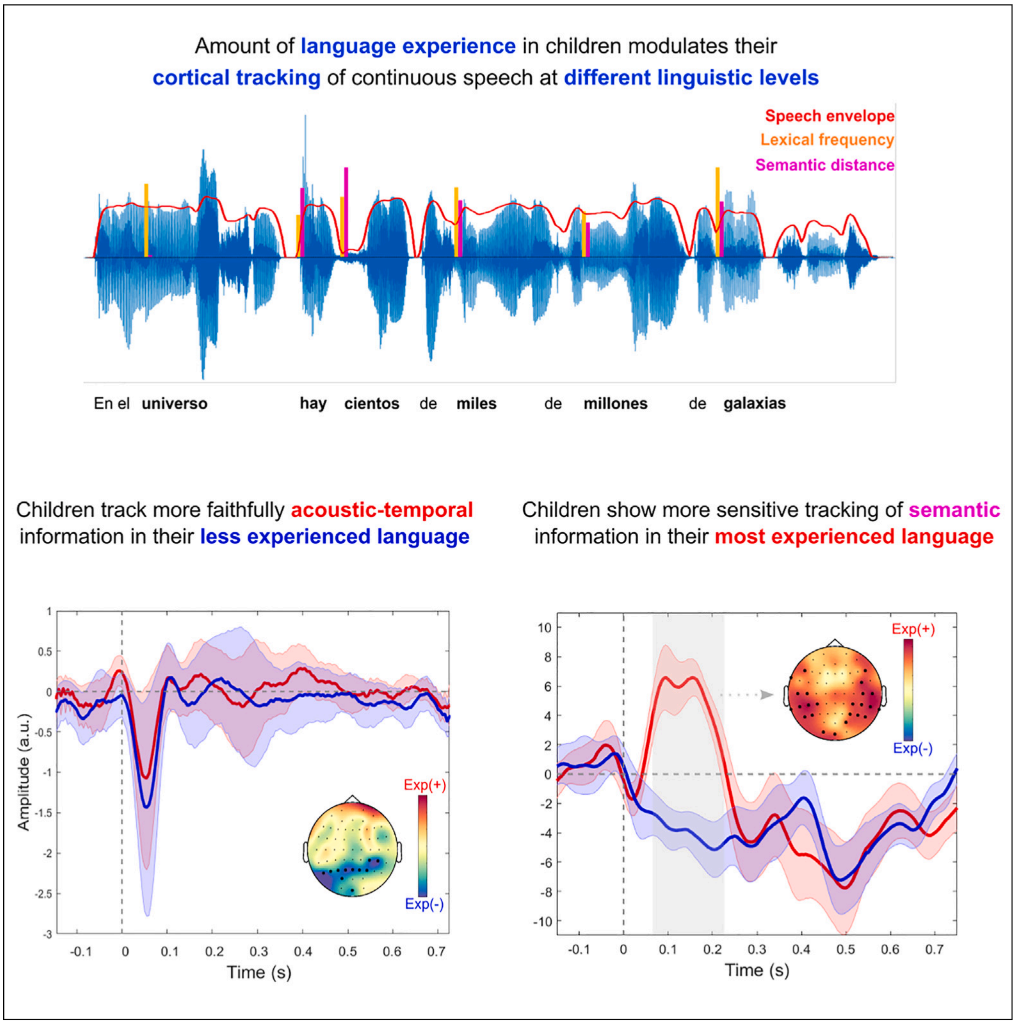
iScience

CellPress
OPEN ACCESS

Article

Early language experience modulates the tradeoff between acoustic-temporal and lexico-semantic cortical tracking of speech



Jose Pérez-Navarro, Anastasia Klimovich-Gray, Mikel Lizarazu, Giorgio Piazza, Nicola Molinaro, Marie Lallier

jperez@bcbl.eu

Highlights

Amount of linguistic exposure in childhood modulates cortical tracking of speech

Cortical tracking of the envelope is more robust is less experienced language

Lexico-semantic cortical tracking is more sensitive in most experienced language

Envelope tracking is linked phonology, lexico-semantic tracking to vocabulary

Pérez-Navarro et al., iScience *27*, 110247
July 19, 2024 © 2024 Published by Elsevier Inc.
https://doi.org/10.1016/j.isci.2024.110247

iScience

CellPress OPEN ACCESS

Article

Early language experience modulates the tradeoff between acoustic-temporal and lexico-semantic cortical tracking of speech

Jose Pérez-Navarro,[1,5,*] Anastasia Klimovich-Gray,[2] Mikel Lizarazu,[1] Giorgio Piazza,[1] Nicola Molinaro,[1,3,4] and Marie Lallier[1,4]

SUMMARY

Cortical tracking of speech is relevant for the development of speech perception skills. However, no study to date has explored whether and how cortical tracking of speech is shaped by accumulated language experience, the central question of this study. In 35 bilingual children (6-year-old) with considerably bigger experience in one language, we collected electroencephalography data while they listened to continuous speech in their two languages. Cortical tracking of speech was assessed at acoustic-temporal and lexico-semantic levels. Children showed more robust acoustic-temporal tracking in the least experienced language, and more sensitive cortical tracking of semantic information in the most experienced language. Additionally, and only for the most experienced language, acoustic-temporal tracking was specifically linked to phonological abilities, and lexico-semantic tracking to vocabulary knowledge. Our results indicate that accumulated linguistic experience is a relevant maturational factor for the cortical tracking of speech at different levels during early language acquisition.

INTRODUCTION

Continuous exposure to spoken language in a wide variety of contexts makes its acquisition appear spontaneous and effortless. Despite this apparent ease, developmental evidence shows that complex brain systems and processes supporting language comprehension emerge between perinatal stages and around three years of age, and become language-selective between birth and the start of primary school (for reviews see in the study by Gervain et al., Kuhl et al., and Skeide et al.[1–3]). However, little is known about how these neurocognitive processes evolve and mature as a function of language experience, an established contributor to language development.[4–9]

Accurate multidimensional models of language development require understanding how different levels of language exposure and proficiency modulate fundamental brain mechanisms underlying language acquisition, including speech comprehension. We focus on the cortical tracking of speech (CTS),[10] a neurocognitive process relevant for understanding speech as it unfolds over time. CTS typically refers to the dynamic alignment of brain activity to the temporal modulations of the speech envelope.[11] In this study we refer to such speech envelope tracking as *acoustic-temporal* CTS, which takes place within delta (aligning with prosodic phrasing) and theta (syllable timing) frequency bands and has shown to support speech comprehension in adults.[12–16] Beyond the speech envelope, a growing number of studies are investigating CTS at higher order lexical and semantic levels.[17–19] Here, we term this process as *lexico-semantic* CTS. We hypothesize that early language experience shapes the maturation of CTS at both acoustic-temporal and lexico-semantic levels. In fact, behavioral evidence shows that language experience supports increasingly efficient processing and comprehension skills for continuous speech. For example, in monolingual environments, the amount of child-directed input is positively linked to word processing speed, the ability to encode lexical information from continuous speech input.[20,21] In addition, in bilingual contexts, the proportional exposure to both languages correlates with language-specific word processing speed.[22]

We propose that language experience modulates CTS at two different levels. First, accumulated input in a language should be crucial for building up knowledge about the temporal statistics embedded in the speech envelope, which contributes to the efficient dynamic alignment of cortical oscillatory activity to relevant phonological units in the speech signal.[10,14] Accordingly, greater language experience should be associated with more efficient acoustic-temporal CTS. However, the two studies that have directly assessed this relationship in adult acoustic-temporal CTS show to an extent divergent results, with language experience and CTS being either positively[23] or negatively related.[24] Thus, the directionality of the relationship between language experience and acoustic-temporal CTS is unclear and worth further exploration.

[1]Basque Center on Cognition, Brain and Language (BCBL), 20009 Donostia-San Sebastian, Spain
[2]School of Psychology, University of Aberdeen, King's College, Aberdeen AB24 3FX, UK
[3]Ikerbasque, Basque Foundation for Science, 48009 Bilbao, Spain
[4]These authors contributed equally
[5]Lead contact
*Correspondence: jperez@bcbl.eu
https://doi.org/10.1016/j.isci.2024.110247

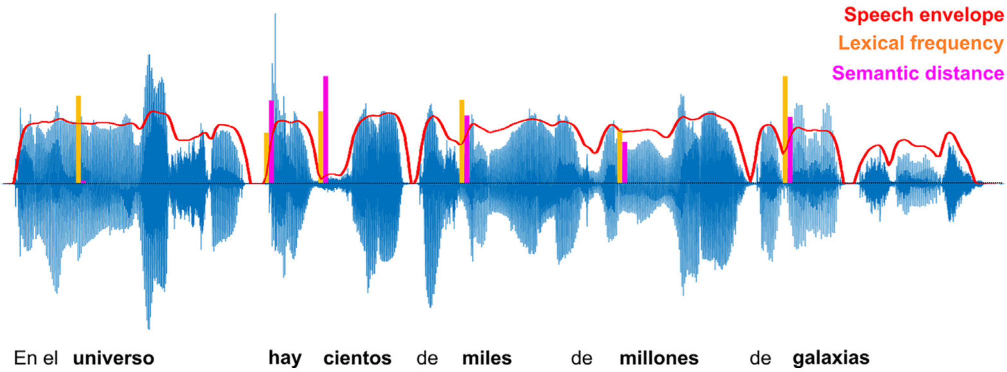


Figure 1. Graphical summary of the CTS analyses

In blue, the speech waveform of "En el universo, hay cientos de miles de millones de galaxias" ("In the universe, there are hundreds of billions of galaxies"). Speech-brain coherence and envelope-level mTRF models are based on the relationship between the speech envelope (in red) and EEG activity. Lexical frequency and semantic distance mTRFs are obtained from the EEG response to bursts of different amplitude (lexical frequency, orange; sentence-level semantic distance, pink) at the onset of each content word.

Second, richer language experience should support the development of a neurocognitive model guiding the extraction of higher order linguistic information via lexico-semantic CTS.[19,25] Therefore, linguistic experience should also be positively associated with the cortical tracking of lexico-semantic information.

Language experience supports the development of language knowledge across linguistic domains.[6,26–28] In particular, accumulated language experience enhances phonological acquisition (Nittrouer,[29] 1996; Nittrouer & Burton,[30] 2005; for a review, see Nittrouer,[31] 2002) which, in turn, supports the comprehension of continuous speech. Of particular relevance for the purpose of our study, Goswami[32,33] proposes that the CTS amplitude modulations within the speech envelope (acoustic-temporal CTS) supports phonological development during infancy and childhood. Within this framework, the accurate processing of the speech envelope is thought to guide and shape the emergence of phonological representations.[34,35] Accordingly, a growing number of studies have now shown that acoustic-temporal CTS is at play as early as 4 months of age and developing throughout childhood.[36–38]

Studies of atypical language development also suggest that CTS mediates the contribution of language experience to the acquisition of adequate linguistic abilities. In addition, several electrophysiological studies have shown that phonological deficits in children with dyslexia are tightly linked to atypical CTS within the delta and theta bands[39–43] in comparison to both chronological-age-matched and reading-age-matched younger peers who are also matched on how much they have been exposed to written inputs.[40,43] This suggests that impaired CTS could be directly implicated in atypical phonological development, and be somehow independent from linguistic experience (specially for written language). However, Destoky et al.[39] reported poorer CTS in children with dyslexia only when compared to chronological and not to reading age-matched controls. It is thus unclear whether impaired CTS is a cause or consequence of phonological deficits, as well as to what extent accumulated language experience contributes to the maturation of CTS during phonological development. Our study tackles the latter question by exploiting the largely different experiences that bilingual children have within their two languages. Importantly, this group provides a unique opportunity to quantify the effect of accumulated language experience on acoustic envelope-level CTS and phonological development within the same participants. To gain a more comprehensive understanding of the relevance of CTS for the development of language skills, we go a step forward by exploring whether CTS is also linked to lexico-semantic knowledge through accumulated experience within a particular language.

Lexico-semantic knowledge, whose development is tightly related to accumulated language experience during childhood,[4,8,44,45] has been linked to the efficiency of cortical mechanisms for speech processing. Recent studies have shown that, in adults, lexico-semantic CTS is modulated by context-driven word predictability.[19,46–48] Of particular relevance for our hypotheses, these studies converge on the finding that the predictability of lexico-semantic information shapes CTS in adults, possibly as they have developed efficient neurocognitive language models as a function of accumulated language experience. Indeed, a recent study in adults showed that the encoding of acoustic features by magnetoencephalographic activity was boosted in a language unknown to the listener, while linguistic features were in turn more efficiently encoded in the native language.[49] However, the contribution of language experience to the development of lexico-semantic CTS during childhood is an open question that, to our knowledge, remains poorly understood. By addressing such a question, we will be able to add evidence about a potential tradeoff between acoustic-temporal and lexico-semantic CTS from a developmental perspective. In other words, whether depending on accumulated language experience, children either rely more on basic acoustic-temporal speech features or on higher-order lexico-semantic information to comprehend continuous speech in the language (or languages) they are learning.

Overall, the evidence reviewed previous points to the relevance of accumulated linguistic experience, known to be crucial for language knowledge, in the tuning also of CTS at both the acoustic-temporal level (linked to phonology) and higher order linguistic levels (associated with lexico-semantics) (Figure 1). Here, we study the relationship between input and CTS at both levels by capitalizing on a single group of bilingual children whose accumulated experience varies greatly between their two languages. Therefore, we are able to explore the role of

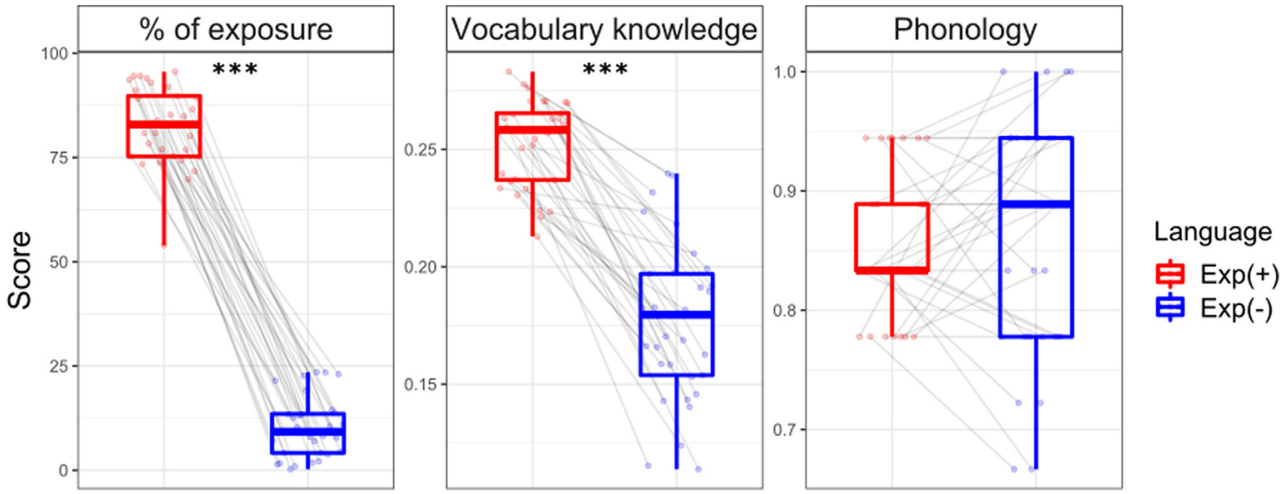


Figure 2. Amount of linguistic exposure and performance on vocabulary (picture naming) and phonology (nonword repetition) tasks in each language
Points represent each individual score in the different measures and languages. Boxplots represent group estimates, with horizontal lines within each box marking the median score. Upper and lower hinges mark the first and third quartile, and whiskers show 1.5 * interquartile range. Lines connect the scores of each participant across both languages. Asterisks indicate a significant difference between languages (***$p < 0.001$).

greatly different language experiences on the cortical tracking of acoustic-temporal speech information (i.e., the envelope), and on the much less investigated mechanism of cortical tracking of lexico-semantic information in continuous speech. In addition, we assess whether the cortical tracking of these two types of speech information is related to phonological and lexico-semantic skills measured behaviorally.

RESULTS

In line with their significantly higher language experience (i.e., percentage of exposure) within Exp(+) than within Exp(−), $t(63) = 35.13$, $p < 0.001$ ($\beta = 72.298$, $SE = 2.058$), participants showed significantly higher vocabulary knowledge in Basque than in Spanish, $t(32) = 13.18$, $p < 0.001$ ($\beta = 0.077$, $SE = 0.006$; see Figure 2). However, there was no between-language difference in phonological abilities, $t(30.1) = -0.84$, $p > 0.05$ ($\beta = -0.014$, $SE = 0.017$), nor in the comprehension of the stories that participants listened to during the electroencephalography (EEG) session, $t(28) = 0.12$, $p > 0.05$, ($\beta = 0.005$, $SE = 0.039$). Indeed, Bayesian t-tests yielded moderate evidence for between-language similarities in phonological abilities, $BF_{10} = 0.285$, *error* = 0.034%, *median difference* = −0.151, *CI* [−0.492 0.184], and story comprehension respectively, $BF_{10} = 0.22$, *error* = 0.024%, *median difference* = −0.04, *CI* [−0.417 0.334].

Increased language experience is linked to reduced acoustic-temporal CTS

Coherence

Within the delta frequency band, speech-brain coherence was significant between 0.5 and 1.5 Hz in Exp(+), *cluster statistic* = 271.73, $p < 0.001$, $SD = 0.001$, and Exp(−), *cluster statistic* = 278.17, $p = 0.001$, $SD = 0.001$ (Figure 3A). However, there was no between-language difference. Coherence in this delta range had considerably overlapping topographies in both languages (Figure 3B).

In the theta range (4–7 Hz), we did not find significant speech-brain coherence in any of the languages (all cluster-corrected *p values* > 0.05). We also did not observe significant coherence in the specific 5.5 Hz bin that aligned closely to the syllable rate of both languages ($p > 0.05$).

mTRF

Our control tests, one-sample t-tests (one-tailed) against no different prediction correlation from 0, showed that our three regressors of interest contributed above chance level to the multivariate temporal response function (mTRF) model (speech envelope: $t(65) = 12.78$, $p < 0.001$, $d = 1.57$; lexical frequency: $t(65) = 6.50$, $p < 0.001$, $d = 0.8$; semantic distance: $t(65) = 5.27$, $p < 0.001$, $d = 0.65$; Figure S3), which enabled us to conduct between-languages comparisons of their fit.

Regarding acoustic-temporal CTS, CBPT yielded significantly higher prediction correlation coefficients (Pearson's r) between envelope and EEG mTRF in the less exposed language (Exp(-)) than in the more exposed one (Exp(+)), *cluster statistic* = −27.65, $p = .01$, $SD = .004$ (Figure 4B). Regarding the temporal weights in response to the speech envelope, CBPT did not show significant differences between languages (*cluster-corrected p-values* > .05; Figure 4A). Although we did not plan a statistical comparison of the temporal response to the envelope in the electrodes that showed a significant difference in prediction correlation, a visual inspection suggests that a more sustained response to the speech envelope (Figure S4) in Exp(-) could be driving such significantly bigger prediction correlation. This suggests differences in how strongly children rely on envelope tracking between their two languages (i.e., stronger reliance on envelope tracking in their less

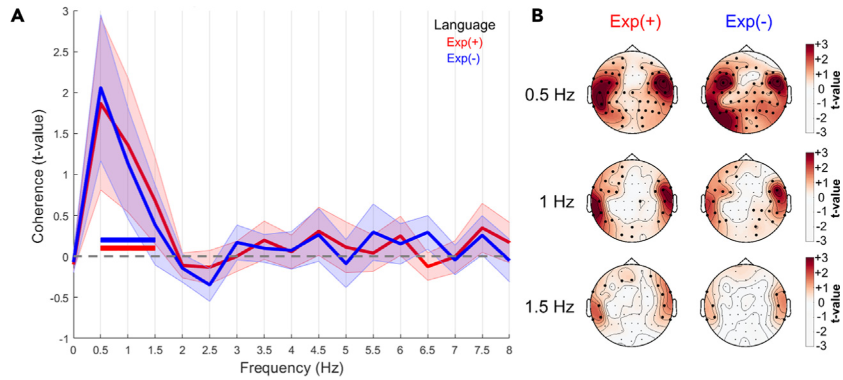


Figure 3. Speech-brain coherence

(A) Speech-brain coherence across frequency bands for Exp(+) (red), Exp(−) (blue) against their random surrogate (flipped version). The continuous horizontal lines mark the frequency range in which there was significant coherence for Exp(+) (red), and Exp(−) (blue). Red and blue shaded areas represent the standard error.

(B) Topography of speech-brain coherence in the 3 significant frequencies in Exp(+) (left) and Exp(−) (right). The colormap marks the size of the difference in coherence (normalized) between genuine speech and its flipped version (i.e., red, higher coherence; white, lower). Bigger black circles within each topographic map signal electrodes with significant speech-brain coherence from CBPT.

experienced language), but no widespread between-language differences in the temporal characteristics of the response to the envelope. A potential source of such divergent findings in the prediction correlation and the temporal latency of envelope-level TRFs is the employment of CBPT in the whole TRF window as a statistical test of significance, contrary to the a-priori or visual selection of a specific time window for testing the presence of a significant effect that is typically employed in mTRF research (e.g., Broderick et al.[19]). CBPTs apply a spatiotemporal correction for multiple comparisons that could limit small effects (i.e., the 100 ms peak) from reaching significance.

Language experience increases semantic sensitivity of CTS

Cluster-based permutation test (CBPT) did not yield significant between-language differences in the prediction correlation coefficients (model-fit r values derived from mTRF) between lexical frequency and EEG activity (*cluster-corrected p values* > 0.05). The temporal weights in response to lexical frequency showed a similar latency to lexical-level TRFs in adults[19] and did not show significant between-language differences (*cluster-corrected pvalues* > 0.05; Figure 5A).

Regarding EEG responses to semantic distance, we found relevant differences between languages. While mTRF prediction correlation coefficients (R values) for semantic distance did not differ between languages (*cluster-corrected p values* > 0.05); CBPT yielded a significant difference in the temporal response (i.e., weights) to semantic distance between the Exp(+) and Exp(−) languages (*cluster statistic* = 2827.4, $p = 0.009$, $SD = 0.003$). Thus, within this group of participants, EEG signal showed significantly higher sensitivity to semantic information in Exp(+) than in Exp(−) between 70 and 230 ms after word onset (Figure 5B). This suggests earlier tracking of the relevant contextual-semantic information in the language to which bilingual children were exposed the most. The divergent results in prediction correlation (non-significant) and temporal response (significant) to semantic distance should be interpreted carefully, since prediction correlations in mTRF models explain the unique predictive power of a regressor, and temporal responses' unique contribution to an mTRF model cannot be as easily disentangled from related regressors.[50] However, the considerably different TRFs to lexical frequency and semantic distance regressors point to a genuinely different response to semantic distance in Exp(+) driving our significant between-languages difference.

Acoustic-temporal CTS is specifically linked to phonological abilities, lexico-semantic CTS to vocabulary knowledge

Regarding the relationship between acoustic-temporal CTS and behavioral language measures, there was a significant interaction between phonological abilities and language as predictors of acoustic-temporal CTS ($F[2, 46.52] = 4.69$, $p < 0.014$). This interaction was driven by the significant and positive relationship between acoustic-temporal CTS and phonological abilities (nonword repetition) in Exp(+) ($t[45.69] = 3.05$, $p = 0.004$ [$\beta = 0.62$, $SE = 0.21$]), but not in Exp(−) ($t[48.93] = 0.62$, $p > 0.05$ [$\beta = 0.09$, $SE = 0.14$]) (Figure 6). There was no significant relationship between envelope-CTS and vocabulary knowledge in either of the two languages (all *p values* > 0.05).

In parallel to the relationship between phonological abilities and envelope-level CTS, we found that vocabulary knowledge was positively related to lexico-semantic CTS. In this case the interaction between vocabulary knowledge and language was only marginally significant ($F[2, 57] = 2.81$, $p = 0.068$). Nonetheless, this interaction was also driven by an underlying significant and positive relationship between vocabulary knowledge and lexico-semantic CTS only in Exp(+), $t(57) = 2.37$, $p = 0.021$ ($\beta = 0.41$, $SE = 0.17$), and not in Exp(−), $t(57) = -0.006$, $p > 0.05$ ($\beta = -0.001$, $SE = 0.14$) (Figure 6). There was no significant relationship between lexico-semantic CTS and phonological abilities in either of the two languages (all *p values* > 0.05).

It is relevant to note that, while acoustic-temporal and lexico-semantic CTS composite scores were not significantly correlated in Exp(+) (*Pearson's r coefficient* = 0.29, $p > 0.05$), they showed a significant and positive correlation in Exp(−) ($r = 0.41$, $p = 0.02$).

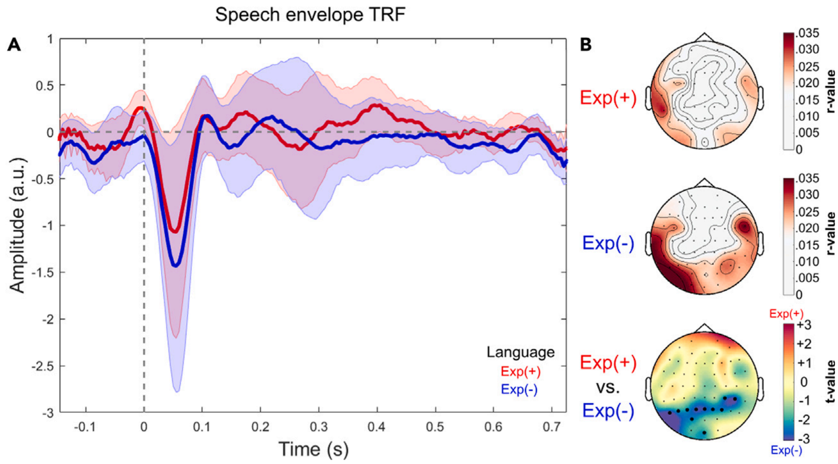


Figure 4. Speech envelope temporal response function

(A) Temporal weights of the EEG signal TRF to the speech envelope in Exp(+) (red) and Exp(−) (blue) across all electrodes. Red and blue shaded areas represent the standard error.

(B) Topography of the prediction correlation coefficients (R values) between speech envelope and EEG TRF for Exp(+) (top), Exp(−) (middle), and their difference (bottom). The colormap for the two top topographies marks the degree of correlation (R values); while the bottom topography is color mapped according to the size of the difference (t-value) between Exp(+) (relatively higher correlation coefficients in red) and Exp(−) (relatively higher correlation coefficients in blue). Bigger black circles signal electrodes in which CBPT yielded a significantly bigger prediction correlation at envelope level in Exp(−).

DISCUSSION

The main goal of the present study was to estimate the impact of early linguistic experience on the CTS, by assessing the CTS of 6-year-old bilingual children with considerably different amounts of experience between both of their languages. We found relevant differences between the more experienced language (Basque) and the less experienced language (Spanish) in the cortical tracking of acoustic-temporal (speech envelope) and lexico-semantic speech features. Such differences can be understood from a developmental perspective in which linguistic input plays a relevant role in the maturation of the cortical tracking of different types of linguistic information encoded in the speech signal. Additionally, and given that CTS was linked to language knowledge (in vocabulary and phonology) at behavioral level, our results inform the relevance of CTS for indexing the development of specific language abilities. Moreover, the present study is, to our knowledge, the first to assess the cortical tracking of two different languages (and at different levels of acoustic-linguistic complexity) within the same group of bilingual children.

We found robust acoustic-temporal CTS in both of the languages of bilinguals through two different metrics, namely speech-brain coherence and envelope-level mTRF models. Significant speech-brain coherence was found within the 0.5 to 1.5 Hz delta range, the timescale of prosodic stress occurrences, whose tracking has been shown to be relevant for efficient lexical segmentation.[51] In line with previous research in children,[38] we did not find significant theta-band speech-brain coherence, which has been related to the phase alignment of brain oscillatory activity to syllabic speech units in continuous speech and might be more relevant at later stages of development, such as adulthood.[16,52] It is possible that the developing brain oscillatory activity of children is more sensitive to the slower delta frequency timescales for tracking continuous speech, and only starts exploiting faster syllabic amplitude modulations once their phonological representations are more developed, possibly through sufficient experience and proficiency. This interpretation is in line with Bertels et al.,[53] who found that delta (phrasal) speech-brain coherence was comparable to adults from 5 years of age, while theta matured later between 7.5 and 10 years of age. However, in another recent study, Menn et al.[54] did find significant speech-brain coherence at the syllable rate in infants as young as 9 months of age. Task requirement differences could explain these divergent findings between Menn et al.[54] and our study at the theta frequency band. In Menn and collaborators' study, infants' CTS was measured while they were listening to their mothers in relatively short (40 s) live interactions (audiovisual stimuli) which could have boosted infants' attention during those transient interactions, whereas in our study, CTS metrics were extracted from considerably longer 15-min continuous stories with static images (very limited visual information). It is therefore possible that the length of the tasks and/or the use (or lack) of additional visual cues modulated CTS differently between studies, in line with the attested supportive effect of visual information for theta-band CTS.[55,56]

The lack of significant between-language differences in speech-brain coherence analyses shows that the cortical activity of children in our study robustly phase-aligned to the speech envelope in both their more and less experienced languages. Nonetheless, when it comes to envelope-level TRFs, we found that linguistic experience did shape children's CTS envelope information. Namely, the continuous EEG activity of this group of children encoded the speech envelope of their less experienced language (Spanish) more faithfully than that of their more experienced one (Basque). This was observed despite the fact that, at the time of testing, children had developed similar receptive language abilities in both their languages, illustrated by the lack of relevant between-language differences in phonological abilities and in story comprehension. Our interpretation of such divergent findings between speech-brain coherence and envelope-level TRFs is that each approach maps

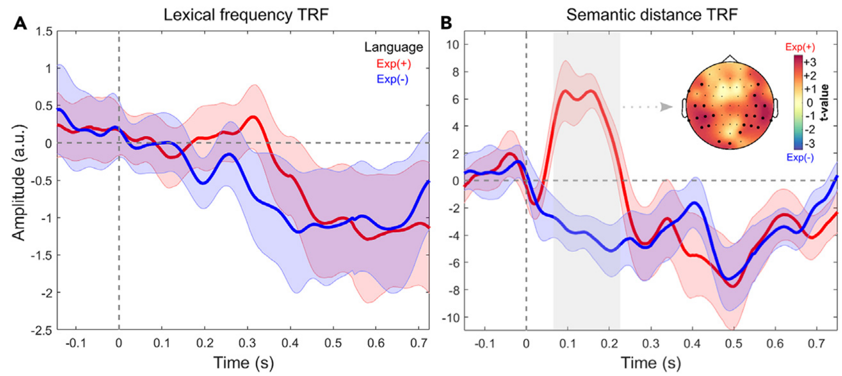


Figure 5. Lexical frequency and semantic distance TRFs

(A) Temporal weights of the EEG signal TRF to lexical frequency in Exp(+) (red) and Exp(−) (blue) across all electrodes. Red and blue shaded areas represent the standard deviation of the temporal weights.

(B) Temporal weights of the EEG signal TRF to semantic distance in Exp(+) (red) and Exp(−) (blue) for the electrodes that showed a significant between-language difference in CBPT. Red and blue shaded areas represent the standard error of the temporal weights. The gray rectangle marks the latency of significant between-languages differences. The distribution over the scalp of such a significant difference in semantic distance (Exp(+) > Exp(−)) in this time window is represented by bigger dots in the top right topographic map.

onto to an extent different aspects of the speech signal. Speech-brain coherence restricts the mapping to consistent phase alignment ("rhythmic" entrainment) between the EEG signal and the speech envelope within narrow frequency bands. Envelope-level TRFs take into account the relationship between amplitude changes in the speech envelope and broadband EEG activity (0.2–15 Hz), and thus could be capturing EEG sensitivity to the acoustic-temporal structure of the speech signal beyond the frequency bands typically used in speech-brain coherence.

These findings help contextualize the role of cortical tracking of the speech envelope for language development. During language acquisition, envelope tracking is tightly linked to phonological development and broader language outcomes.[57–59] Relatedly, envelope CTS by adult second-language learners increases with second language proficiency.[23] Therefore, the fact that the children encoded envelope-level information more faithfully in their less experienced language suggests they relied to a larger extent on low-level acoustic information to achieve a similar level of comprehension as in their more exposed language. We propose that acoustic-temporal CTS could thus serve as a route for speech comprehension during language learning, possibly by supporting the mapping of salient linguistic units in the speech signal while higher-order linguistic knowledge is being learned. Thus, it is likely that as individuals can exploit an increasingly proficient language model for CTS, they gradually shift from relying on low-level temporal information toward tracking higher order linguistic units and structures. Such interpretation is in line with a previous study in adults showing that brain activity shifts from encoding acoustic to linguistic features as a function of increasing language comprehension.[49] Relevantly, the positive correlation between acoustic-temporal and lexico-semantic CTS in the less experienced language in our study suggests that they develop hand-in-hand during language learning, and not that acoustic-temporal CTS compensates for a low lexico-semantic CTS.

Beyond the speech envelope, we report initial evidence on the cortical tracking of abstract linguistic features from continuous speech in children. Such evidence provides a developmental perspective to the role of language knowledge in the cortical tracking of higher order linguistic information, which is being increasingly studied in adults.[19,60] Specifically, we found that children showed an early increased temporal sensitivity to semantic-level information in their more experienced language (Basque). Although we did not statistically test for the presence of peaks at specific latencies, the TRF to semantic distances indicates a negative deflection peaking between 400 and 600 ms in the temporal response to semantic distance that did not differ significantly between languages.[19] Additionally, the temporal response to semantic distance showed a peak between 70 and 230 ms in the more experienced language (Basque), but not in the less experienced one (Spanish). We did not expect the cortical response to the semantic context to show a peak at such early latency. However, it falls within a similar latency range (~50–200 ms) to several adult studies that found early contextual (typically semantic) influence on the cortical tracking of continuous speech.[61–64]

A plausible interpretation of such early sensitivity to semantic information is that, when children have more established language knowledge (only in the more experienced language here), they can capitalize on the semantic context to predict and efficiently track the phonological and/or lexical information from an upcoming word in continuous speech. A study in ~2-year-old bilingual children assessing their event-related potentials (ERPs) in response to lexico-semantic information (known vs. unknown single words in that case) already showed that the latency of such ERPs was earlier in their dominant versus their non-dominant language.[65] In the specific case of our study, two interrelated factors suggest that such difference between languages in the cortical exploitation of semantic information might be driven by lexical, but not phonological, facilitation. First, Basque and Spanish overlap to a great extent in their phonological systems, but not in their vocabularies.[66] Second, the children showed similarly developed phonological abilities in both languages, but significantly greater lexical knowledge in their more experienced language. Therefore, it is plausible that their increased cortical sensitivity to contextual semantic information in their more experienced language is related to a more efficient lexical (rather than phonological) processing of the upcoming word in

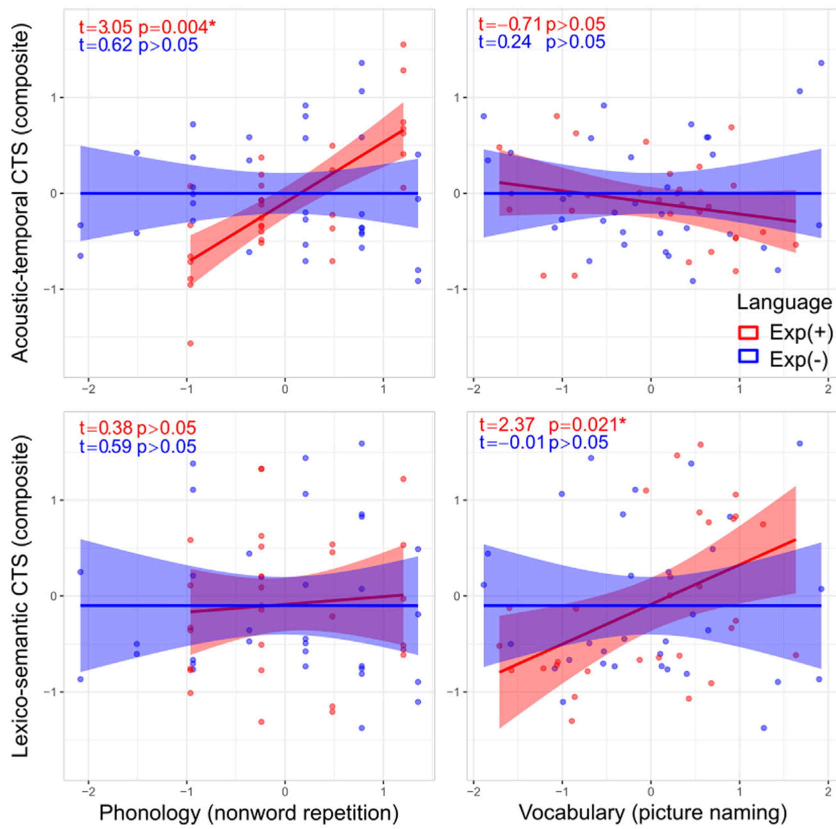


Figure 6. Relationship between CTS (acoustic-temporal and lexico-semantic) and language knowledge (phonology and vocabulary)
Lines represent the slopes of linear regressions between CTS and behavioral metrics in each language respectively, with shaded areas around the lines marking 95% confidence intervals. Estimates in top left corners are the t- and *p* values of linear mixed effect models for each relationship; asterisks indicate a significant effect.

continuous speech. This interpretation is consistent with the positive relationship found between the cortical tracking of lexico-semantic information and vocabulary knowledge.

The relevance of acoustic-temporal and lexico-semantic CTS for explaining behavior was evaluated by assessing their link to individual differences in language knowledge within both phonology and vocabulary domains. Indeed, there was a specific relationship between acoustic-temporal (envelope) CTS and phonological abilities, only in the more experienced language (Basque). This finding is in line with Molinaro et al.,[42] and Di Liberto et al.,[40] who, respectively, reported that the cortical tracking of the envelope and phonetic features were related to phonological abilities in dyslexic children. Moreover, our finding is coherent with the hypothesis that sensitivity to prosodic and syllabic speech amplitude modulations (assessed directly here through speech-brain coherence) subserves the temporal sampling of phonological units within these timescales (phrases, words, and syllables) during phonological development.[32–35] Accordingly, bilingual participants who exhibited higher CTS to such speech amplitude modulations were those who showed more developed phonological processing abilities. Thus, our results support that phonological abilities and acoustic-temporal CTS develop hand in hand,[40,43,59,67,68] with CTS also predicting later reading outcomes.[69]

It is relevant to note that, while tracking speech amplitude modulations contributes to the development of phonological skills,[35,42,70] studies in adults highlight that CTS detaches from faithfully tracking the envelope, at least in normal and effortless listening environments.[71,72] Therefore, the role of acoustic-temporal CTS for language development might decrease as soon as individuals can rely on sufficient lexico-semantic knowledge to track continuous speech.[48,60] A complementary possibility is that, along with language development, delta oscillatory activity shifts gradually from acoustic-temporal CTS (at envelope-level) toward semantic and syntactic CTS. This possibility aligns with a suggested developmental shift when exploiting prosodic information, namely from recognizing word boundaries early in life toward extracting semantic and syntactic information later in childhood.[73–75] Indeed, Kaufeld et al.[60] found that linguistic content modulated delta neural oscillatory activity beyond stimulus-driven prosodic timing. It is therefore possible that, during language development, delta CTS helps map the acoustically driven prosodic timing during language learning, while concurrently allowing the accumulation of knowledge about the linguistic units and structures present at similar acoustic timescales, such as semantic and syntactic information. While our results offer an initial proof of an envelope-to-semantic tradeoff in CTS, future developmental research is needed to attest whether the acoustic-to-syntactic developmental tradeoff observed by previous studies also takes place in CTS. Interestingly, Menn, Ward et al.[37] reported that 10-month-old infants delta (prosodic) CTS predicted vocabulary knowledge at 24 months. It is therefore possible that, in our study, children relied more on extracting envelope-level phonological information in their less familiar language, hence the significantly stronger prediction correlation of envelope-level TRFs in that language; and utilized the same delta tracking to extract increasingly more abstract linguistic information in their more exposed language. Future longitudinal studies assessing the different reliance on acoustic-temporal or linguistic features for tracking the speech signal as a function of age and language knowledge, will help clarify whether and when this potential developmental shift takes place.

Parallel to the acoustic-temporal CTS-phonology relationship, we observed a significant positive relationship between lexico-semantic CTS and vocabulary knowledge, also only in the most experienced language (Basque). Indeed, these results provide initial evidence for the behavioral relevance of the cortical tracking of lexico-semantic information during language development. Previous research in monolingual children showed that ERPs linked to semantic processing were positively correlated with speech comprehension as well as with non-word decoding.[76] In bilingual infants, ERPs to known words showed an earlier latency than for unknown words,[65] possibly reflecting a facilitatory effect of language knowledge on lexico-semantic neurocognitive processing. Such ERP evidence and our CTS findings converge on highlighting the relevance of lexico-semantic neurocognitive processing for explaining phonological, and lexico-semantic behavioral performance in speech comprehension. Interestingly, CTS was linked to both phonological and vocabulary performance only in the more experienced language, suggesting that the more stable language representations and processes in this language might map more consistently and faithfully onto individual differences in CTS. In contrast, weaker knowledge in the less experienced language might not involve as efficiently the neural resources for CTS. Accordingly, the language knowledge of children was overall more variable in their less experienced language (Figure 2). It is also plausible that significant behavioral-CTS relationships in the less exposed language had been detected with a larger sample size, by accounting for the higher intra- and inter-individual variability.

Overall, the present study informs theoretical, empirical, and computational accounts of the relevant role that linguistic experience plays in continuously shaping cortical mechanisms that extract linguistic information from speech. Given our and previous findings, we propose that acoustic-temporal (envelope) CTS serves as phonological route for speech processing while a language is being learnt. Later on, and as a function of maturational factors such as age and linguistic experience, expert neurocognitive language models start exploiting lexical, semantic, and syntactic information to organize linguistic meaning,[77,78] group words within or between phrases,[79] and predict upcoming linguistic information.[25,48]

In summary, our findings shed new light on the role that linguistic experience plays in the CTS, a neurocognitive mechanism supporting speech comprehension. For the first time, we show that, during language development, the maturation of the cortical tracking of the temporal speech envelope and lexico-semantic speech features is tightly linked to the amount of exposure to a language and the subsequent knowledge attained in that language. Moreover, in the less experienced language, we show that acoustic-temporal and lexico-semantic CTS are positively linked. This, together with our findings of higher lexico-semantic CTS in the most experienced language, suggests that, while lexico-semantic CTS needs accumulated experience to develop, both types of speech tracking develop hand in hand, and not that acoustic-temporal compensates for lexico-semantic CTS. We also highlight the behavioral relevance of CTS, by showing that individual differences in specific language abilities are linked to the cortical tracking of specifically related linguistic features in continuous speech.

Limitations of the study

There are two main study limitations that could be helped by future research. First, our single time point study only allows us to observe that acoustic-temporal and lexico-semantic CTS is indeed sensitive to differences in language exposure. However, longitudinal data would help us further understand whether the developmental tradeoff that we point toward is in place. In other words, whether the shift from relying on acoustic-temporal toward lexico-semantic information for tracking continuous speech takes place as a function of accumulated linguistic exposure in addition to other crucial maturational factors for language development, such as chronological age.

Also, a bigger sample size would have allowed us to compensate for the higher variability in behavioral language knowledge in the less experienced language (Figure 2), and test whether the specific links between CTS and behavioral language knowledge take place also in the second language. Additionally, a broader range of exposure to each language would have allowed us to test the correlation between language experience and CTS. A minor caveat of our study, specific to the topographic interpretation of our lexico-semantic results, is that our reference electrode was located on the scalp midline central position (Cz), which could have boosted the signal-to-noise ratio for sensors located apart from such position. And, thus, yield more observable effects toward lateral temporal and parietal areas.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Behavioral session

- Electroencephalography (EEG) session
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Participants
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavioral language measures
  - Speech-brain coherence
  - mTRFs
  - Relationship between CTS and behavioral language measures

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110247.

ACKNOWLEDGMENTS

We would like to thank the participants and their families, as well as the BCBL lab department, for their contribution and commitment over the years of this longitudinal study. Special thanks to Araitz Garnika, Leire Eizagirre, Ioanna Taouki, for their support with stimuli preparation and EEG data collection, to Giovanni Di Liberto for his support with mTRF modeling, and to Caroline Handley for her proofreading and suggestions on this manuscript. This research was supported by the FPI grant BES-2016-078125 (to M.L. and J.P.N.) and by the PID2022-136989OB-I00 (to M.L.) and PID2022-136991NB-I00 (to N.M.) by Ministerio Español de Economía, Industria y Competitividad MINECO and Fondo Social Europeo; through project RTI2018-096242-B-I00(MCIU/AEI/FEDER, UE) funded by Ministerio de Ciencia, Innovación y Universidades (MCIU), the Agencia Estatal de Investigación and Fondo Europeo de Desarrollo Regional; by the Basque Government through the BERC 2018-2021 program and by the Spanish State Research Agency through BCBL Severo Ochoa excellence accreditation SEV-2015-0490.

AUTHOR CONTRIBUTIONS

J.P.-N.: conceptualization, methodology, formal analysis, writing – original draft, writing – review & editing investigation, data curation, visualization, funding acquisition; A.K.-G.: methodology, formal analysis, writing – review & editing, data curation, resources, visualization; M.L.: methodology, formal analysis, writing – review & editing, data curation, resources, visualization; G.P.: data collection, data curation, writing – review & editing; N.M.: conceptualization, writing – original draft, writing – review & editing, supervision; M. L.: conceptualization, writing – original draft, writing – review & editing, supervision.

DECLARATION OF INTERESTS

The authors have declared that no competing financial or nonfinancial interests existed.

Received: November 3, 2023
Revised: March 14, 2024
Accepted: June 7, 2024
Published: June 11, 2024

REFERENCES

1. Gervain, J. (2015). Plasticity in early language acquisition: the effects of prenatal and early childhood experience. Curr. Opin. Neurobiol. *35*, 13–20. https://doi.org/10.1016/j.conb.2015.05.004.
2. Kuhl, P.K. (2004). Early language acquisition: cracking the speech code. Nat. Rev. Neurosci. *5*, 831–843. https://doi.org/10.1038/nrn1533.
3. Skeide, M.A., and Friederici, A.D. (2016). The ontogeny of the cortical language network. Nat. Rev. Neurosci. *17*, 323–332. https://doi.org/10.1038/nrn.2016.23.
4. Carbajal, M.J., and Peperkamp, S. (2020). Dual language input and the impact of language separation on early lexical development. Infancy *25*, 22–45. https://doi.org/10.1111/infa.12315.
5. Gámez, P.B., Griskell, H.L., Sobrevilla, Y.N., and Vazquez, M. (2019). Dual Language and English-Only Learners' Expressive and Receptive Language Skills and Exposure to Peers' Language. Child Dev. *90*, 471–479. https://doi.org/10.1111/cdev.13197.
6. Gathercole, V.C.M., and Thomas, E.M. (2005). Minority language survival: Input factors influencing the acquisition of Welsh. In ISB4 (Cascadilla Press).
7. Paradis, J., Tulpar, Y., and Arppe, A. (2016). Chinese L1 children's English L2 verb morphology over time: individual variation in long-term outcomes. J. Child Lang. *43*, 553–580. https://doi.org/10.1017/S0305000915000562.
8. Pearson, B.Z., Fernandez, S.C., Lewedeg, V., and Oller, D. (1997). The relation of input factors to lexical learning by bilingual infants. Appl. Psycholinguist. *18*, 41–58. https://doi.org/10.1017/S0142716400009863.
9. Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. Int. J. Billing. *15*, 426–445. https://doi.org/10.1177/1367006911403202.
10. Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. *15*, 511–517. https://doi.org/10.1038/nn.3063.
11. Obleser, J., and Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. Trends Cognit. Sci. *23*, 913–926. https://doi.org/10.1016/j.tics.2019.08.004.
12. Ding, N., and Simon, J.Z. (2012). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J. Neurophysiol. *107*, 78–89. https://doi.org/10.1152/jn.00297.2011.
13. Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., and Garrod, S. (2013). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. PLoS Biol. *11*, e1001752. https://doi.org/10.1371/journal.pbio.1001752.
14. Luo, H., and Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably

Discriminate Speech in Human Auditory Cortex. Neuron 54, 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004.
15. Molinaro, N., and Lizarazu, M. (2018). Delta(but not theta)-band cortical entrainment involves speech-specific processing. Eur. J. Neurosci. 48, 2642–2650. https://doi.org/10.1111/ejn.13811.
16. Peelle, J.E., Gross, J., and Davis, M.H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. Cerebr. Cortex 23, 1378–1387. https://doi.org/10.1093/cercor/bhs118.
17. Brodbeck, C., Bhattasali, S., Cruz Heredia, A.A.L., Resnik, P., Simon, J.Z., and Lau, E. (2022). Parallel processing in speech perception with local and global representations of linguistic context. Elife 11, e72056. https://doi.org/10.7554/eLife.72056.
18. Broderick, M.P., Anderson, A.J., Di Liberto, G.M., Crosse, M.J., and Lalor, E.C. (2018). Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech. Curr. Biol. 28, 803–809.e3. https://doi.org/10.1016/j.cub.2018.01.080.
19. Broderick, M.P., Di Liberto, G.M., Anderson, A.J., Rofes, A., and Lalor, E.C. (2021). Dissociable electrophysiological measures of natural language processing reveal differences in speech comprehension strategy in healthy ageing. Sci. Rep. 11, 4963. https://doi.org/10.1038/s41598-021-84597-9.
20. Hurtado, N., Marchman, V.A., and Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. Dev. Sci. 11, F31–F39. https://doi.org/10.1111/j.1467-7687.2008.00768.x.
21. Weisleder, A., and Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. Psychol. Sci. 24, 2143–2152. https://doi.org/10.1177/0956797613488145.
22. Hurtado, N., Grüter, T., Marchman, V.A., and Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. Biling. 17, 189–202. https://doi.org/10.1017/S136672891300014X.
23. Lizarazu, M., Carreiras, M., Bourguignon, M., Zarraga, A., and Molinaro, N. (2021). Language Proficiency Entails Tuning Cortical Activity to Second Language Speech. Cerebr. Cortex 31, 3820–3831. https://doi.org/10.1093/cercor/bhab051.
24. Reetzke, R., Gnanateja, G.N., and Chandrasekaran, B. (2021). Neural tracking of the speech envelope is differentially modulated by attention and language experience. Brain Lang. 213, 104891. https://doi.org/10.1016/j.bandl.2020.104891.
25. ten Oever, S., and Martin, A.E. (2021). An oscillating computational model can track pseudo-rhythmic speech by using linguistic predictions. Elife 10, e68066. https://doi.org/10.7554/eLife.68066.
26. Oller, D.K., and Eilers, R.E. (2002). Language and Literacy in Bilingual Children (Multilingual Matters).
27. Paradis, J., and Jia, R. (2017). Bilingual children's long-term outcomes in English as a second language: language environment factors shape individual differences in catching up with monolinguals. Dev. Sci. 20, e12433. https://doi.org/10.1111/desc.12433.
28. Pérez-Navarro, J., and Lallier, M. (2023). The Contribution of the Amount of Linguistic Exposure to Bilingual Language Development: Longitudinal Evidence from Preschool Years. Preprint at PsyArXiv. https://doi.org/10.31234/osf.io/jwkt5.
29. Nittrouer, S. (1996). The relation between speech perception and phonemic awareness: evidence from low-SES children and children with chronic OM. J. Speech Hear. Res. 39, 1059–1070. https://doi.org/10.1044/jshr.3905.1059.
30. Nittrouer, S., and Burton, L.T. (2005). The role of early language experience in the development of speech perception and phonological processing abilities: evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. J. Commun. Disord. 38, 29–63. https://doi.org/10.1016/j.jcomdis.2004.03.006.
31. Nittrouer, S. (2002). 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. https://doi.org/10.1044/0161-1461(2002/020.
32. Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. Trends Cognit. Sci. 15, 3–10. https://doi.org/10.1016/j.tics.2010.10.001.
33. Goswami, U. (2017). A Neural Basis for Phonological Awareness? An Oscillatory Temporal-Sampling Perspective. Curr. Dir. Psychol. Sci. 27, 56–63. https://doi.org/10.1177/0963721417727520.
34. Goswami, U., and Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? Lab. Phonol. 4, 67–92. https://doi.org/10.1515/lp-2013-0004.
35. Leong, V., and Goswami, U. (2014). Assessment of rhythmic entrainment at multiple timescales in dyslexia: Evidence for disruption to syllable timing. Hear. Res. 308, 141–161. https://doi.org/10.1016/j.heares.2013.07.015.
36. Attaheri, A., Choisdealbha, Á.N., Di Liberto, G.M., Rocha, S., Brusini, P., Mead, N., Olawole-Scott, H., Boutris, P., Gibbon, S., Williams, I., et al. (2022). Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. Neuroimage 247, 118698. https://doi.org/10.1016/j.neuroimage.2021.118698.
37. Menn, K.H., Ward, E.K., Braukmann, R., van den Boomen, C., Buitelaar, J., Hunnius, S., and Snijders, T.M. (2022). Neural Tracking in Infancy Predicts Language Development in Children With and Without Family History of Autism. Neurobiol. Lang. 3, 495–514. https://doi.org/10.1162/nol_a_00074.
38. Ríos-López, P., Molinaro, N., Bourguignon, M., and Lallier, M. (2020). Development of neural oscillatory activity in response to speech in children from 4 to 6 years old. Dev. Sci. 23, e12947. https://doi.org/10.1111/desc.12947.
39. Destoky, F., Bertels, J., Niesen, M., Wens, V., Vander Ghinst, M., Leybaert, J., Lallier, M., Ince, R.A.A., Gross, J., De Tiège, X., and Bourguignon, M. (2020). Cortical tracking of speech in noise accounts for reading strategies in children. PLoS Biol. 18, e3000840. https://doi.org/10.1371/journal.pbio.3000840.
40. Di Liberto, G.M., Peter, V., Kalashnikova, M., Goswami, U., Burnham, D., and Lalor, E.C. (2018). Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. Neuroimage 175, 70–79. https://doi.org/10.1016/j.neuroimage.2018.03.072.
41. Granados Barbero, R., Ghesquière, P., and Wouters, J. (2022). Development of Atypical Reading at Ages 5 to 9 Years and Processing of Speech Envelope Modulations in the Brain. Front. Comput. Neurosci. 16, 894578.
42. Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., and Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. Hum. Brain Mapp. 37, 2767–2783. https://doi.org/10.1002/hbm.23206.
43. Power, A.J., Colling, L.J., Mead, N., Barnes, L., and Goswami, U. (2016). Neural encoding of the speech envelope by children with developmental dyslexia. Brain Lang. 160, 1–10. https://doi.org/10.1016/j.bandl.2016.06.006.
44. Oller, D.K., Pearson, B.Z., and Cobo-Lewis, A.B. (2007). Profile effects in early bilingual language and literacy. Appl. Psycholinguist. 28, 191–230. https://doi.org/10.1017/S0142716407070117.
45. Paradis, J., and Genesee, F. (1996). Syntactic Acquisition in Bilingual Children: Autonomous or Interdependent? Stud. Sec. Lang. Acquis. 18, 1–25. https://doi.org/10.1017/S0272263100014662.
46. Klimovich-Gray, A., Barrena, A., Agirre, E., and Molinaro, N. (2021). One Way or Another: Cortical Language Areas Flexibly Adapt Processing Strategies to Perceptual And Contextual Properties of Speech. Cerebr. Cortex 31, 4092–4103. https://doi.org/10.1093/cercor/bhab071.
47. Koskinen, M., Kurimo, M., Gross, J., Hyvärinen, A., and Hari, R. (2020). Brain activity reflects the predictability of word sequences in listened continuous speech. Neuroimage 219, 116936. https://doi.org/10.1016/j.neuroimage.2020.116936.
48. Molinaro, N., Lizarazu, M., Baldin, V., Pérez-Navarro, J., Lallier, M., and Ríos-López, P. (2021). Speech-brain phase coupling is enhanced in low contextual semantic predictability conditions. Neuropsychologia 156, 107830. https://doi.org/10.1016/j.neuropsychologia.2021.107830.
49. Tezcan, F., Weissbart, H., and Martin, A.E. (2023). A tradeoff between acoustic and linguistic feature encoding in spoken language comprehension. Elife 12, e82386. https://doi.org/10.7554/eLife.82386.
50. Brodbeck, C., Das, P., Gillis, M., Kulasingham, J.P., Bhattasali, S., Gaston, P., Resnik, P., and Simon, J.Z. (2023). Eelbrain, a Python toolkit for time-continuous analysis with temporal response functions. Elife 12, e85012. https://doi.org/10.7554/eLife.85012.
51. Kooijman, V., Hagoort, P., and Cutler, A. (2009). Prosodic Structure in Early Word Segmentation: ERP Evidence From Dutch Ten-Month-Olds. Infancy 14, 591–612. https://doi.org/10.1080/15250000903263957.
52. Doelling, K.B., Arnal, L.H., Ghitza, O., and Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. Neuroimage 85, 761–768. https://doi.org/10.1016/j.neuroimage.2013.06.035.

53. Bertels, J., Niesen, M., Destoky, F., Coolen, T., Vander Ghinst, M., Wens, V., Rovai, A., Trotta, N., Baart, M., Molinaro, N., et al. (2023). Neurodevelopmental oscillatory basis of speech processing in noise. Dev. Cogn. Neurosci. *59*, 101181. https://doi.org/10.1016/j.dcn.2022.101181.
54. Menn, K.H., Michel, C., Meyer, L., Hoehl, S., and Männel, C. (2022). Natural infant-directed speech facilitates neural tracking of prosody. Neuroimage *251*, 118991. https://doi.org/10.1016/j.neuroimage.2022.118991.
55. Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. Science *320*, 110–113. https://doi.org/10.1126/science.1154735.
56. Power, A.J., Mead, N., Barnes, L., and Goswami, U. (2013). Neural entrainment to rhythmic speech in children with developmental dyslexia. Front. Hum. Neurosci. *7*, 777. https://doi.org/10.3389/fnhum.2013.00777.
57. Goswami, U., Wang, H.-L.S., Cruz, A., Fosker, T., Mead, N., and Huss, M. (2011). Language-universal Sensory Deficits in Developmental Dyslexia: English, Spanish, and Chinese. J. Cognit. Neurosci. *23*, 325–337. https://doi.org/10.1162/jocn.2010.21453.
58. Kalashnikova, M., Goswami, U., and Burnham, D. (2019). Sensitivity to amplitude envelope rise time in infancy and vocabulary development at three years: A significant relationship. Dev. Sci. *22*, e12836. https://doi.org/10.1111/desc.12836.
59. Vanvooren, S., Poelmans, H., De Vos, A., Ghesquière, P., and Wouters, J. (2017). Do prereaders' auditory processing and speech perception predict later literacy? Res. Dev. Disabil. *70*, 138–151. https://doi.org/10.1016/j.ridd.2017.09.005.
60. Kaufeld, G., Bosker, H.R., Ten Oever, S., Alday, P.M., Meyer, A.S., and Martin, A.E. (2020). Linguistic structure and meaning organize neural oscillations into a content-specific hierarchy. J. Neurosci. *40*, 9467–9475. https://doi.org/10.1523/JNEUROSCI.0302-20.2020.
61. Brodbeck, C., Presacco, A., and Simon, J.Z. (2018). Neural source dynamics of brain responses to continuous stimuli: Speech processing from acoustics to comprehension. Neuroimage *172*, 162–174. https://doi.org/10.1016/j.neuroimage.2018.01.042.
62. Brodbeck, C., Kandylaki, K.D., and Scharenborg, O. (2024). Neural representations of non-native speech reflect proficiency and interference from native language knowledge. J. Neurosci. *44*. https://doi.org/10.1523/JNEUROSCI.0666-23.2023.
63. Broderick, M.P., Anderson, A.J., and Lalor, E.C. (2019). Semantic Context Enhances the Early Auditory Encoding of Natural Speech. J. Neurosci. *39*, 7564–7575. https://doi.org/10.1523/JNEUROSCI.0584-19.2019.
64. Donhauser, P.W., and Baillet, S. (2020). Two Distinct Neural Timescales for Predictive Speech Processing. Neuron *105*, 385–393.e9. https://doi.org/10.1016/j.neuron.2019.10.019.
65. Conboy, B.T., and Mills, D.L. (2006). Two languages, one developing brain: event-related potentials to words in bilingual toddlers. Dev. Sci. *9*, F1–F12. https://doi.org/10.1111/j.1467-7687.2005.00453.x.
66. Ezeizabarrena, M.-J., and Alegria, A. (2015). Early coda production in bilingual Spanish and Basque. In The Acquisition of Spanish in Understudied Language Pairings, J. Tiffany and P. Silvia, eds. (John Benjamins Publishing Company).
67. Boets, B., Wouters, J., van Wieringen, A., and Ghesquière, P. (2007). Auditory processing, speech perception and phonological ability in pre-school children at high-risk for dyslexia: A longitudinal study of the auditory temporal processing theory. Neuropsychologia *45*, 1608–1620. https://doi.org/10.1016/j.neuropsychologia.2007.01.009.
68. Lundberg, I., Frost, J., and Petersen, O.-P. (1988). Effects of an Extensive Program for Stimulating Phonological Awareness in Preschool Children. Read. Res. Q. *23*, 263–284.
69. Ríos-López, P., Molinaro, N., Bourguignon, M., and Lallier, M. (2021). Right-hemisphere coherence to speech at pre-reading stages predicts reading performance one year later. J. Cognit. Psychol. *0*, 1–15. https://doi.org/10.1080/20445911.2021.1986514.
70. Leong, V., and Goswami, U. (2017). Difficulties in auditory organization as a cause of reading backwardness? An auditory neuroscience perspective. Dev. Sci. *20*, e12457. https://doi.org/10.1111/desc.12457.
71. Hauswald, A., Keitel, A., Chen, Y.-P., Rösch, S., and Weisz, N. (2022). Degradation levels of continuous speech affect neural speech tracking and alpha power differently. Eur. J. Neurosci. *55*, 3288–3302. https://doi.org/10.1111/ejn.14912.
72. Schmidt, F., Chen, Y.-P., Keitel, A., Rösch, S., Hannemann, R., Serman, M., Hauswald, A., and Weisz, N. (2023). Neural speech tracking shifts from the syllabic to the modulation rate of speech as intelligibility decreases. Psychophysiology *60*, e14362. https://doi.org/10.1111/psyp.14362.
73. Friederici, A.D., and Männel, C. (2014). Neural correlates of the development of speech perception and comprehension. In The Oxford Handbook of Cognitive Neuroscience (Oxford University Press), pp. 171–192.
74. Guasti, M.T. (2002). Language Acquisition: The Growth of Grammar (The MIT Press).
75. Männel, C., Schipke, C.S., and Friederici, A.D. (2013). The role of pause as a prosodic boundary marker: Language ERP studies in German 3- and 6-year-olds. Dev. Cogn. Neurosci. *5*, 86–94. https://doi.org/10.1016/j.dcn.2013.01.003.
76. Henderson, L.M., Baseler, H.A., Clarke, P.J., Watson, S., and Snowling, M.J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary and decoding. Brain Lang. *117*, 88–99. https://doi.org/10.1016/j.bandl.2010.12.003.
77. Martin, A.E. (2020). A Compositional Neural Architecture for Language. J. Cognit. Neurosci. *32*, 1407–1427. https://doi.org/10.1162/jocn_a_01552.
78. Martin, A.E., and Doumas, L.A.A. (2017). A mechanism for the cortical computation of hierarchical linguistic structure. PLoS Biol. *15*, e2000663. https://doi.org/10.1371/journal.pbio.2000663.
79. Meyer, L., Henry, M.J., Gaston, P., Schmuck, N., and Friederici, A.D. (2017). Linguistic Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. Cerebr. Cortex *27*, 4293–4302. https://doi.org/10.1093/cercor/bhw228.
80. van Heuven, W.J.B., Mandera, P., Keuleers, E., and Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English: Q. J. Exp. Psychol.
81. Acha, J., Laka, I., Landa, J., and Salaburu, P. (2014). EHME: a New Word Database for Research in Basque Language. Spanish J. Psychol. *17*, E79. https://doi.org/10.1017/sjp.2014.79.
82. Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., and Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. Behav. Res. Methods *45*, 1246–1258. https://doi.org/10.3758/s13428-013-0326-1.
83. Thordardottir, E., Rothenberg, A., Rivard, M.-E., and Naves, R. (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? J. Multiling. Commun. Disord. *4*, 1–21. https://doi.org/10.1080/14769670500215647.
84. Nation, K., Dawson, N.J., and Hsiao, Y. (2022). Book Language and Its Implications for Children's Language, Literacy, and Development. Curr. Dir. Psychol. Sci. *31*, 375–380. https://doi.org/10.1177/09637214221103264.
85. Destoky, F., Philippe, M., Bertels, J., Verhasselt, M., Coquelet, N., Vander Ghinst, M., Wens, V., De Tiège, X., and Bourguignon, M. (2019). Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope. Neuroimage *184*, 201–213. https://doi.org/10.1016/j.neuroimage.2018.09.006.
86. MathWorks (2014). MATLAB. Version 8.4.0.150421 (R2014b) (The MathWorks Inc).
87. Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Comput. Intell. Neurosci. *2011*, 156869. https://doi.org/10.1155/2011/156869.
88. Bortel, R., and Sovka, P. (2007). Approximation of statistical distribution of magnitude squared coherence estimated with segment overlapping. Signal Process. *87*, 1100–1117. https://doi.org/10.1016/j.sigpro.2006.10.003.
89. Halliday, D.M., Rosenberg, J.R., Amjad, A.M., Breeze, P., Conway, B.A., and Farmer, S.F. (1995). A framework for the analysis of mixed time series/point process data–theory and application to the study of physiological tremor, single motor unit discharges and electromyograms. Prog. Biophys. Mol. Biol. *64*, 237–278. https://doi.org/10.1016/s0079-6107(96)00009-0.
90. Crosse, M.J., Di Liberto, G.M., Bednar, A., and Lalor, E.C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. Front. Hum. Neurosci. *10*, 604. https://doi.org/10.3389/fnhum.2016.00604.
91. de Jong, N.H., and Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. Behav. Res. Methods *41*, 385–390. https://doi.org/10.3758/BRM.41.2.385.
92. Boersma, P., and Weenink, D. (2021). Praat: Doing Phonetics by Computer. Version 6.1.46. http://www.praat.org/.

93. Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., and Mikolov, T. (2016). FastText.zip: Compressing text classification models. Preprint at arXiv. https://doi.org/10.48550/arXiv.1612.03651.

94. Crosse, M.J., Zuk, N.J., Di Liberto, G.M., Nidiffer, A.R., Molholm, S., and Lalor, E.C. (2021). Linear Modeling of Neurophysiological Responses to Speech and Other Continuous Stimuli: Methodological Considerations for Applied Research. Front. Neurosci. *15*, 705621. https://doi.org/10.3389/fnins.2021.705621.

95. Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. Nat. Rev. Neurosci. *16*, 234–244. https://doi.org/10.1038/nrn3924.

96. Goswami, U., and Bryant, P. (1990). Phonological Skills and Learning to Read (Lawrence Erlbaum Associates, Inc)), pp. 1–26.

97. Morais, J., Cary, L., Alegria, J., and Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? Cognition *7*, 323–331. https://doi.org/10.1016/0010-0277(79)90020-9.

98. Douglas Bates, M.M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Software *67*, 1–48. https://doi.org/10.18637/jss.v067.i01.

99. R Core Team (2022). R: A Language and Environment for Statistical Computing.

100. JASP team (2022). JASP. Version 0.16.4.

101. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods *164*, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
MATLAB_R2014B	MathWorks	https://es.mathworks.com/products/matlab.html
R (version 4.2.1)	R Core Team	R Core Team. (2020). R: A Language and Environment for Statistical Computing [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/
Deposited data		
Project's repository	Open Science Framework	https://osf.io/bdtaq/?view_only=a3768484b1e7471fb7fbcdfc459c774f

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Jose Pérez-Navarro (jperez@bcbl.eu).

Materials availability

This study did not generate new reagents.

Data and code availability

- Data and materials (stimuli and their derivate regressors) necessary to reproduce the analyses presented here are publicly accessible at the Open Science Framework repository of the project upon publication.
- All original code necessary for reproducing the analyses is available at the repository.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Behavioral session

The relative amount of linguistic experience within each language was assessed through a comprehensive linguistic questionnaire completed by the children's parents. From this questionnaire, we extracted a composite index of the proportional amount of experience within each language (Formulas S1–S3). In addition to the amount of language experience, we assessed children's phonological and lexico-semantic language knowledge.

Phonological abilities were evaluated with a nonword repetition task consisting of 18 items for each language, ranging from 3- to 6-syllable nonwords (6 nonwords per syllable length and language). In this task, participants were randomly presented with the 36 auditory nonwords, each constructed to follow either the stress and syllabic phonotactic constraints of Basque or those of Spanish. Upon participant's oral repetition of each nonword, an experimenter coded each repetition as either correct or incorrect. Mean accuracy for each language was taken as a measure of language-specific phonological abilities.

Lexico-semantic knowledge was assessed through a picture-naming task consisting of 45 items for each language. The items were different in each language version, although matched in terms of word frequency. In each language version of the task, participants were asked to name the pictures that appeared on screen in a random order, without any time constraints. The experimenter coded each trial as either correct or incorrect upon participant's response. The order of presentation of the two languages was counterbalanced across participants. In order to harmonize participants' performance in the picture-naming task across both languages, we weighted performance on the different items (1, correct; 0, incorrect) by the inverse of their Zipf lexical frequency.[80] For Basque words, lexical frequency was extracted from EHME database[81]; and for Spanish words from EsPal database.[82] This way, more frequent (and presumably easier) words had a smaller weight on participants' overall performance than less frequent ones that were less likely to be known by children. Thus, our measure of performance in the picture-naming task was Zipf-frequency-weighted accuracy.

Importantly, our phonological and lexico-semantic measures have previously proven sensitive to differential amounts of language exposure in bilinguals.[5,7,83]

Electroencephalography (EEG) session

EEG task: speech listening

Participants listened to continuous streams of natural speech in the form of storytelling. We used two stories that were adaptations of two short books targeted at 6-year-old children, and followed a very similar narrative structure. Importantly, this literary style targeted at children offers a great number of different words and, consequently, considerable variability in their lexical frequencies (Nation et al.,[84] 2022; see Figure S1), which was relevant to explore the effect of language experience and knowledge on the cortical tracking of lexico-semantic information.

The stories were about the history of outer space exploration and the evolution of life on Earth. For each story, both a Exp(+) (Basque) and a Exp(−) (Spanish) version was narrated by the same female native speaker of both languages in a child-directed speech register. Each story lasted about 15 min, because this duration has proven sufficient to robustly estimate cortical tracking of speech with EEG.[85] Half of the participants listened to the 'outer space' story in Basque, and the 'life on Earth' story in Spanish, and vice versa for the other half of participants. The order of language presentation was counterbalanced across participants. Participants were asked to listen attentively to the stories that were presented to them over speakers. They were sitting in a comfortable upright position and asked to look at static images depicting the story narrative, presented on the center of a computer screen positioned at 80 cm from their eyes. Every 5 min approximately, participants were asked three simple yes/no questions (9 per story in total) to check whether they were paying attention to, and comprehending, the stories.

EEG preprocessing

EEG data were recorded using a 64 Ag-AgCl electrodes standard setting (actiCAP, Brain Products GmbH, Germany). One electrode was placed over the outer canthus of each eye, and one below the left eye to monitor eye movements and blinks. Electrode impedance was always below 15 kΩ and remained below 10 kΩ in the vast majority of electrodes across participants. During data collection, raw EEG signal was amplified (BrainAmp DC, Brain Products GmbH, Germany), online high-pass filtered at 0.05 Hz, digitized using a sampling rate of 1000 Hz, and referenced to the midline central electrode (Cz). Such reference was also kept offline in order to amplify the signal-to-noise ratio in temporal areas, which have shown to be the most sensitive to acoustic information in speech-brain coherence analyses (e.g., Ríos-López et al.,[38] 2020; Schmidt et al.,[72] 2023).

To obtain the best possible temporal alignment between acoustic stimuli and EEG signal, we digitized the speakers' acoustic signal, with a sampling rate of 1000 Hz (Polybox, Brain Products GmbH, Germany). This allowed us to compensate for varying lags between the digitized trigger and the actual presentation of acoustic stimuli. We then cross-correlated the amplitude values of the speakers' acoustic signal and its corresponding audio file to ensure an optimal alignment. Thus, before EEG data preprocessing, triggers that marked the onset of each speech fragment were realigned to the time of maximum correlation with the actual presentation of the acoustic signal.

All EEG data preprocessing steps, and later data transformations were conducted at sensor level in MATLAB (version R2014B, MathWorks,[86] 2014), using both custom code and functions from FieldTrip toolbox (version 20180604, Oostenveld et al.,[87] 2011). First, we downsampled EEG and audio signals to 200 Hz. Second, we bandpass filtered the signal between 0.2 and 40 Hz with a zero-phase fourth-order Butterworth infinite impulse response filter, using the default transition bandwidth in the FieldTrip toolbox for bandpass filtering. Third, we detected physiological artifacts through independent component analysis (ICA, runica method) on the filtered signal. After visual inspection of ICA, we subtracted independent components related to eye movements and blinks from the EEG signal (*mean number of rejected components per participant* = 2.06, *SD* = 0.5). Interpolation of bad channels was achieved using the weighted average of their neighbors. Our EEG data exclusion criteria was to not further analyze participant datasets for which more than 30% of the data were rejected. Although no participant exceeded such a threshold, two participants were excluded, as one of them did not want to remain seated while listening to the stories and the other fell asleep during the recording. Thus, we ended up with a sample of 33 analyzable EEG and behavioral datasets in each language.

For speech-brain coherence analyses, we divided the continuous EEG signal from each storytelling and language into 2000-ms epochs (the inverse of our lowest frequency of interest as well as frequency resolution, 0.5 Hz) with a temporal overlap of 1000 ms. Such window and overlap were chosen to balance between frequency resolution and signal-to-noise ratio for computing coherence spectra and optimize the calculation of speech-brain coherence, as demonstrated by previous studies.[23,88] We discarded epochs and channels in which overall voltage departed more than 3 z-values from the average of all epochs and channels respectively (*mean percentage of epochs removed per participant* = 1.14%, *SD* = 0.66; *mean number of channels removed per participant* = 1 out of 64, *SD* = 1.93). For mTRF analyses, we further low-pass filtered the continuous EEG signal (without epoching) to 15 Hz with a zero-phase fourth-order Butterworth infinite impulse response filter.

CTS indexes

We used two types of CTS metrics (Figure 1). The first one was coherence,[89] which measures the phase correlation between two signals (here the brain and the speech signals, thus termed speech-brain coherence hereafter), which was used to extract the brain tracking of acoustic temporal speech information (i.e., the speech envelope). The second set of CTS metrics were extracted from the multivariate temporal response function (mTRF, Crosse et al.,[90] 2016) of continuous brain activity to acoustic temporal and linguistic (i.e., lexico-semantic) information. mTRFs consist of the linear mapping of the values of continuous stimuli (acting as regressors) on continuous brain activity (acting as response, EEG in our case).

Speech-brain coherence

We computed speech-brain coherence as the phase correlation between EEG and speech envelope (i.e., the Hilbert-transformed audio signal). To achieve this, we used custom functions in MATLAB following the specifications described in Molinaro and Lizarazu.[15] We circumscribed our speech-brain coherence analysis to the 0.5–10 Hz frequency range, which spans over the delta (0.5–4 Hz) and theta (4–7 Hz) frequency bands, which align to the timescales of prosodic phrasing (∼1000–2000 ms) and syllables (∼150–200 ms) respectively. In order to test whether speech-brain coherence within theta band was specifically related to syllable tracking, we estimated the syllable rate of the stories (based on an automatic algorithm, (de Jong & Wempe,[91] 2009; in Praat, Boersma & Weenink,[92] 2021). The overall average syllable rate was 5.63 Hz (*SD* = 0.33) and was highly similar in Basque (5.7 Hz, *SD* = 0.28) and Spanish (5.56 Hz, *SD* = 0.36). Given that our speech-brain coherence analysis had a frequency resolution of 0.5 Hz, the frequency bin that aligned most closely to the syllable rate was 5.5 Hz.

Coherence values vary between 0 (no linear phase relation) and 1 (total linear phase relation). To find the moment of maximum speech-brain synchronicity, we computed coherence between both signals at 6 different time lags (ranging from 40 to 140 ms in steps of 20 ms) in two arrays of sensors that have previously shown speech-brain coherence effects, located symmetrically within the left (i.e., T7, C3, TP7, and CP3) and right (i.e., T8, C4, TP8, CP4) temporal hemispheres. This *a priori* selection of sensors was only used to determine the time of maximum coherence, and not for localizing statistical coherence effects in our analyses (which were located through cluster-based permutation tests). A 60 ms lag of the EEG with respect to the speech signal was the time point of maximum coherence across participants and conditions, to which we circumscribed our speech-brain coherence estimates for statistical analyses.

mTRFs

We computed one mTRF encoding model per participant and language, which included four regressors: word onsets, the speech envelope, lexical frequency, and sentence-level semantic distance. Word onsets were employed as a control regressor, to account for the presumably high variability in the mTRF models associated with the mere responses to word onsets. Speech envelope served as a regressor for acoustic-temporal CTS, and lexical frequency and semantic distance for lexico-semantic CTS. Importantly, computing envelope-level mTRF also served to verify that it was related to the envelope tracking measured with coherence and to establish a positive control for mTRF analyses of lexico-semantic features, which are less salient in the signal. The speech envelope regressors have the same operationalization (i.e., the Hilbert-transformed audio signal) as in the speech-brain coherence analyses (also following the specifications in Molinaro & Lizarazu,[15] 2018).

Lexical and semantic regressors, as well as the word onset control, consisted of discrete vectors (one per language and story) containing bursts at the onset of every content word in each story. The regressor containing bursts of the same magnitude at word order had the purpose of accounting for EEG responses to the mere start of words, which is temporally correlated with lexico-semantic temporal responses to words but not of higher-order relevance. In the lexico-semantic regressors, the amplitude of such bursts corresponded to latent variables that were used as proxies for either lexical or semantic information. For the lexical regressor, we used lexical frequency: the amplitude of each burst was the inverse of the Zipf lexical frequency of its corresponding word (van Heuven et al.,[80] 2014; see Figure S1). For the semantic regressor, we computed sentence-level semantic distance vectors following a similar approach to Broderick et al.[19] (2021; see Figure S2). We first obtained the semantic representation of each story word in a 300-dimensional space through the fasttext Python package for text representations.[93] Then, the amplitude of each burst was computed as 1 minus the Pearson correlation of the semantic dimensions of a word with the average of all its preceding words within a sentence. This way, lexical items with a bigger semantic correlation with the preceding words within a sentence had lower amplitude bursts, and less semantically related words had higher amplitude bursts relative to their preceding context.

After obtaining all the regressors of interest, we used the mTRF toolbox[90] in MATLAB to fit one mTRF encoding model with the EEG response with all features (word onsets, envelope, lexical frequency, and sentence-level semantic distance), and for each language and participant. Our mTRF analysis time window was 900 ms, spanning from −150 ms pre-to 750 ms post-stimulus. In order to train and test the encoding model, we split our continuous EEG signal and each corresponding feature vector (∼16 min) into 8-folds of equal length (∼2 min). We trained the encoding model in 7-folds and tested its accuracy on the remaining one, repeating this process for 30 iterations per fold. We then extracted two mTRF metrics for between-languages comparisons of CTS. First, we assessed whether and to what extent each individual regressor contributed significantly to the model fit (through *prediction correlation*, r-values). Thus, within 30 iterations per fold for each regressor of interest (envelope, lexical frequency, and semantic distance), we contrasted the full model fit (all regressors) versus the fit of the model where one regressor was permuted, and averaged such r-values across iterations. The resulting fold-average r-values, the above-chance prediction correlation coefficient between each feature, and the mTRF (at electrode level) were used in further statistical analyses as proxies for the extent to which each feature was linearly mapped by EEG activity and to confirm that this feature contributed to the model-fit above the chance level across all subjects (see Figure S3). For the main analysis, the prediction correlation coefficient (the difference between the full model R value and the specific feature-permuted model R value) served as an electrode-level estimate of how faithfully EEG activity mapped a given speech feature (e.g., lexical information).

The second mTRF metric was the *temporal weights* of the EEG response to each regressor. Such weights form an ERP-like response that indexes EEG temporal sensitivity to changes in a given regress or see.[94] Thus, the temporal weights allowed us to test whether there were relevant between-languages differences in the amplitude or the latency of the temporal response to the envelope, lexical, or semantic information.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Participants

Participants were 35 (18 females) Basque-Spanish bilingual children between 6 and 7 years of age (*Mean age* = 6.92; *SD* = 0.11). We selected them from a broader longitudinal study (Pérez-Navarro & Lallier, under review) based on their amount of exposure to each language when they were 6.4 years old (the last stage of that longitudinal study). Namely, they were the group of children who had the most experience within their most exposed language (>70% of their waking hours exposed to Basque, referred to as *Exp(+)* henceforth) and the least within their least exposed language (<30% of their waking hours exposed to Spanish, *Exp(−)* henceforth). This selection criterion enabled us to investigate the influence of differential linguistic experience on CTS of two different languages within the same participants, thus limiting inter-subject variability for our language comparisons. Importantly, since participants had started formal reading instruction only a few months prior to participation in the study, the amount of written input was not sufficient to be considered a main source of language input at the time of testing. Thus, the potential influence of written language exposure on phonological abilities[95–97] and CTS in our study was considered to be minimal.

Participants had normal hearing, no history of neurological disorders, nor familial risk of developmental language disorder or any other cognitive-related genetic pathology. The study was approved by the BCBL Ethics Committee and complied with the Declaration of Helsinki.

QUANTIFICATION AND STATISTICAL ANALYSIS

Behavioral language measures

In order to assess between-languages differences in language performance, we fitted linear mixed effect (LME) models with language (Exp(+), Exp(−)) as predictor for each behavioral language measure (language experience, vocabulary knowledge, and phonological ability), and participants as random intercepts, to account for between-participant differences. LME models were fitted using the lmer function from lme4 package[98] in R (version 4.2.1, R Core Team,[99] 2022). In addition, we used Bayesian t-tests in JASP (version 0.16.4, JASP team,[100] 2022) to test the strength of the evidence for between-languages similarities in our behavioral measures in the case of non-significant between-languages differences.

Speech-brain coherence

We used cluster-based permutation tests (CBPTs) to statistically estimate speech-brain coherence in our 0.5–10 Hz frequency band of interest. CBPT is an efficient way of estimating the presence of a statistical effect in a high dimensional space, as it is able to account for the spatial adjacency of electrodes and test for significant effects that are shared across a group of electrodes (a cluster).[101] In our case, we ran dependent-samples two-tailed CBPTs in FieldTrip with 1000 permutations. Our CBPTs required at least 2 electrodes to form a cluster as a way to limit the possibility of single-electrode false-alarm effects, and we corrected for multiple comparisons based on the number of *a-priori* significant clusters. We first assessed whether there was above-chance speech-brain coherence in each language by contrasting through CBPT the phase alignment of EEG and genuine speech envelope versus the phase alignment of EEG and a surrogate version of the envelope that did not follow the original speech order (i.e., flipped speech envelope surrogate). Then, we contrasted through CBPT the speech-brain coherence values of Exp(+) and Exp(−) to analyze whether there was a significant between-languages effect on this CTS metric.

mTRFs

Before our analyses of interest, we conducted a control test to verify that, at the group level, there were significant prediction correlations between our three regressors of interest and the mTRF model. Namely, we conducted a one-sample t-test (one-tailed) against the null hypothesis (no different prediction correlation from 0) for each of the regressors. After such control, and similarly to when contrasting speech-brain coherence between languages, we compared Exp(+) and Exp(−) through 6 CBPTs, one per mTRF metric (i.e., prediction correlation and temporal weights) and regressor (i.e., envelope, lexical frequency, and semantic distance). Both prediction correlation and temporal weights were analyzed at electrode level.

Relationship between CTS and behavioral language measures

Acoustic-temporal CTS and lexico-semantic CTS composite scores were computed by aggregating EEG metrics for a more robust estimate of CTS-behavior relationships, while limiting the possibility of inflating type I error due to multiple tests. More specifically, an acoustic-temporal CTS composite was composed by averaging the normalized values of: 1) speech-brain coherence at delta (0.5–1.5 Hz) of the electrodes that showed significant delta coherence, 2) speech-brain coherence at theta (4–7 Hz) of the same group of electrodes, and 3) the average prediction correlation values for envelope TRF across all electrodes. The lexico-semantic CTS composite consisted of the average across all electrodes of: 1) prediction correlation values for lexical frequency TRF, and 2) prediction correlation values for semantic distance TRF. These composite scores were validated by the strong positive correlations found between (i) coherence to the speech envelope at delta and theta, and the envelope TRF (*Pearson's r coefficients* in Exp(+) delta-theta = 0.68; delta-envelope TRF = 0.48; theta-envelope TRF = 0.63; Exp(−) delta-theta = 0.68; delta-envelope TRF = 0.55; theta-envelope TRF = 0.44; all *p-values* < 0.01) and (ii) lexical frequency and semantic distance mTRFs (Exp(+) = 0.81; Exp(−) = 0.78; all *p-values* < 0.01).

In order to assess the correspondence between CTS and behavioral language performance in the phonological and lexico-semantic domains, we fitted two LME models, one with acoustic-temporal CTS and one with lexico-semantic CTS as outcome measure. In the first LME model, phonological abilities and vocabulary knowledge, as well as their interaction with language, served as predictors of acoustic-temporal CTS, the composite score of speech-brain coherence and envelope-level TRFs. In the second LME model, the same set of measures were employed as predictors of lexico-semantic CTS, the composite score of lexical- and semantic-level TRFs. We included participants as random intercepts in both LME models, which allowed us to account for baseline individual differences in performance in the behavioral language measures in each of the languages. As for our behavioral language analyses, we used the lmer function from lme4 package in R. In order to test for omnibus main effects and interactions of the predictors, we used the anova function in base R.