

Novel ways of identifying children at risk for language disorders

Being able to use language efficiently is one of the most sophisticated human achievements. It is vital for cognitive development, social inclusion, and education, and talking to each other underpins large parts of our cultural experience.

Yet language problems are common in children in all languages. Estimates are ~7% incidence of significant spoken language problems (Developmental Language Disorder, DLD) and a further ~7% incidence for written language problems (Developmental Dyslexia, DD). Only half of 'late talkers', toddlers who are not speaking by age 2, will go on to develop a language disorder. Currently, there is no way to identify which late talkers are the children at risk. This makes it difficult to deliver targeted early intervention, leading to significant developmental difficulties. DLD and DD also require different interventions, yet share many overlapping behavioural features. Children with DLD or DD thus lose out on many levels.

The Disordered-Speech project capitalises on one significant behavioural overlap between children with DLD or DD: difficulties in hearing speech rhythm patterns (linguistic prosody). By using novel acoustic measures of how preschool children say familiar words aloud, the project hopes to identify key markers of atypical speech. These markers could indicate who is at risk for which disorder across European languages.

The Disordered-Speech project is based on Temporal Sampling theory (Goswami, 2011). This is a sensory-neural theory that builds on how the brain processes speech rhythm patterns. English-speaking children with either DLD or DD have difficulty in discriminating acoustic cues to speech rhythm patterns, particularly the rates-of-change in loudness (amplitude 'rise times') which mark the beginning of syllables. A strong or stressed syllable will have a more salient amplitude rise time.

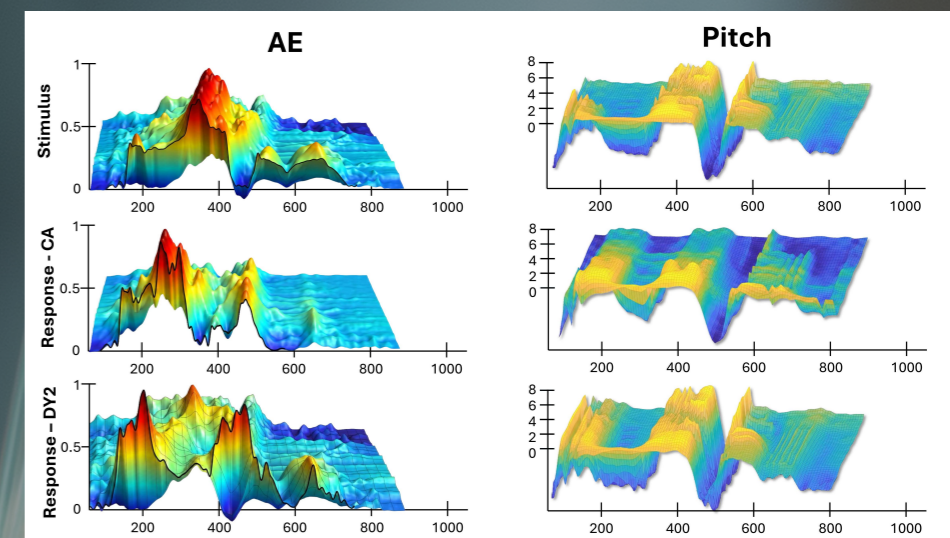
This rise time difficulty is mirrored by difficulties in hearing syllable stress patterns in words. Children with DLD or DD have difficulty in deciding whether a word like 'zebra' has a louder first syllable or a louder second syllable ('zebra' has a louder first syllable, as in most English bisyllabic nouns). Amplitude rise times also serve a neural purpose. Networks of cells in the auditory cortex with preferred rhythmic (oscillation) rates automatically use differences in amplitude rise times

to align their signalling to different parts of the speech signal. We can think of the brain waves 'surfing' the sound waves.

More technically, when we speak, we create sound pressure waves that move energy through the air. The overall 'amplitude envelope' reaching the ear contains a range of energy patterns, as our articulators (larynx, tongue, jaw, lips) modulate the amplitude (loudness) that we produce at different temporal rates or different speeds. The brain

aligns its own intrinsic cellular rhythms to these different temporal rates, phase-resetting its ongoing oscillatory activity using amplitude rise times as a trigger mechanism. Speech-brain alignment is an automatic aspect of how we listen. Temporal Sampling theory proposes that this automatic neural alignment process is inaccurate for children with both DLD and DD at lower frequencies (slower temporal rates). This inaccurate alignment leads to specific difficulties with speech prosody.

Aladdin



Alligator

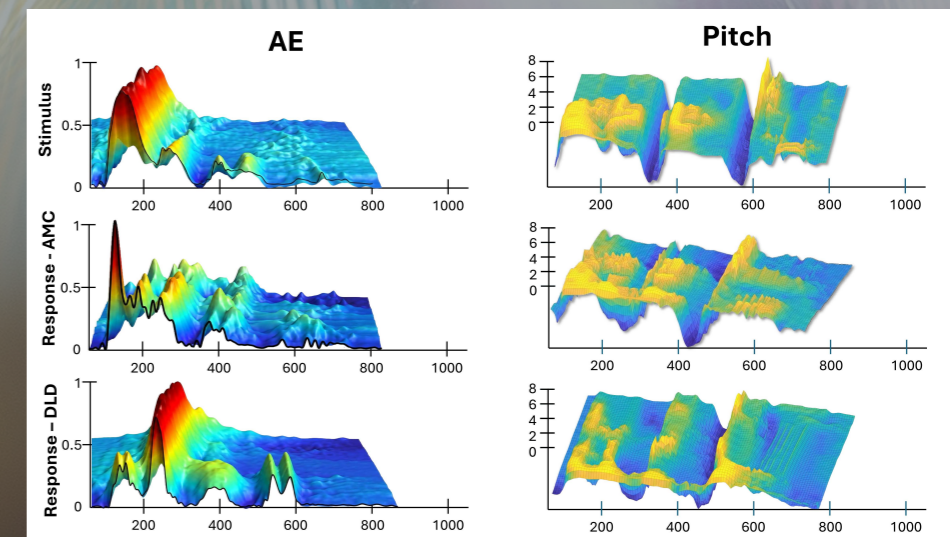


Figure 1: This figure is a collection of 3D visuals looking at the amplitude envelopes (AE) (left) and pitch contours (right) produced by selected individual children in each group for the target stimulus. The top set shows the results for the target stimulus 'Aladdin.' Row A is the target word spoken by an adult, row B is for an example control child (CA) and row C is a DY2 child (a child with dyslexia waiting for intervention). The bottom set shows the results for the target stimulus 'Alligator.' Row A is the target word spoken by an adult, row B depicts the information for the representative age matched controlled child (AMC) and row C depicts this information for the representative participant with developmental language disorder (DLD).

Prior studies of English-speaking children with either DLD or DD have used the neural technique of electroencephalography (EEG) to test Temporal Sampling theory. EEG can measure the automatic neural tracking of speech rhythms with millisecond accuracy. In these studies, children listen to a story while their brain responses are recorded. New signal-processing methods for recreating heard speech from brain responses are then used to estimate speech-brain alignment. These methods show that dyslexic brains are less accurate in encoding slower rhythm patterns during natural speech listening. Dyslexic child brains are fine with faster rhythm patterns in natural speech (Mandke et al., 2022). DLD brains are also less accurate at encoding slower rhythm patterns during story listening, as demonstrated using magnetoencephalography (MEG; Keshavarzi et al., 2026). We can think of the language-disordered brain as always coming in slightly too early (or too late) in terms of surfing the sound wave, but only for the slower energy patterns in speech.

These processing differences are subtle. They do not mean that children with DLD or with DD cannot learn spoken language. Perceiving some of the temporal rhythm patterns in the sound wave differently is a bit like being colour blind. If you are colour blind, you can still see, but your sensitivity to certain wavelengths of light is reduced. You cannot really distinguish reds, greens, browns, and oranges; they look very similar. But you can still navigate your environment. In both DLD and DD, affected children can still hear. They pass medical hearing screens, and they still learn language. However, their sensitivity to syllable stress patterns is reduced. They cannot easily distinguish whether a word like 'rhinoceros' has primary syllable stress on the second syllable (it does). This perceptual difficulty has subtle effects on all levels of linguistic processing. It affects word syllabification and rhyme judgements as well as prosodic processing. It even affects hearing individual sounds ('phonemes', which are approximately the sound units represented by the alphabet) in spoken words.

In the Disordered-Speech project, we hypothesise that these difficulties in hearing syllable stress patterns should affect how children produce syllable stress patterns. Careful acoustic measurements should reveal that when children with DLD or DD say a word like 'zebra' or 'rhinoceros', either the amplitude envelope or the intonation contour (the pitch contour) should be inaccurate. When we say words aloud, we are essentially producing 3D patterns of sound—variations in pitch and loudness (amplitude) that are extended in time. But children with DLD or DD should produce measurably different 3D sound patterns. The acoustic measurements that we will use in the Disordered-Speech project model words as 3D objects in terms of pitch, amplitude, and time, and look for systematic differences.

The project is based on data from English-speaking children. Speech produced by 8–9-year-old children with DLD versus DD has been acoustically assessed in English. The children simply copied aloud multisyllabic words like 'zebra' and 'rhinoceros', and we recorded their voices. While children with DD used highly accurate pitch contours, they generated inaccurate amplitude envelopes. By contrast, children with DLD generated both inaccurate pitch contours and inaccurate amplitude envelopes. This is difficult to hear with the naked ear. However, novel acoustic analysis techniques revealed significant group differences. Some examples of the 3D objects produced by the children in our earlier project are shown in Figure 1.

The major question asked in the ERC Disordered-Speech project is whether similar patterns of inaccuracy will identify children with DLD or DD aged 8–9 years in other European languages, even though these languages sound different. We also investigate whether younger children (2–5 years) will already show differences in how they speak that could be used to identify early risk. Linguists have documented many differences in the prosodic structures and speech-rhythm typologies of the languages we have chosen to compare. However, prior speech modelling studies

conducted to test Temporal Sampling theory suggest that the perceptual foundations of language systems may be highly similar.

Temporal Sampling theory was also tested in the Cambridge UK BabyRhythm project. This project used EEG to measure neural tracking of speech rhythms in 120 infants from 2 months of age. These 120 infants were listening to nursery rhymes being sung in infant-directed speech or 'babytalk'. They were not at risk for language disorders. When we re-created the heard speech from the brain responses of the infants, we could 'read out' which linguistic units were being encoded. The data showed that rhythmic information was recorded with high accuracy from the first measurement periods. Indeed, rhythmic accuracy was comparable to that of the adult brain. By contrast, neural tracking of phonetic features emerged rather slowly during the first year of life and remained sparse at 11 months, when most infants were saying their first words.

The slower rhythms of speech (the amplitude modulations) are unconsciously exaggerated during babytalk, even though the speaker is not aware of this. When we talk to infants, we automatically increase the amplitude changes signifying stressed (stronger) syllables in our voice. We also produce stressed syllables more regularly in temporal terms. In effect, with babytalk, we present the rhythm patterns required for the brain to learn language in an optimal format. Another optimal format is singing to infants. There are 6000+ world languages, yet most infants learn to speak and comprehend these languages without difficulty. The Cambridge BabyRhythm project suggested that automatic speech-brain alignment to rhythm patterns is at the centre of this achievement.

Computational modelling of babytalk in both English and Japanese has demonstrated a core set of acoustic statistics supporting speech-brain alignment (Leong et al., 2017; Daikoku and Goswami, 2025). These statistics depend on the phase alignment of slower

amplitude modulations at ~2 Hz and ~5 Hz. These core rhythmic statistics repeat across different sentences and also characterise child-directed speech in Spanish (Perez-Navarro et al., 2022) and English nursery rhymes (Leong and Goswami, 2015). Nursery rhymes are often perfect metrical poems, and computational modelling of metrical poetry in French versus German reveals the same set of acoustic statistics (Daikoku, Lee and Goswami, 2024).

In essence, the statistics describe consistent dependencies between amplitude modulation patterns at different temporal rates in English, Japanese, Spanish, French, and German. Accordingly, disorders of language acquisition may manifest in similar ways across languages.

The duplication of these acoustic statistics across languages also helps to explain the long-standing puzzle of how the human brain acquires language. The infant studies and computational modelling studies suggest that speech rhythm patterns are the hidden 'glue' underpinning the development of a well-functioning language system. By using matched speech copying tasks with children speaking English, French, Spanish, and Basque, we can test the production side of this idea. If the 'hidden glue' cannot be detected accurately by your brain, you are unlikely to produce rhythm patterns accurately. And you are likely to develop a developmental language disorder.

References

Daikoku, T., Lee, C. and Goswami, U. (2024) 'Amplitude modulation structure in French and German poetry: Universal acoustic physical structures underpin different poetic rhythm structures', *Royal Society Open Science*, 11: 232005. Available at: <https://doi.org/10.1098/rsos.2320>.

Daikoku, T. and Goswami, U. (2025) 'The amplitude modulation structure of Japanese infant-directed speech: Longitudinal data reveal universal acoustic physical structures that accommodate both syllabic and moraic timing', *Neurobiology of Language*. Advance publication. Available at: <https://doi.org/10.1162/NOL.a.226>.

Goswami, U. (2011) 'A temporal sampling framework for developmental dyslexia', *Trends in Cognitive Sciences*, 15(1), 3-10. Available at: <https://doi.org/10.1016/j.tics.2010.10.001>.

Keshavarzi, M. et al. (2026) 'Slow-rate temporal sampling deficits during natural speech listening in children with Developmental Language Disorder', *BioRxiv*. Available at: <https://doi.org/10.64898/2026.04.16.718920>.

Leong, V. and Goswami, U. (2015). Acoustic-emergent phonology in the amplitude envelope of child-directed speech. *PLOS ONE*, 10(12), e0144411. Available at: <https://doi.org/10.1371/journal.pone.0144411>.

Leong, V. et al. (2017) 'The temporal modulation structure of infant-directed speech', *Open Mind*, 1(2), 78-90. Available at: https://doi.org/10.1162/OPMI_a_00008.

Mandke, K. et al. (2022) 'Neural sampling of the speech signal at different timescales by children with dyslexia', *Neuroimage*, 253, 119077. <https://doi.org/10.1016/j.neuroimage.2022.119077>.

Perez-Navarro, J. et al. (2022) 'Local temporal regularities in child-directed speech in Spanish', *Journal of Speech, Language and Hearing Research*, 65 (10), 3776-3788.



PROJECT SUMMARY

When we say words, we are essentially producing 3D patterns of sound—variations in pitch and loudness that are extended in time. When English children with a language disorder say words, they produce measurably different sound patterns. The project studies this phenomenon in 4 languages, English, Spanish, Basque and French, aiming to improve early identification of children at risk.

PROJECT PARTNERS

Professor Marie Lallier, Ikerbasque Research Professor, Head, Educational Neuroscience and Developmental Disorders group at the Basque Centre on Cognition, Brain and Language (BCBL), Spain.

Professor Marina Kalashnikova, Ikerbasque Research Professor, Head, Infant Language and Cognition group, BCBL, Spain.

Professor Sophie Bouton, Permanent Researcher, Head, Neural Coding and Neural Engineering of Speech Functions group, Institut Pasteur, Paris, France.

PROJECT LEAD PROFILE

Usha Goswami CBE, FRS, FBA, FAAAS, Member Leopoldina, Yidan Prize (2019), Professor of Cognitive Developmental Neuroscience at the University of Cambridge. She is a global leader in research into the sensory/neural mechanisms that underpin language acquisition and developmental dyslexia. Her 'Temporal Sampling' theory of language acquisition and language disorders, proposed in 2011, underpins the project and is now the dominant framework in the field.

PROJECT CONTACTS

Professor Usha Goswami
Department of Psychology, Downing Street,
Cambridge CB2 3EB, UK

- ✉ ucg10@cam.ac.uk
- 🌐 www.cne.psychol.cam.ac.uk/people/usha-goswami
- 🌐 www.neuroscience.cam
- 🌐 www.europeandisseminatio.eu/disordered-speech-by-usha-goswami/24438

FUNDING



This project has been funded by The European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme - grant agreement No. 101199408.

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union, or the ERC. Neither the European Union nor the granting authorities can be held responsible for them.