modulation in nCTS. Furthermore, the result

429    that the informational modulation in nCTS correlated more with the reading strategy index than

430    the score on the irregular words suggests that the key elements at the basis of this relation are

431    the processes needed to read irregular words that are not needed to read pseudo-words. Our

432    results in dyslexia support this relation although they cannot rule out the possibility that it is

433    due to variability in reading experience since phrasal nCTS in non-visual informational noise

434    conditions was reduced in dyslexic readers compared with age-matched but not reading-level-

435    matched controls.

436        The relation between the development of the lexical route and the level of phrasal nCTS

437    in babble noise could be explained by a positive influence of good SiN abilities on reading

438    acquisition. Let us take as an example the situation of being faced for the first time with a

439    written word that is read by a teacher while some classmates are making noise. SiN abilities

440    will naturally determine the odds of hearing that word properly and hence the odds of building

441    up the orthographic lexicon. When reading again the word alone, only children with good SiN

442    abilities will have the opportunity to train their lexical route for that specific word. Of course

443    the same chain of action could be posited for the training of grapheme–phoneme

444    correspondence. But there are many more words than phonemes and syllables, so that good

445    SiN abilities might be more important to successfully learn the correspondence between

446    irregular words' orthographic and phonological representations. Indeed, grapheme–phoneme

447    correspondence is intensively trained when learning to read. Children are repeatedly exposed

448    to examples of successful grapheme–phoneme correspondence, some with noise, and some

449    without noise. Accordingly, no matter what children's SiN abilities are, they will learn the

450    grapheme–phoneme correspondence and develop their sublexical route, provided that they

451    have adequate phonological awareness. Supporting this, phonological awareness does not

452    predict SiN abilities in typical readers.[21]

21

453     Alternatively, the relation between the ability to read irregular words (which tags the

454     development of the lexical route) and nCTS in babble noise could be mediated by the degree

455     of maturation of the mental lexicon.[66,67] The mental lexicon integrates and binds the

456     orthographic, semantic and phonological representations of words. Its proper development is

457     important for reading acquisition. Indeed, reading acquisition entails creating a new

458     orthographic lexicon and binding it to the preexisting semantic and phonological lexicons.[68]

459     Development of such binding (i) is indispensable to read irregular words,[69] (ii) benefits reading

460     of regular words, and (iii) does not contribute to reading pseudowords. The proper development

461     of the mental lexicon is also important for SiN comprehension. Indeed, SiN comprehension

462     strongly depends on vocabulary knowledge.[21,70,71] And the level of CTS in noise relates to the

463     listeners' level of comprehension.[35,40,41] This therefore suggests that the robustness of CTS to

464     babble noise depends on the level of comprehension, which in turn depends on how developed

465     is the mental lexicon. The development of the mental lexicon could therefore be the hidden

466     factor mediating the relation between SiN and lexical reading ability. This is also perfectly in

467     line with our result that altered phrasal nCTS in babble noise in dyslexia may result from

468     reduced reading experience. In brief, reading difficulties in dyslexia would reduce their reading

469     experience, which would impair building up the mental lexicon, and in turn impede SiN

470     perception. Still, future studies on the association between SiN processing and reading should

471     include measures of the development of the mental lexicon to carefully analyse the interrelation

472     between SiN perception, reading abilities and the development of the mental lexicon.

473

**Audiovisual integration and reading abilities**

475     We found significant relations between reading abilities and the ability to leverage

476     visual speech to maintain phrasal and syllabic CTS in noise. Visual speech cues (articulatory

477     mouth and facial gestures) are well known to benefit SiN comprehension [55] and CTS in noise.[72–]

478  [76] Obviously, the auditory signal carries much more fine-grained information about the

479  phonemic content of speech than the visual signal. But the effect of audiovisual speech

480  integration is quite evident in SiN conditions, where it affords a substantial comprehension

481  benefit.[55,56,77,78] Mirroring this perceptual benefit, it is already well documented that phrasal

482  and syllabic CTS in noise is boosted in adults when visual speech information is available.[72–

483  76,79–82]

484      We found that the visual modulation in phrasal nCTS correlated globally positively

485  with reading speed (significantly so for the pseudowords) but not accuracy. However, our

486  dyslexic readers (compared with both control groups) did not have any alteration in their

487  phrasal nCTS in babble noise when visual speech was provided. Instead, they successfully

488  relied on visual speech information to restore their phrasal CTS in babble noise (which was

489  altered without visual speech information). In other words, reliance on lip-reading to maintain

490  appropriate phrasal CTS in babble noise appeared as a protection factor in our group of dyslexic

491  readers.

492      We also found that the visual modulation in syllabic nCTS correlated globally

493  positively with reading abilities. More interestingly, our dyslexic readers (compared with both

494  control groups) did not have any significant alteration in their syllabic nCTS in noise when

495  visual speech was not provided. However, compared with age-matched typically developing

496  children, they benefited significantly less from visual speech to boost syllabic CTS in noise.

497  Instead, they behaved more like reading-level-matched typically developing children.

498  Accordingly, our results cannot argue against the view that poor audiovisual integration in

499  dyslexia is caused by reduced reading experience.[57,83,84] Notwithstanding, the pattern of results

500  (see Fig. 4B left) is even suggestive of an alteration in dyslexia in comparison with reading-

501  level-matched children. More statistical power would be needed to confirm/infirm the trend.

502    Our result that audiovisual integration abilities correlate with reading abilities is in line

503    with existing literature. Indeed, individuals with dyslexia benefit less from visual cues to

504    perceive SiN than typical readers[85–89]. Audiovisual integration and reading could be altered in

505    dyslexia simply because both rely on similar mechanisms. Indeed, reading relies on the ability

506    to bind visual (graphemic) and auditory (phonemic) speech representations.[90,91] And according

507    to some authors, suboptimal audiovisual integration mechanisms could reduce reading

508    fluency.[92] Importantly, the finding that individuals with dyslexia benefit normally from visual

509    speech to boost phrasal but not syllabic CTS in noise brings important information about the

510    nature of the audiovisual integration deficit in dyslexia. Following the functional roles

511    attributed to CTS, individuals with dyslexia would properly integrate visual speech information

512    to optimize processing of syntactic information,[36] but not to support acoustic/phonemic

513    processing.[65] This could be explained by their preserved ability to extract and integrate the

514    temporal dynamics of visual speech, but not the lip configuration,[89] two aspects of audiovisual

515    speech integration currently thought to be supported by distinct neuronal pathways.[93] This

516    inability to rely on lip configuration to improve auditory phonemic perception in SiN

517    conditions may be caused by a supra-modal phonemic categorization deficit, as already

518    proposed for children with specific language impairment.[94] Finally, the fact that the visual

519    modulation in syllabic nCTS brought a limited amount of unique information about reading

520    with respect to classical behavioral predictors of reading, but that all of them brought more

521    information in synergy, suggests that a broad set of low-level processing abilities contribute to

522    determine reading abilities and alterations in dyslexia.[95,96]

523

524    **Classical behavioral predictors related to global reading abilities**

525    Our results confirm that classical behavioral predictors of reading (RAN, phonological

526    memory, and metaphonological abilities) are directly related to the global reading level rather

24

527    than reading strategy. We draw this conclusion since the optimal model for reading score

528    contained a common slope for all reading subtests. This means that the model was not

529    significantly improved by optimizing the slope for each of the 5 reading subtests separately.

530    Accordingly, univariate correlation coefficients presented in Table 5 were roughly similar

531    across the 5 reading scores.

532          Phonological memory (assessed with forward digit span) was significantly positively

533    correlated with the global reading level. That phonological memory relates to global reading

534    abilities rather than reading strategy is well documented.[4] Poor readers, regardless of their

535    reading profile, typically perform poorly on phonological memory tests involving digits,

536    letters,[97,98] or words.[99]

537          Performance on the RAN task was also related to the global reading level, in line with

538    existing literature.[6–8,100–103] RAN performance has indeed a moderate-to-strong relationship

539    with all classical reading measures alike, including word, non-word and text reading, as well

540    as text comprehension.[100] It is a consistent predictor of reading fluency in various alphabetic

541    orthographies independently of their complexity.[104] RAN performance even predicts reading

542    performance at 2 years interval,[105] similarly well for reading performance assessed with tasks

543    tagging lexical and sublexical routes. It is thought that RAN and reading performances correlate

544    because both involve serial processing and oral production,[103] two processes that are common

545    to both reading routes.

546          Finally, phonological awareness assessed with phoneme suppression and fusion tasks

547    was significantly related to reading abilities. However, the information it brought about reading

548    was less, and essentially redundant with that brought by RAN and phonological memory. This

549    is not surprising given that children tested in the present study had at least one year of reading

550    experience. Phonological awareness indeed plays a key role in the early stages of reading

551     acquisition, i.e., when learning grapheme-to-phoneme conversion,[106–108] and undergoes a

552     substantial maturation during that period.[109]

553

554     **Phonological awareness**

555     Our results indicate that, in typical readers, phonological awareness mediates at best

556     part of the relation between the cortical processing of SiN and reading abilities. Indeed, the

557     information about reading brought by phonological awareness was redundant with that brought

558     by the visual modulation in syllabic nCTS, but not with that brought by the informational and

559     visual modulations in phrasal nCTS. This finding illustrates the importance of separating the

560     different processes involved in SiN processing and reading to seek associations. It also provides

561     a potential reason why contradictory reports exist on the topic.[19–21]

562

563     **Conclusion**

564     Overall, these results significantly further our understanding of the nature of the relation

565     between SiN processing abilities and reading abilities. They demonstrate that cortical

566     processing of SiN and reading abilities are related in several specific ways, and that some of

567     these relations translate into alterations in dyslexia that are attributable to reading experience.

568     They also demonstrate that classical behavioral predictors of reading (including phonological

569     awareness) mediate relations involving the processing of acoustic/phonemic but not syntactic

570     information in natural SiN conditions. This contrasts with the classically assumed mediating

571     role of phonological awareness. Instead, the ability to process speech syntax in babble noise

572     could directly modulate skilled reading acquisition. Finally, the information about reading

573     abilities brought by cortical markers of syntactic processing of SiN was complementary to that

574     provided by classical behavioral predictors of reading. This implies that such markers of SiN

575     processing could serve as novel electrophysiological markers of reading abilities.

## Material and Methods

### Participants

Seventy-three typical and 26 dyslexic readers enrolled in elementary-school took part in this experiment (see Table 1 for participants' characteristics). All were native French speakers, reported being right-handed, had normal hearing according to pure tone audiometry (normal hearing thresholds between 0–25 dB HL for 250, 500, 1000, 2000, 4000 and 8000 Hz) and normal SiN perception as revealed by a SiN test (Lafon 30) from a French language central auditory battery.[110] We used a French translation of the Family Affluence Scale [111] to evaluate participants' socio-economic level.

This study was approved by the Ethics Committee of the CUB Hôpital Erasme (Brussels, Belgium). Participants and their legal representatives signed a written informed consent before participation. Participants were compensated with a gift card worth 50 euros.

### Behavioral assessment

Participants underwent a comprehensive behavioral assessment intended to appraise their reading abilities and some cognitive abilities related to reading or speech perception.

### *Reading abilities*

Children completed the word reading (regular, irregular and pseudowords) tasks of a dyslexia detection tool (ODEDYS-2; [112] and the Alouette-R reading task [113]).

For each of the word reading tasks (regular, irregular or pseudowords), participants had to read as rapidly and accurately as possible a list of 20 words. Each task provided a *reading score* computed as the number of words correctly read divided by the reading time (in seconds).

In the Alouette-R task,[113] children had 3 minutes to read as rapidly and accurately as possible a text of 256 words. This text is composed by a succession of words which do not tell

601   a meaningful story. This peculiarity forces children to solely rely on their reading skills and

602   prevents children from using anticipation or inference strategies that could boost the reading

603   scores. An *accuracy* score was computed as the number of words correctly read divided by the

604   total number of words read, and a *speed* score as the number of words correctly read multiplied

605   by the ratio of 180 s (maximal reading time) to the effective reading time.

606

607   ***Phonological processing***

608   The initial phoneme suppression and initial phonemes fusion tasks of the ODEDYS-2

609   [112] was used to assess phonological processing.

610   In the initial phoneme suppression task, children had to repeat orally presented words

611   while intentionally suppressing the initial phoneme of the word (*i.e.* dog -> og). In total, 10

612   words were presented, and performance was quantified as percentage correct.

613   In the initial phoneme fusion task, children had to combine the initial phoneme of two

614   orally presented words to create a new (non-)word (*i.e., Big & Owen -> /bo/*). In total, 10 pairs

615   of words were presented, and performance was quantified as percentage correct.

616

617   ***Rapid Automatized Naming***

618   We used the RAN task of the ODEDYS-2.[112] Children had to name as rapidly and

619   accurately as possible 25 pictures (5 different pictures randomly repeated 5 times).

620   Performance was quantified as the total time to complete the task, meaning that the lower the

621   score, the better the performance.

622

623   ***Phonological memory***

624   The forward and backward digit repetition task from the ODEDYS-2 [112] was used to

625   assess phonological memory.

28

626      In the forward digit repetition task, children were asked to repeat orally presented

627    numbers series in the same order as presented. The series are different at every trial. The first

628    series contains 3 digits, and the size of the series is incremented by one every second trial. The

629    task ends after a failure to repeat the 2 series of a given size. Forward digit span score was

630    taken as the number of digits in the last correctly repeated series.

631      The backward digit repetition task is akin to the forward one. The only difference is

632    that digit series have to be repeated in the exact reverse order (e.g., children presented 1 2 3 4

633    have to repeat 4 3 2 1).

634

635    ***Attention abilities***

636      The Bells test [114] was used to assess visual attention and the TAP auditory attention

637    subtest [115] to assess the auditory attentional level.

638      In the Bells test, children had 2 minutes to find as many bells as possible on a sheet

639    comprising 35 bells scattered among 280 visual distractors. Performance was quantified as the

640    number of bells found divided by the time needed.

641      In the TAP auditory attention subtest, a, children had to focus their attention during 3

642    min 20 s on an auditory stream. Children were hearing a train of 200 pure tone stimuli lasting

643    500 ms with a 1000-ms stimulus onset asynchrony. Tones alternated between high (1073 Hz)

644    and low (450 Hz) pitch. There were 16 occurrences in which 2 high or low pitch tones were

645    following one another. Only in this case, participants had to press a response button as fast as

646    possible. A performance score was quantified as the number of correct responses, a speed score

647    as the mean response time, and a failure score as the number of responses to tones differing in

648    pitch with the preceding one.

649

650     ***Non-verbal intelligence***

651     The brief version of the Weschler Nonverbal (WNV) Scale of Ability [116] was used to

652     assess non-verbal intelligence.

653     This assessment consisted of matrices and recognition subtests for children younger

654     than 8 years. Older children were assessed with matrices and spatial memory subtests.

655     In matrices subtest, children were presented with incomplete visual matrices and had to

656     select the good missing portion among 4 or 5 response options. The subtest ended when 4

657     mistakes were made in the last 5 trials. A raw score was taken as the number of correctly

658     completed matrices. This raw score was converted to a *T* score by comparison with values

659     provided in a table of norms.

660     In recognition subtest, children had to carefully look at visual geometric designs that

661     were presented one by one for three seconds. After each presentation, they had to identify the

662     previously seen design among four or five response options. The subtest ended when 4 mistakes

663     were made in the last 5 recognition trials. A raw score was taken as the number of correctly

664     recognized drawings. This raw score was converted to a *T* score by comparison with values

665     provided in a table of norms.

666     In spatial memory subtest, children were presented with a board with 10 cubes spread

667     on it, and were asked to mimic the examiner's tapping sequence. The sequences are different

668     on every trial. The first sequence consists in tapping on 2 cubes, and the size of the sequences

669     is incremented by one every second trial. The task ends after a failure to repeat 2 sequences of

670     a given size. This task was performed twice, in the forward and backward directions. For each

671     direction, a raw score was taken as the number of correctly repeated sequences. Raw scores

672     were summed and converted to a *T* score by comparison with values provided in a table of

673     norms.

674    Total non-verbal IQ was computed as the sum of both *T* scores, which compared to a

675    table of norms provided a *Total non-verbal* IQ score.

676

677    **Neuroimaging assessment**

678    *Stimuli*

679    The stimuli were derived from 12 audiovisual recordings of 4 native French-speaking

680    actors (2 females, 3 recordings per actor) telling a story for ~6-min (mean ± SD, 6.0 ± 0.8 min)

681    (see Supplementary Methods for details on recording of video stimuli). In each video, the first

682    5 s were kept unaltered to enable children to unambiguously identify the actor's voice and face

683    they were requested to attend to. The remainder of the video was divided into 10 consecutive

684    blocks of equal size that were assigned to 9 conditions. Two blocks were assigned to the

685    *noiseless* condition in which the audio track was kept but the video was replaced by static

686    pictures illustrating the story (mean ± SD picture presentation time across all videos, 27.7 ±

687    10.8 s). The remaining 8 blocks were assigned to 8 conditions in which the original sound was

688    mixed with a background noise at 3 dB signal-to-noise ratio (SNR). There were 4 different

689    types of noise, and each type of noise was presented once with the original video, thereby

690    giving access to lip-read information (*lips* visual conditions), and once with the static pictures

691    illustrating the story (*pics* visual conditions). The different types of noise differed in the degree

692    of energetic and informational interference they introduced.[54] The non-energetic non-

693    informational noise was a white noise filtered through 100–10000-Hz. The (maximally-

694    )energetic non-informational noise had its spectral properties dynamically adapted to mirror

695    those of the actor's voice ~1 s around (see Supplementary Methods for the procedure used to

696    build the energetic non-informational noise). The non-(or least-)energetic informational noise

697    was a 5-talker cocktail party noise recorded by individuals of gender opposite to the actor's

698    (i.e., a 5-man for female actors). The (maximally-)energetic informational noise was a 5-talker

31

699  cocktail party noise recorded by individuals of gender identical to the actor's. The assignment

700  of conditions to blocks was random, with the constraint that each of the 5 first and last blocks

701  contained exactly 1 *noiseless* audio, 2 *lips* videos, 2 energetic noises, and 2 informational

702  noises. Smooth audio and video transitions between blocks was ensured with 2-s fade-in and

703  fade-out. Ensuing videos were grouped in 3 disjoint sets featuring one video of each of the

704  actors (total set duration: 23.0, 24.3, 24.65 min), and there were 4 versions of each set differing

705  in condition random ordering.

706

### *Experimental paradigm*

708  During the imaging session, participants were laying on a bed with their head inside the

709  MEG helmet. Their brain activity was recorded while they were attending 4 videos (separate

710  recording for each video) of a randomly selected set and ordering of the videos presented in a

711  random order, and finally while they were at *rest* (eyes opened, fixation cross) for 5 min. They

712  were instructed to watch the videos attentively, listen to the actors' voice while ignoring the

713  interfering noise, and remain as still as possible. After each video, they were asked 10 yes/no

714  comprehension questions related to each of the 10 blocks/conditions (data not analyzed here).

715  Videos were projected onto a back-projection screen placed vertically, ~120 cm away from the

716  MEG helmet. The inner dimensions of the black frame were 35.2 cm (horizontal) and 28.8 cm

717  (vertical), and actors face spanned ~15 cm (horizontal) and ~20 cm (vertical). Participants

718  could see the screen through a mirror placed above their head. In total the optical path from the

719  screen to participants' eyes was of ~150 cm. Sounds were delivered at 60 dB (measured at ear-

720  level) through a MEG-compatible front-facing flat-panel loudspeaker (Panphonics Oy, Espoo,

721  Finland) placed ~1 m behind the screen.

722

723 *Data acquisition*

724        During the experimental conditions, participants' brain activity was recorded with MEG

725    at the CUB Hôpital Erasme. Neuromagnetic signals were recorded with a whole-scalp-covering

726    MEG system (Triux, Elekta) placed in a lightweight magnetically shielded room (Maxshield,

727    Elekta), the characteristics of which being described elsewhere.[117] The sensor array of the MEG

728    system comprised 306 sensors arranged in 102 triplets of one magnetometer and two

729    orthogonal planar gradiometers. Magnetometers measure the radial component of the magnetic

730    field, while planar gradiometers measure its spatial derivative in the tangential directions. MEG

731    signals were band-pass filtered at 0.1–330 Hz and sampled at 1000 Hz.

732        We used 4 head-position indicator coils to monitor subjects' head position during the

733    experimentation. Before the MEG session, we digitized the location of these coils and at least

734    300 head-surface points (on scalp, nose, and face) with respect to anatomical fiducials with an

735    electromagnetic tracker (Fastrack, Polhemus).

736        Finally, subjects' high-resolution 3D-T1 cerebral images were acquired with a magnetic

737    resonance imaging (MRI) scanner (MRI 1.5T, Intera, Philips) after the MEG session.

738

739 *Data pre-processing*

740        Continuous MEG data were first preprocessed off-line using the temporal signal space

741    separation method implemented in MaxFilter software (MaxFilter, Neuromag, Elekta;

742    correlation limit 0.9, segment length 20 s) to suppress external interferences and to correct for

743    head movements.[118,119] To further suppress physiological artifacts, 30 independent components

744    were evaluated from the data band-pass filtered at 0.1–25 Hz and reduced to a rank of 30 with

745    principal component analysis. Independent components corresponding to heartbeat, eye-blink,

746    and eye-movement artifacts were identified, and corresponding MEG signals reconstructed by

747    means of the mixing matrix were subtracted from the full-rank data. Across subjects and

33

748 conditions, the number of subtracted components was $3.45 \pm 1.23$ (mean $\pm$ SD across subjects

749 and recordings). Finally, a window time of 1 s time points at timings 1 s around remaining

750 artifacts were set to bad. Data were considered contaminated by artifacts when MEG amplitude

751 exceeded 5 pT in at least one magnetometer or 1 pT/cm in at least one gradiometer.

752 We extracted the temporal envelope of the attended speech (actors' voice) using the

753 optimal approach proposed by Biesmans et al.[120]. Briefly, audio signals were bandpass filtered

754 using a gammatone filter bank (15 filters centered on logarithmically-spaced frequencies from

755 150 Hz to 4000 Hz), and subband envelopes were computed using Hilbert transform, elevated

756 to the power 0.6, and averaged across bands.

757

758 ***Accuracy of speech envelope reconstruction and normalized CTS***

759 For each condition and participant, a global value of cortical tracking of the attended

760 speech was evaluated for all left-hemisphere sensors at once, and for all right-hemisphere

761 sensors at once. Using the mTRF toolbox,[58] we trained a decoder on MEG data to reconstruct

762 speech temporal envelope, and estimated its Pearson correlation with real speech temporal

763 envelope. This correlation is often referred to as the reconstruction accuracy, and it provides a

764 global measure of cortical tracking of speech. See Supplementary Methods for a full description

765 of the procedure and statistical assessment. A similar approach has been used in previous

766 studies on the cortical tracking of speech.[47,51,60,61]

767 Based on CTS values in noiseless condition ($CTS_{noiseless}$) and in each SiN condition

768 ($CTS_{SiN}$), we estimated nCTS as follows:

$$nCTS = (CTS_{SiN} - CTS_{noiseless})/(CTS_{SiN} + CTS_{noiseless}).$$

770 This index can however be misleading when derived from negative CTS values (which may

771 happen since CTS is an unsquared correlation values). For this reason, and for the sake of nCTS

34

772    computation only, CTS values below a threshold of 10% of the mean CTS across all subjects,

773    conditions and hemispheres were set to that threshold prior to nCTS computation.

774

775    **Linear mixed-effects modeling of nCTS and reading values**

776    All behavioral and nCTS measures were corrected for IQ, age, time spent at elementary

777    school, and for outliers (see Supplementary Methods for details on this procedure).

778    We performed linear mixed-effects analysis with R [121] and *lme4* [122] to identify how

779    different fixed effects modulate nCTS. We started with a null model that included only a

780    different random intercept for each subject. The model was iteratively compared to models

781    incremented with simple fixed effects of *hemisphere*, *noise* (non-energetic non-informational,

782    energetic non-informational, non-energetic informational, and energetic informational), and

783    *visual* (*lips* vs. *pics*) added one by one. At every step, the most significant fixed effect was

784    retained, until the addition of the remaining effects did not improve the model any further ($p >$

785    $0.05$). The same procedure was then repeated to refine the ensuing model with the interactions

786    of the simple fixed effects of order 2 (e.g., *hemisphere × noise*) and then 3 (*hemisphere × noise*

787    *× visual*).

788    We followed the same approach to identify how reading abilities (5 standardized

789    scores) relate to classical behavioral predictors of reading and features of nCTS. In that

790    analysis, we first considered a non-zero slope for the classical behavioral predictors identical

791    for all reading scores, then a non-zero slope for the classical behavioral predictors different for

792    all reading scores, then a non-zero slope for the features of nCTS identical for all reading

793    scores, and finally a non-zero slope for the features of nCTS different for all reading scores.

794

**Partial information decomposition**

We used PID to appraise without *a priori* the relation between reading abilities, cortical measures of SiN processing, and classical behavioral predictors of reading. In general, PID decomposes the mutual information (MI) quantifying the relationship between two explanatory variables (or sets of explanatory variables) and a single target, into four constituent terms: the unique information about the target which is available separately from each explanatory variable alone, the redundant or shared information which is common to the two explanatory variables, and synergistic information, which is information about the target that is available only when both explanatory variables are observed together (e.g. the relationship between their values is informative about the target).[62–64] In our analysis, the 5 reading scores were used as the target, the features nCTS as the first set of explanatory variables, and behavioral scores as the second set of explanatory variables. PID was also used to better understand the nature of some other statistical associations we uncovered. See Supplementary Methods for further details on PID and its statistical assessment.

## Data availability

The data and the code that support the findings of this study are available at "a link to a OSF repository will be provided upon positive review".

36

# 5. References

1.  Leppänen, P. H. T. *et al.* Infant brain responses associated with reading-related skills before school and at school age. *Neurophysiol. Clin.* **42**, 35–41 (2012).

2.  Share, D. L., Jorm, A. F., Maclean, R. & Matthews, R. Sources of individual differences in reading acquisition. *Journal of Educational Psychology* vol. 76 1309–1324 (1984).

3.  Caravolas, M., Hulme, C. & Snowling, M. J. The Foundations of Spelling Ability: Evidence from a 3-Year Longitudinal Study. *Journal of Memory and Language* vol. 45 751–774 (2001).

4.  Muter, V. & Snowling, M. Concurrent and Longitudinal Predictors of Reading: The Role of Metalinguistic and Short-Term Memory Skills. *Reading Research Quarterly* vol. 33 320–337 (1998).

5.  Gathercole, S. E. & Baddeley, A. D. Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education* vol. 8 259–272 (1993).

6.  Manis, F. R., Doi, L. M. & Bhadha, B. Naming speed, phonological awareness, and orthographic knowledge in second graders. *J. Learn. Disabil.* **33**, 325–33, 374 (2000).

7.  Wimmer, H., Mayringer, H. & Landerl, K. The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology* vol. 92 668–680 (2000).

8.  Wimmer, H., Mayringer, H. & Landerl, K. Poor Reading: A Deficit in Skill-Automatization or a Phonological Deficit? *Scientific Studies of Reading* vol. 2 321–340 (1998).

9.  Samuelsson, S. & Lundberg, I. The impact of environmental factors on components of reading and dyslexia. *Annals of Dyslexia* vol. 53 201–217 (2003).

10. Hooper, S. R., Roberts, J., Sideris, J., Burchinal, M. & Zeisel, S. Longitudinal predictors of reading and math trajectories through middle school for African American versus Caucasian students across two samples. *Dev. Psychol.* **46**, 1018–1029 (2010).

11. Klatte, M., Bergström, K. & Lachmann, T. Does noise affect learning? A short review on noise effects on cognitive performance in children. *Frontiers in Psychology* vol. 4 (2013).

12. Stockman, J. A. Aircraft and Road Traffic Noise and Children's Cognition and Health: A Cross-

National Study. *Yearbook of Pediatrics* vol. 2007 69–71 (2007).

13. McDermott, J. H. The cocktail party problem. *Current Biology* vol. 19 R1024–R1027 (2009).

14. Anderson, S. & Kraus, N. Sensory-cognitive interaction in the neural encoding of speech in noise: a review. *J. Am. Acad. Audiol.* **21**, 575–585 (2010).

15. White-Schwoch, T. *et al.* Auditory Processing in Noise: A Preschool Biomarker for Literacy. *PLoS Biol.* **13**, e1002196 (2015).

16. Calcus, A., Colin, C., Deltenre, P. & Kolinsky, R. Informational masking of speech in dyslexic children. *The Journal of the Acoustical Society of America* vol. 137 EL496–EL502 (2015).

17. Ziegler, J. C., Pech-Georgel, C., George, F. & Lorenzi, C. Speech-perception-in-noise deficits in dyslexia. *Developmental Science* vol. 12 732–745 (2009).

18. Dole, M., Hoen, M. & Meunier, F. Speech-in-noise perception deficit in adults with dyslexia: Effects of background type and listening configuration. *Neuropsychologia* vol. 50 1543–1552 (2012).

19. Nittrouer, S. From Ear to Cortex: A Perspective on What Clinicians Need to Understand About Speech Perception and Language Processing. *Lang. Speech Hear. Serv. Sch.* **33**, 237–252 (2002).

20. Fallon, M., Trehub, S. E. & Schneider, B. A. Children's perception of speech in multitalker babble. *The Journal of the Acoustical Society of America* vol. 108 3023–3029 (2000).

21. Lewis, D., Hoover, B., Choi, S. & Stelmachowicz, P. Relationship between speech perception in noise and phonological awareness skills for children with normal hearing. *Ear Hear.* **31**, 761–768 (2010).

22. Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* **108**, 204–256 (2001).

23. Coltheart, M., Curtis, B., Atkins, P. & Haller, M. Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review* vol. 100 589–608 (1993).

24. Fiez, J. A. & Petersen, S. E. Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences* **95**, 914–921 (1998).

25. Turkeltaub, P. E., Eden, G. F., Jones, K. M. & Zeffiro, T. A. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage* **16**, 765–780 (2002).

26. McCandliss, B. D., Cohen, L. & Dehaene, S. The visual word form area: expertise for reading in the

fusiform gyrus. *Trends Cogn. Sci.* **7**, 293–299 (2003).

27. Dehaene, S. & Cohen, L. The unique role of the visual word form area in reading. *Trends Cogn. Sci.* **15**, 254–262 (2011).

28. Destoky, F. *et al.* Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope. *Neuroimage* **184**, 201–213 (2019).

29. Vander Ghinst, M. *et al.* Cortical Tracking of Speech-in-Noise Develops from Childhood to Adulthood. *J. Neurosci.* **39**, 2938–2950 (2019).

30. Bourguignon, M. *et al.* The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Human Brain Mapping* vol. 34 314–326 (2013).

31. Ahissar, E. *et al.* Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 13367–13372 (2001).

32. Gross, J. *et al.* Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol.* **11**, e1001752 (2013).

33. Luo, H. & Poeppel, D. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* **54**, 1001–1010 (2007).

34. Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M. & Carreiras, M. Out-of-synchrony speech entrainment in developmental dyslexia. *Hum. Brain Mapp.* **37**, 2767–2783 (2016).

35. Peelle, J. E., Gross, J. & Davis, M. H. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb. Cortex* **23**, 1378–1387 (2013).

36. Meyer, L. & Gumbert, M. Synchronization of Electrophysiological Responses with Speech Benefits Syntactic Information Processing. *J. Cogn. Neurosci.* **30**, 1066–1074 (2018).

37. Meyer, L., Henry, M. J., Gaston, P., Schmuck, N. & Friederici, A. D. Linguistic Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. *Cereb. Cortex* **27**, 4293–4302 (2017).

38. Vander Ghinst, M. *et al.* Left Superior Temporal Gyrus Is Coupled to Attended Speech in a Cocktail-Party Auditory Scene. *J. Neurosci.* **36**, 1596–1606 (2016).

39. Ding, N., Melloni, L., Zhang, H., Tian, X. & Poeppel, D. Cortical tracking of hierarchical linguistic structures in connected speech. *Nat. Neurosci.* **19**, 158–164 (2016).

40. Riecke, L., Formisano, E., Sorger, B., Başkent, D. & Gaudrain, E. Neural Entrainment to Speech

Modulates Speech Intelligibility. *Curr. Biol.* **28**, 161–169.e5 (2018).

41. Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z. & Francart, T. Speech intelligibility predicted from neural entrainment of the speech envelope. *Journal of the Association for Research in Otolaryngology* (2018) doi:10.1101/246660.

42. Ding, N. & Simon, J. Z. Cortical entrainment to continuous speech: functional roles and interpretations. *Front. Hum. Neurosci.* **8**, 311 (2014).

43. Fuglsang, S. A., Dau, T. & Hjortkjær, J. Noise-robust cortical tracking of attended speech in real-world acoustic scenes. *Neuroimage* **156**, 435–444 (2017).

44. Puschmann, S. *et al.* The Right Temporoparietal Junction Supports Speech Tracking During Selective Listening: Evidence from Concurrent EEG-fMRI. *J. Neurosci.* **37**, 11505–11516 (2017).

45. Rimmele, J. M., Golumbic, E. Z., Schröger, E. & Poeppel, D. The effects of selective attention and speech acoustics on neural speech-tracking in a multi-talker scene. *Cortex* vol. 68 144–154 (2015).

46. Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J. & Lalor, E. C. Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech. *Curr. Biol.* **28**, 803–809.e3 (2018).

47. Ding, N. & Simon, J. Z. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11854–11859 (2012).

48. Ding, N. & Simon, J. Z. Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *J. Neurosci.* **33**, 5728–5735 (2013).

49. Horton, C., D'Zmura, M. & Srinivasan, R. Suppression of competing speech through entrainment of cortical oscillations. *J. Neurophysiol.* **109**, 3082–3093 (2013).

50. Mesgarani, N. & Chang, E. F. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* **485**, 233–236 (2012).

51. O'Sullivan, J. A. *et al.* Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cereb. Cortex* **25**, 1697–1706 (2014).

52. Simon, J. Z. The encoding of auditory objects in auditory cortex: insights from magnetoencephalography. *Int. J. Psychophysiol.* **95**, 184–190 (2015).

53. Zion-Golumbic, E. & Schroeder, C. E. Attention modulates 'speech-tracking' at a cocktail party.

*Trends in Cognitive Sciences* vol. 16 363–364 (2012).

54. Pollack, I. Auditory informational masking. *The Journal of the Acoustical Society of America* vol. 57 S5–S5 (1975).

55. Sumby, W. H. & Pollack, I. Visual Contribution to Speech Intelligibility in Noise. *The Journal of the Acoustical Society of America* vol. 26 212–215 (1954).

56. Schwartz, J.-L., Berthommier, F. & Savariaux, C. Seeing to hear better: evidence for early audio-visual interactions in speech identification. *Cognition* **93**, B69–78 (2004).

57. Goswami, U. Sensory theories of developmental dyslexia: three challenges for research. *Nat. Rev. Neurosci.* **16**, 43–54 (2015).

58. Crosse, M. J., Di Liberto, G. M., Bednar, A. & Lalor, E. C. The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Front. Hum. Neurosci.* **10**, 604 (2016).

59. Bourguignon, M., Baart, M., Kapnoula, E. C. & Molinaro, N. Lip-reading enables the brain to synthesize auditory features of unknown silent speech. *J. Neurosci.* (2019) doi:10.1523/JNEUROSCI.1101-19.2019.

60. Zion-Golumbic, E. M. *et al.* Mechanisms underlying selective neuronal tracking of attended speech at a 'cocktail party'. *Neuron* **77**, 980–991 (2013).

61. Lalor, E. C. & Foxe, J. J. Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *Eur. J. Neurosci.* **31**, 189–193 (2010).

62. Daube, C., Ince, R. A. A. & Gross, J. Simple Acoustic Features Can Explain Phoneme-Based Predictions of Cortical Responses to Speech. *Curr. Biol.* **29**, 1924–1937.e9 (2019).

63. Ince, R. Measuring Multivariate Redundant Information with Pointwise Common Change in Surprisal. *Entropy* vol. 19 318 (2017).

64. Ince, R. A. A. *et al.* A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Human Brain Mapping* **38**, 1541–1573 (2017).

65. Molinaro, N. & Lizarazu, M. Delta(but not theta)-band cortical entrainment involves speech-specific processing. *European Journal of Neuroscience* vol. 48 2642–2650 (2018).

66. Allport, D. A. & Funnell, E. Components of the Mental Lexicon. *Philosophical Transactions of the*

41

*Royal Society B: Biological Sciences* vol. 295 397–410 (1981).

67. McClelland, J. L. & Rogers, T. T. The parallel distributed processing approach to semantic cognition. *Nat. Rev. Neurosci.* **4**, 310–322 (2003).

68. Ramus, F. The neural basis of reading acquisition. in *The Cognitive Neurosciences (3rd ed.)* (ed. Gazzaniga, M. S.) 815–824 (2004).

69. Ricketts, J., Davies, R., Masterson, J., Stuart, M. & Duff, F. J. Evidence for semantic involvement in regular and exception word reading in emergent readers of English. *J. Exp. Child Psychol.* **150**, 330–345 (2016).

70. Kaandorp, M. W., De Groot, A. M. B., Festen, J. M., Smits, C. & Goverts, S. T. The influence of lexical-access ability and vocabulary knowledge on measures of speech recognition in noise. *Int. J. Audiol.* **55**, 157–167 (2016).

71. Carroll, R., Warzybok, A., Kollmeier, B. & Ruigendijk, E. Age-Related Differences in Lexical Access Relate to Speech Recognition in Noise. *Front. Psychol.* **7**, 990 (2016).

72. Golumbic, E. Z., Zion Golumbic, E., Cogan, G. B., Schroeder, C. E. & Poeppel, D. Visual Input Enhances Selective Speech Envelope Tracking in Auditory Cortex at a 'Cocktail Party'. *Journal of Neuroscience* vol. 33 1417–1426 (2013).

73. Park, H., Ince, R. A. A., Schyns, P. G., Thut, G. & Gross, J. Representational interactions during audiovisual speech entrainment: Redundancy in left posterior superior temporal gyrus and synergy in left motor cortex. *PLoS Biol.* **16**, e2006558 (2018).

74. Park, H., Kayser, C., Thut, G. & Gross, J. Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *eLife* vol. 5 (2016).

75. Bourguignon, M., Baart, M., Kapnoula, E. C. & Molinaro, N. Hearing through lip-reading: the brain synthesizes features of absent speech. doi:10.1101/395483.

76. Giordano, B. L. *et al.* Contributions of local speech encoding and functional connectivity to audio-visual speech perception. *eLife* vol. 6 (2017).

77. MacLeod, A. & Summerfield, Q. Quantifying the contribution of vision to speech perception in noise. *Br. J. Audiol.* **21**, 131–141 (1987).

78. Helfer, K. S. & Freyman, R. L. The role of visual speech cues in reducing energetic and

informational masking. *J. Acoust. Soc. Am.* **117**, 842–849 (2005).

79. Crosse, M. J., Di Liberto, G. M. & Lalor, E. C. Eye Can Hear Clearly Now: Inverse Effectiveness in Natural Audiovisual Speech Processing Relies on Long-Term Crossmodal Temporal Integration. *J. Neurosci.* **36**, 9888–9895 (2016).

80. Hauswald, A., Lithari, C., Collignon, O., Leonardelli, E. & Weisz, N. A Visual Cortical Network for Deriving Phonological Information from Intelligible Lip Movements. *Curr. Biol.* **28**, 1453–1459.e3 (2018).

81. O'Sullivan, A. E., Crosse, M. J., Di Liberto, G. M. & Lalor, E. C. Visual Cortical Entrainment to Motion and Categorical Speech Features during Silent Lipreading. *Front. Hum. Neurosci.* **10**, 679 (2016).

82. Crosse, M. J. & Lalor, E. C. The cortical representation of the speech envelope is earlier for audiovisual speech than audio speech. *J. Neurophysiol.* **111**, 1400–1408 (2014).

83. Baart, M., de Boer-Schellekens, L. & Vroomen, J. Lipread-induced phonetic recalibration in dyslexia. *Acta Psychol.* **140**, 91–95 (2012).

84. Keetels, M., Bonte, M. & Vroomen, J. A Selective Deficit in Phonetic Recalibration by Text in Developmental Dyslexia. *Front. Psychol.* **9**, 710 (2018).

85. van Laarhoven, T., Keetels, M., Schakel, L. & Vroomen, J. Audio-visual speech in noise perception in dyslexia. *Developmental Science* vol. 21 e12504 (2018).

86. Bastien-Toniazzo, M., Stroumza, A. & Cavé, C. Audio-visual perception and integration in developmental dyslexia: An exploratory study using the McGurk effect. *Curr. Psychol. Lett.* **25**, (2010).

87. Rüsseler, J., Gerth, I., Heldmann, M. & Münte, T. F. Audiovisual perception of natural speech is impaired in adult dyslexics: an ERP study. *Neuroscience* **287**, 55–65 (2015).

88. Ramirez, J. & Mann, V. Using auditory-visual speech to probe the basis of noise-impaired consonant-vowel perception in dyslexia and auditory neuropathy. *J. Acoust. Soc. Am.* **118**, 1122–1133 (2005).

89. Campbell, R. The processing of audio-visual speech: empirical and neural bases. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 1001–1010 (2008).

90. van Atteveldt, N., Formisano, E., Goebel, R. & Blomert, L. Integration of Letters and Speech Sounds in the Human Brain. *Neuron* vol. 43 271–282 (2004).

91. Raij, T., Uutela, K. & Hari, R. Audiovisual Integration of Letters in the Human Brain. *Neuron* vol. 28 617–625 (2000).

92. Blomert, L. The neural signature of orthographic–phonological binding in successful and failing reading development. *NeuroImage* vol. 57 695–703 (2011).

93. Bernstein, L. E. & Liebenthal, E. Neural pathways for visual speech perception. *Frontiers in Neuroscience* vol. 8 (2014).

94. Leybaert, J. *et al.* Atypical audio-visual speech perception and McGurk effects in children with specific language impairment. *Front. Psychol.* **5**, 422 (2014).

95. Hood, M. & Conlon, E. Visual and auditory temporal processing and early reading development. *Dyslexia* **10**, 234–252 (2004).

96. Boets, B., Wouters, J., van Wieringen, A., De Smedt, B. & Ghesquière, P. Modelling relations between sensory processing, speech perception, orthographic and phonological ability, and literacy achievement. *Brain Lang.* **106**, 29–40 (2008).

97. Katz, R. B., Healy, A. F. & Shankweiler, D. Phonetic coding and order memory in relation to reading proficiency: A comparison of short-term memory for temporal and spatial order information. *Applied Psycholinguistics* vol. 4 229–250 (1983).

98. Shankweiler, D. The speech code and learning to read. *Journal of Experimental Psychology: Human Learning & Memory* vol. 5 531–545 (1979).

99. Brady, S., Shankweiler, D. & Mann, V. Speech perception and memory coding in relation to reading ability. *J. Exp. Child Psychol.* **35**, 345–367 (1983).

100. Araújo, S., Reis, A., Petersson, K. M. & Faísca, L. Rapid automatized naming and reading performance: A meta-analysis. *Journal of Educational Psychology* vol. 107 868–883 (2015).

101. Lervåg, A. & Hulme, C. Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychol. Sci.* **20**, 1040–1048 (2009).

102. Norton, E. S. & Wolf, M. Rapid automatized naming (RAN) and reading fluency: implications for understanding and treatment of reading disabilities. *Annu. Rev. Psychol.* **63**, 427–452 (2012).

103. Georgiou, G. K., Parrila, R., Cui, Y. & Papadopoulos, T. C. Why is rapid automatized naming related to reading? *J. Exp. Child Psychol.* **115**, 218–225 (2013).

104. Landerl, K. *et al.* Phonological Awareness and Rapid Automatized Naming as Longitudinal Predictors of Reading in Five Alphabetic Orthographies with Varying Degrees of Consistency. *Scientific Studies of Reading* vol. 23 220–234 (2019).

105. Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S. & Hecht, S. Contributions of Phonological Awareness and Rapid Automatic Naming Ability to the Growth of Word-Reading Skills in Second-to Fifth-Grade Children. *Scientific Studies of Reading* vol. 1 161–185 (1997).

106. Sprenger-Charolles, L., Siegel, L. S., Béchennec, D. & Serniclaes, W. Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: a four-year longitudinal study. *J. Exp. Child Psychol.* **84**, 194–217 (2003).

107. Elhassan, Z., Crewther, S. G. & Bavin, E. L. The Contribution of Phonological Awareness to Reading Fluency and Its Individual Sub-skills in Readers Aged 9- to 12-years. *Front. Psychol.* **8**, 533 (2017).

108. Boets, B. *et al.* Intact but less accessible phonetic representations in adults with dyslexia. *Science* **342**, 1251–1254 (2013).

109. Perfetti, C. A., Beck, I., Bell, L. C. & Hughes, C. Phonemic Knowledge and Learning to Read are Reciprocal: A Longitudinal Study of First Grade Children. *Merrill. Palmer. Q.* **33**, 283–319 (1987).

110. Demanez L, Dony-Closon B, Lhonneux-Ledoux E, Demanez JP. Central auditory processing assessment: a French-speaking battery. *Acta Otorhinolaryngol. Belg.* **57**, 275–290 (2003).

111. Currie, C. E., Elton, R. A., Todd, J. & Platt, S. Indicators of socioeconomic status for adolescents: the WHO Health Behaviour in School-aged Children Survey. *Health Educ. Res.* **12**, 385–397 (1997).

112. Jacquier-Roux, M., Valdois, S. & Zorman, M. *Odédys: outil de dépistage des dyslexies*. (2005).

113. Lefavrais, P. *Manuel du test de l'alouette*. (2005).

114. Gauthier, L., Dehaut, F. & Joanette, Y. Bells Test. *PsycTESTS Dataset* (1989) doi:10.1037/t28075-000.

115. Fimm, B. & Zimmermann, P. A test battery for attentional performance. in *Applied Neuropsychology of Attention. Theory, Diagnosis and Rehabilitation* 110–151 (2002).

116. Wechsler, D. & Naglieri, J. A. *WNV: Wechsler Nonverbal Scale of Ability ; Administration and Scoring Manual*. (2006).

117. De Tiège, X. *et al.* Recording epileptic activity with MEG in a light-weight magnetic shield. *Epilepsy Res.* **82**, 227–231 (2008).

118. Taulu, S. & Simola, J. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Phys. Med. Biol.* **51**, 1759–1768 (2006).

119. Taulu, S., Simola, J. & Kajola, M. Applications of the signal space separation method. *IEEE Trans. Signal Process.* **53**, 3359–3372 (2005).

120. Biesmans, W., Das, N., Francart, T. & Bertrand, A. Auditory-Inspired Speech Envelope Extraction Methods for Improved EEG-Based Auditory Attention Detection in a Cocktail Party Scenario. *IEEE Trans. Neural Syst. Rehabil. Eng.* **25**, 402–412 (2017).

121. R Core Team. *R: A language and environment for statistical computing*. vol. 3 (R Foundation for Statistical Computing, Vienna, Austria, 2018).

122. Bates, D. M., Maechler, M. & Bolker, B. *lme4: Linear mixed-effects models using S4 classes*. (2011).

## 6. Acknowledgements

## Supplementary Material

## Supplementary methods

### *Recording of video stimuli*

The 12 video stimuli of actors telling a story were recorded with a digital camera (Sony handycam, HDR-CX115E). Audio signals were recorded with both the internal microphone of the camera, and an independent high quality microphone (Sony linear PCM recorder, PCM-D50). Audio tracks were synchronized, and only the high-quality audio was kept. Video recordings were framed as a head-shots, and recorded at 50 frames per second (videos were 1920 × 1080 pixels in size, 24 bits/pixel, with an auditory sampling rate of 44100 Hz). The camera was placed ~1 m away from the actors, and the face spanned about half of the vertical field of view. Final images were resized to a resolution of 1152 × 864 pixels. A black old-style-TV-monitor frame was then added to the image.

### *Building the energetic non-informational noise*

The (maximally-)energetic non-informational noise was derived from the actual actors' audio recording by i) Fourier transforming the sound in 2-s-long windows sliding by step of 0.5 s, ii) replacing the phase by random numbers, iii) inverse Fourier transforming the Fourier coefficients in each window, iv) multiplying these phase-shuffled sound segments by a sine window (i.e., half a sine cycle with 0 at edges, and 1 in the middle), and v) summing the contribution of each overlapping window. As a result, the spectral properties of this noise dynamically changed on a ~0.5-s time scale to mirror those of the actor's voice ~1 s around.

### *Accuracy of speech envelope reconstruction*

For each condition and participant, a global value of cortical tracking of the attended speech was evaluated for all left-hemisphere sensors at once, and for all right-hemisphere sensors at once.

The decoder tested on a given condition was built based on MEG data from all the other conditions. This procedure was preferred over a more conventional cross-validation approach in which the decoder is trained and tested on separate chunks of data from the same condition because of the paucity of data (*i.e.*, at most ~2.4 min of data per condition). It is based on the rationale that the different conditions do modulate response amplitude but not its topography and temporal dynamics. In practice, electrophysiological data were band-pass filtered at 0.2–1.5 Hz (phrasal rate) or 2–8 Hz (syllabic rate), resampled to 10 Hz (phrasal) or 40 Hz (syllabic) and standardized. The decoder was build based on MEG data from –500 ms to 1000 ms (phrasal) or from 0 ms to 250 ms (syllabic) with respect to speech temporal envelope. Filtering and delay ranges were as in previous studies for phrasal,[28,59] and syllabic CTS.[47,51,60,61] Regularization was applied to limit the norm of the derivative of the reconstructed speech temporal envelope,[58] by estimating the decoder for a fixed set of ridge values ($\lambda = 2^{-10}, 2^{-8}, 2^{-6}, 2^{-4}, 2^{-2}, 2^{0}$). The regularization parameter was determined with a classical 10-fold cross-validation approach: the data is split into 10 segments of equal length, the decoder is estimated for 9 segments and tested on the remaining segment, and this procedure is repeated 10 times until all segments have served as test segment. The ridge value yielding the maximum mean RA is then retained. The ensuing decoder was then used to reconstruct speech temporal envelope in the left-out condition. RA was then estimated in 10 disjoint consecutive segments. We then retained the mean of this RA, leaving us with one value for all combinations of subjects, conditions, hemispheres, and frequencies of interest.

Significance of RA in each participant, condition, hemisphere and frequency range was assessed with a *t*-test on the RA values evaluated on 10 disjoint segments.

### *Preprocessing of brain and behavioral indices*

All behavioral and nCTS measures were corrected for IQ, age, and time spent at elementary school, and for outliers. For simplicity, we refer to the standardized IQ, age, and time spent at elementary school as the regressors.

An iterative procedure was used to simultaneously control for regressors and deal with outliers. First regressors were regressed out of each measure with an amount of regularization equal to 0.1% the maximal eigenvalue of the regressors' covariance matrix. Then measures deviating by more than 3 standard deviations from the mean were removed from the distribution. This procedure was repeated until there were no more outliers. Discarded data points were then set to the mean plus or minus 3 standard deviations.

### *Extraction of the relevant features of nCTS*

In total, we derived 8 features of nCTS in SiN conditions based on the significant effects of hemisphere and conditions highlighted in Table 3 and Figure 2. There were 4 features for phrasal nCTS: (i) the mean nCTS (the mean standardized nCTS across conditions), (ii) the informational modulation in nCTS (the difference in standardized *nCTS* between informational and non-informational noise conditions averaged across all other factors), (iii) the visual modulation in *nCTS* (the difference in standardized nCTS between *lips* and *pics* visual conditions averaged across all informational noise conditions), and (iv) the hemispheric difference in nCTS (the contrast in standardized nCTS between *left* and *right* hemispheres averaged across all informational noise conditions). The same 4 features were used for syllabic nCTS except that visual and hemispheric modulations were evaluated based on averages across all other factors (and not just across informational noise conditions since visual and hemispheric modulations were seen in all noise conditions for syllabic CTS).

### *Partial information decomposition*

PID was previously used to decompose the information brought by acoustic and visual speech signals about brain oscillatory activity,[73] and to compare auditory encoding models of MEG during speech processing.[62] As in these references, we measure redundancy with pointwise common change in surprisal for Gaussian variables.[63] Continuous data values were first subject to a rank-normalisation (copula-normalised)[64] before being treated as Gaussian variables. A crucial advantage of this redundancy measure as opposed to other PID implementations is that it measures the overlapping information content at the pointwise level and therefore can be interpreted as a within sample (here participant) measure of redundant prediction, directly linked to the coding interpretations of MI. An advantage of the PID over variance partitioning approaches is that unique variance explained might, like conditional mutual information, be confounded by synergistic effects,[123] whereas PID with common change in surprisal gives the true unique contribution. While the PID has so far been applied within subject to trial data, information theoretic quantities can also be applied as a second level analysis, where each participant is a sample.[124]

The statistical significance of the different information values was assessed with a nonparametric permutation test.[125] A permutation distribution was computed for each information value by randomly shuffling (10,000 times) children's target values, and a significance level was computed as the proportion values from the permutation distribution exceeding the observed value.

## Supplementary results

### *Side measures are redundant with RAN and digit span but not with modulations in phrasal nCTS*

The information about reading brought by the 3 "side" measures (visual modulation in syllabic nCTS, phoneme suppression and phoneme fusion) was redundant with that brought by a subset (possibly all) of the 4 "main" measures (RAN, phonological memory, visual and

informational modulation in phrasal nCTS). To identify this subset, we relied on the PID framework. In this analysis, PID assessed the nature of the information about reading abilities (target) brought by the 3 side measures (first set of explanatory variables) and the 4 main measures (second set of explanatory variables). Specifically, we considered as second set of explanatory variables all possible combinations of the 4 main measures (with or without RAN, with or without forward digit span, with or without the informational modulation in phrasal nCTS, and with or without the visual modulation in phrasal nCTS). Also, the analysis was run separately for the modulation in syllabic nCTS, and for both measures of phonological awareness (phoneme suppression and phoneme fusion) at once. Overall, the results tended to show that the visual modulation in syllabic nCTS provided unique information about reading only with regard to the two phrasal nCTS modulations (see Fig. S1A), and synergic information mainly with the two classical behavioral predictors of reading (see Fig. S1B). Noticeably, the level of synergic information about reading brought by the visual modulation in syllabic nCTS and either or both classical behavioral predictors of reading was not influenced by the addition of either or both of the two phrasal nCTS modulations. Similar results were obtained for the two measures of phonological awareness (see Fig. S1C & D).

**Supplementary discussion**

*Limitations*

We did not manipulate the SNR in SiN conditions in the present study. Instead, it was set to 3 dB so that the attended speech was always louder than the noise. Still, one could expect that large effect sizes would be uncovered in more challenging/discriminating listening conditions, as often encountered in classrooms.[126] However, children's SiN perception abilities are lower than adults'. Indeed, SiN perception abilities develop until late childhood ($\geq$ 10 years) due to maturation of the auditory system and attentional abilities.[127,128] Moreover, our data showed that fewer subjects

had significant CTS in informational than non-informational SiN conditions for phrasal CTS, and in all SiN conditions for syllabic CTS, with values in the most challenging SiN condition (energetic and informational noise) that were ~50% lower than those in the noiseless condition. Accordingly, setting speech SNR to 3 dB appeared to have made the task challenging enough for children, while ensuring that they could keep their attention focused throughout the experiment.

The amount of data per condition was limited to 2.5 min. Although it may seem little, we evidenced in a previous study that on average, ~30 s of MEG suffice to uncover significant CTS.[28] Moreover, CTS was significant, when assessed non-parametrically within-participant, in most of our participants in the least challenging conditions (phrasal, 100 %; syllabic, 94%). Of note, more data was used to estimate the regression model mapping MEG data onto reconstructed speech temporal envelope. Indeed, the model used to estimate CTS in each condition was trained on the ~20 min of data from all other conditions. This procedure substantially improved the estimation of CTS compared with a procedure wherein the model was trained and tested in a cross-validation scheme on the data from each condition separately (data not shown). Still, longer data acquisition could have produced more stable CTS estimates, and perhaps stronger associations with reading scores.

## Supplementary references

123.    Daube, C., Giordano, B., Schyns, P. & Ince, R. Quantitatively comparing predictive models with the Partial Information Decomposition. *2019 Conference on Cognitive Computational Neuroscience* (2019) doi:10.32470/ccn.2019.1142-0.

124.    Jaworska, K. *et al.* Healthy ageing delays the neural processing of face features relevant for behaviour by 40 ms. *BioRxiv* **3260092**, (2019).

125.    Nichols, T. E. & Holmes, A. P. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* **15**, 1–25 (2002).

126.    Picard, M. & Bradley, J. S. Revisiting speech interference in classrooms. *Audiology* **40**, 221–244

(2001).

127.    Thompson, E. C., Woodruff Carr, K., White-Schwoch, T., Otto-Meyer, S. & Kraus, N. Individual differences in speech-in-noise perception parallel neural speech processing and attention in preschoolers. *Hear. Res.* **344**, 148–157 (2017).

128.    Sanes, D. H. & Woolley, S. M. N. A behavioral framework to guide research on central auditory development and plasticity. *Neuron* **72**, 912–929 (2011).

## Supplementary tables and figures

**Table S1.** Nature of the information about reading abilities brought by each of the 3 uncovered features of the cortical tracking of speech (CTS) in noise and phonological awareness (mean of the scores for phoneme fusion and suppression). Significant values ($p < 0.05$) are displayed in boldface and marginally significant values are displayed in boldface and italicized.

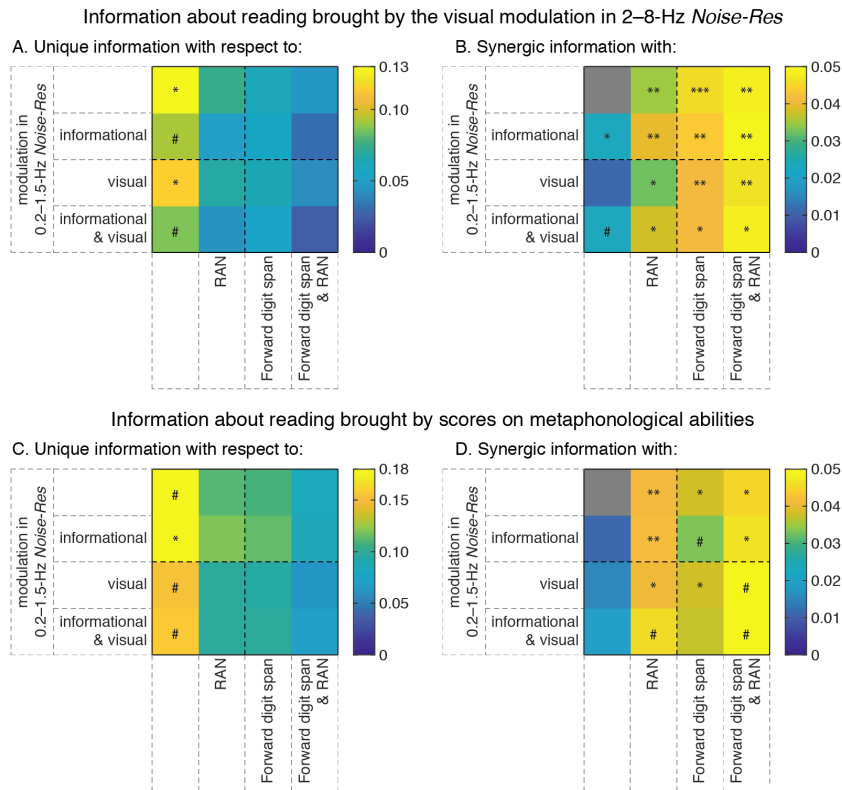| | redundant | | unique for phonological awareness | | unique for the feature of CTS in noise | | synergic | |
|---|---|---|---|---|---|---|---|---|
| | info | *p* | info | *p* | info | *p* | info | *p* |
| informational modulation in phrasal *nCTS* | 0.0011 | 0.57 | **0.148** | **0.013** | **0.160** | **0.006** | *0.0073* | *0.097* |
| visual modulation in phrasal *nCTS* | 0.0066 | 0.30 | *0.109* | *0.057* | **0.155** | **0.011** | 0.0061 | 0.33 |
| visual modulation in syllabic *nCTS* | **0.048** | **0.0006** | 0.085 | 0.16 | **0.116** | **0.036** | **0.032** | **0.001** |

**Figure S1.** Nature of the information about reading brought by the visual modulation in normalized cortical tracking of speech (nCTS) at syllabic rate (A & B) and metaphonological abilities (C & D). **A & C** — Unique information with regard to each possible combination of the 4 regressors included in the final model of reading abilities (without RAN: columns 1 and 3; with RAN: columns 2 and 4; without forward digit span: columns 1 and 2; with forward digit span: columns 3 and 4; without the informational modulation in phrasal *nCTS*: rows 1 and 3; …). **B & D** — Same as A for the synergistic information with each possible combination of the 4 regressors. ** $p < 0.01$, * $p < 0.05$, # $p < 0.1$.