SISSA
Scuola
Internazionale
Superiore di
Studi Avanzati

Neuroscience Area
Ph.D. Program in Cognitive Neuroscience

Statistical Learning across Representational Levels: Neural and Behavioral Expression, Modality, Individual Variability and Lifespan

Candidate:
Claudia Ruzza

Supervisor:
Prof. Davide Crepaldi

Co-supervisors:
Prof. Richard N. Aslin
Dr. Erin S. Isbilen

Academic Year 2025–2026

Aknowledgements

I would first like to thank my supervisor, Prof. Davide Crepaldi. Thank you for guiding me through these years while always allowing me to discover what truly fascinated me about our field and letting my curiosity guide my scientific path. That freedom has been one of the most valuable aspects of my doctoral journey. One of the things I appreciated most was the trust you placed in me: the space to explore, to make mistakes, to follow unexpected directions, and ultimately to grow into an independent researcher. Research is rarely a straightforward process, and this PhD was no exception. Through setbacks, uncertainty, and challenges that at times felt overwhelming, I learned the importance of persistence and adaptability and that even the most intimidating obstacles (and God knows how many we had, considered that I embarked on a EEG project with a humanities background) can be tackled with patience, persistence and a healthy amount of stubbornness. For your guidance, your trust, and the opportunities you gave me to grow, I am deeply grateful.

My thanks also go to the other SISSA professors – Alessandro Treves, Domenica Bueti, Eugenio Piasini, Davide Zoccolan, Mathew Diamond, and Raffaella Rumiati – for their teaching, guidance, and for giving me the opportunity to embark on this journey.

I would also like to express my sincere gratitude to my dissertation committee: Laura Batterink and Tess Forest, for their kind words at the last IASL conference in San Sebastián, and for finding the time and energy to carefully read and provide valuable feedback on this thesis. My thanks also go to Chiara Santolin and Judith Gervain for taking the time to come to SISSA and be present for my dissertation, together with my dear internal committee members, Davide Crepaldi and Domenica Bueti.

This defense would certainly not have been the same without all of you. Thank you for your support, your thoughtful comments, and for contributing to such a meaningful milestone in my academic journey. To Laura, Tess, Chiara, Judith, and Domenica in particular, thank you also for being such inspiring examples of women in science. Seeing your leadership, expertise, and achievements throughout my academic journey has been both motivating and empowering, and a powerful reminder of the importance of continuing to raise high the flag of women in STEM.

When I began this journey, I could not have foreseen that I would one day have the privilege to learn from and be mentored and guided by Prof. Dick Aslin and Dr. Erin Isbilen.

Dick, working with you during my time at Yale was an opportunity I could not have imagined when I first started my academic journey. Like many students entering the field, I knew your name long before I had the chance to meet you, through papers, books, and the countless references to your work that shaped my understanding of cognitive science. What struck me most was not only your scientific insight, but the way you engaged with ideas and with people. Despite your extraordinary accomplishments, you always made me feel a peer. I still remember being surprised by how naturally you treated me as someone whose thoughts were worth listening to. That confidence and carefulness had a greater impact than you probably realize. Beyond your scientific achievements, I am grateful for your kindness, humor, openness and of course funny stories. During my time at Haskins, you and David Lewkowicz helped create an environment that felt welcoming, supportive, and deeply human. The idea that "once a Haskins, always a Haskins" is more than a saying; it is something I experienced firsthand. Thank you for your trust. It has been a privilege to learn from you.

Erin, meeting you has been one of the greatest gifts of my PhD journey. I still remember first seeing your work at the San Sebastián conference, shortly after I had transitioned from Linguistics to Cognitive Neuroscience and Statistical Learning. At the time, I admired your research from afar. Back then, the possibility of working with you and Dick at Yale felt far beyond anything I could realistically imagine. Thank you for believing in me and for helping make that opportunity possible. You made a little Gilmore girl's dream possible. Your mentorship has had a profound impact on my development as a researcher. Despite your remarkable accomplishments, you always treated me with kindness and generosity, making me feel that my ideas and contributions mattered. Beyond your scientific expertise, what I will remember most is the warmth with which you welcomed me. During my time at Yale, you made me feel not only part of a Lab, but part of a community. Sharing holidays, celebrations, and everyday moments with you and your family made being far from home much easier and gave me a sense of belonging that I will always cherish. I am deeply grateful for your guidance and your friendship. I feel incredibly fortunate to have you as both a mentor and a friend.

Speaking of San Sebastián, I want to thank also Louisa Bogaerts and Noam Siegelman for organizing the Consensus Workshop and lead the Consensus paper. Stealing the wording you used in your inspiring and touching Early Career Talk at IASL this year Noam, this experience has definitely on the highlights podium of my graduate journey and probably many of my future works will stem from exchanges on that event (some already have), so thank you and to all the amazing and talented researches than are listed among the authors of that paper!

This work would have never been possible without all the lab networks, both the SISSA and the Yale one.

The Language, Learning and reading (L2R) Lab would not have been the same without my beloved Lab mates. Each of you, in different ways, has shown me how to face the challenges that neuroscience sets before us, and together you have helped me enrich and expand my personal toolkit.

Maria Ktori, thank you for sharing you expertise with EEG. A wise woman once told me to never forget the very first person who gave you the tools you have to do what you do know. That person is you. Thanks for having shared with me all your expertise with your rare kindness and, together with Mara De Rosa and and Yamil Vidal, for letting me taking care of your XinX baby.

Elena Greatti, thank you for sharing with me the endless EEG sessions. You are one of the hardest-working and most determined people I know. I have always admired your perseverance and your ability to keep pushing forward, no matter the obstacles. I cannot wait to see what you will achieve next.

Olga Solaja and Jon Carr, thank you for sharing with me your expertise, your ideas (and your work on VSL task) over the years. I am especially grateful for the way you welcomed me into the group in San Sebastián when I was still a rather lost first-year PhD student, taking my very first steps into the world of Statistical Learning. Your encouragement made a lasting difference.

Anna Marie Liebelt, thank you for the crochet tips (and gifts), the crêpe recipes, the croissant and brioche tutorials, and the running advice. You have a remarkable ability to bring warmth and positivity wherever you go, and it has been a pleasure getting to know you. You are one of the brightest people I have met, and I truly wish you all the best. Keep running toward your goals, I have no doubt you will reach them.

Giulio Severjinen, thank you for being there in both the difficult moments and the happy ones. You will always be my favorite colleague to play cards with on train journeys and to make plans for Harry Potter movie nights with.

Thanks to Timon Kunze for your peaceful eyes and witty advices and to Elsa De Simone, for your contagious smile and optimism.

Doroteja Stamenkovic, thanks for always matching my energy as I told you bad or good or angry episodes happened in the lab and, most of all, for having always an high opinion of me and my work; it is hard to find someone as kind and supportive as you.

Thanks to our dear Master Students, now brilliant graduate students or career women. Sasha Kenjeeva, there since Francesco's and my early days in SISSA, thanks for letting me in and to have shared with me moments of laughter, of sadness, of frustration, and of happiness. It has been such a joy to have you by my side on my first two years of graduate school; you're a brilliant and sensitive mind, rare to find, can't wait to see what comes next.

Emma Peri and Emma Coba, thanks for your help in the lab, endless testing sessions would have been much more endless without our gossip sessions; and to all our other former master students: Ilaria Appel, Emma Icardi, Emma Montecchiari, Nare Haroutynian, Irmak Ergin, Paola Sartori.

Some of the most meaningful moments of this PhD took place during my time at Yale. What began as an academic opportunity quickly became much more.

Thanks also to all the LLAMB Lab and the Yale Child Study Center, Sara Sánchez-Alonso, Virginia Chambers, Nabin Koirala, Anna Püsök, Jinglei Ren, Aditya Chander for including me in scientific debates as well as celebrations and adventures.

Thanks to Roeland Hancock, William Mennie and Alexander Forrence for your help and training with the OP-MEG system at Wu Tsai Institute. Thanks as well to Mitchell Kim, Melissa Vergara and Mary Shamon Ayala for their valuable help with data collection and coding.

Although this thesis may appear to be the product of scientific collaborations, it was also made possible by a much broader network of people, one that is often overlooked in academia precisely because it exists beyond its boundaries.

Over the past few years, I have learned that the future is often unpredictable. Life has taught me that we do not always have the time, the energy, or the strength we imagine we do, and that affection should never remain unspoken. For this reason, when it comes to the people who matter, I would rather say one word too many than one too few. This thesis is already long enough, so a few extra pages will not hurt.

According to the 2026 ISTAT Report, the rate of early school leaving in Italy is 8.2%. Behind this average, however, lies a striking evidence: the rate rises to 20.7% among children whose parents have at most a middle-school education and falls to just 1.1% among those with at least one parent holding a university degree. I am the first person in my family, including my extended family, to go beyond middle school. Statistics would suggest that the odds were not exactly in my favor. And yet, here I am. I made it here also – and perhaps above all – because you, Mum and Dad, believed in education long before it could be measured by a diploma or a degree. You encouraged me to improve, supported me financially as you could, and allowed me to seize opportunities, often by giving up things for yourselves. If I am able to write these words in a doctoral dissertation today, it is because of you. This achievement belongs as much to you as it does to me.

Thanks to my grandmother Sandra, the person whom I spent much of my childhood and adolescence with, before her memories slowly began to fade. Born in 1936, one might expect a woman shaped by the values of another era, when a woman's place in society was measured primarily through her roles as a wife and a mother. Not her. Proud of her 5 years elementary school degree (many people at that time left school at the end of the 3rd year), she has always seen further than many of her peers. She never stopped encouraging me to study, work hard, and pursue my dreams. Hardly a week goes by without me asking whether she would have been proud of what I have achieved. Thanks also to my grandfather Mario, who taught me as a child that we are capable of doing things others do not expect from us and that the only real limit is imagination. On an occasion like this, what I miss most is hearing you say in Venetian dialect (because it was the only language you spoke), "Sandra, go buy some pastries, the baby is graduating."

Thanks also to Grandma Rosetta, Aunt Sonia, and Uncle Johnny for always being there, and to the entire *Monarin* crew – Aunt Teresa, Aunt Daniela and Uncle Lino, Uncle Berto and Aunt Novella – for always having a hug ready for me when my grandmother's were not the same as they were.

Alessandro and Francesco, today it is my turn, but yours will come soon enough. Jokes aside, you are the closest thing I have to brothers. Watching you grow up has been wonderful, and being part of that journey has been an honor. My presence may have been inconsistent over these past few years, but you never judged me for it and were always ready for a Melograno burger whenever I needed one. I love you both dearly.

Beyond the family we are born into, there is also the family we choose.

First among them is Alessio, who began this journey as my boyfriend and is now my husband (or, as my parents like to call him in dialect, "son by soul"). You encouraged and supported me in every possible way, from bringining me food and water (mostly water, since I am essentially a succulent and would probably only remember to drink once a month if left to my own devices), to helping me when Python refused to cooperate (although, more often than not, I was the one refusing to cooperate), to brightening my days and making sure the Shadow never swallowed me whole. I do not think there are words capable of expressing just how much this achievement belongs to you as well. The *Team Cucu* is the best, and I cannot wait to see what adventures (both academic and beyond) the future has in store for us.

There are many families we choose along the way.

Some have always been there, like my Siena family. It is rare to be able to say that there are people in your life with whom nothing is off limits, people you can say absolutely anything to. Besides Alessio, I am lucky enough to have three.

Marta, a brilliant mind and exceptional in more ways than I can count, who guided us (and occasionally dragged us) through our Master's years and well beyond. I do not know whether you realize it, but I doubt I speak only for myself when I say that you have been a role model for all of us in many ways.

Gaia, a vibrant and endlessly entertaining soul, full of stories, colors, and emotions. One of the most refreshing people in my life, someone with laughter and tears come equally naturally. People like you are hard to find.

And Silvia, a creative and sharp-minded spirit. She is music, words, and images all at once; a soul that seems to belong to many generations, and one of the most compassionate and loving gazes I know and always find comfort in returning to.

Other families, instead, are newer. SISSA gifted me several of them.

First, among the SISSA Very Important People, is Francesco. You have been there since my very first day at SISSA and have been an inseparable companion ever since. As you once said, the same is true for me, though with one small correction: there would not have been a first, second, third, or fourth Christmas without you (have I taught you nothing? Time is measured in Christmases). No amount of laughter, complaints, confidences, hugs, or fast-food dinners will ever be enough. I could not have wished for a better friend. Speaking of Fra, it would be impossible not to mention Kendall, one of the most selfless and kind-hearted people I know. Very few people possess your capacity for genuine care and generosity.

Giulio, probably the person most different from me whom I have nevertheless chosen to keep so close. Over the years, we have argued more times than I can remember, yet every disagreement left me with something valuable: a new perspective, an extra question, a different way of looking at the world. When we first met, I could never have imagined the role you would come to play in my life. Behind a personality that can sometimes be rough around the edges and a uniquely personal way of interpreting the world, I discovered someone capable of remarkable sensitivity and depth. Your curiosity does not stop at research, it is the engine through which you approach life, constantly questioning what surrounds you and, remarkably, yourself. Thank you for making these years more stimulating, more challenging, and undoubtedly more interesting.

Ben, one of the most intelligent—and at the same time inexplicably insecure—people I know. I am honored that, over the years, you have considered me both a friend and someone you could confide in. I will always treasure the memories of our breakfasts at Mapo, our afternoon tea-times, and the endless evening conversations about physics (I still strongly suspect your confidence in my understanding of the subject is wildly misplaced), books, films, cats, and, of course, every possible form of gossip. Of all your qualities, however, my favorite remains your extraordinary ability to be, entirely effortlessly, the *potatoest* person I've ever met. I will miss your hugs (and head kisses).

Meriem, the Rory to my Paris (and vice versa). I hardly know where to begin. Over these years, you have been a constant presence in my life, one of those people who somehow make difficult days lighter and happy days even brighter. It is hard to imagine this journey without our conversations, our daily updates, and a friendship that has gradually become one of the most precious things I take away from these years. With you, I have shared cultural references (aka TikTok trends), laughter, sadness, reflections, joys, anxieties, frustrations, and dramas of every kind. The sort of things that only happen with someone who truly knows you. Thank you for being there when things were going well, when they were going badly, when they were somewhere in between, and when they were not going at all. There is something deeply comforting about knowing that someone feels you and understands what you are going through. You are one of the most wonderfully complex and brilliant minds I know. I am lucky to call you my friend.

The Trio Indrio would not be complete without Lodovico Capuano, the group's other Venetian half. Thank you for the gossip sessions and the evening walks at Molo, for our discussions about life's biggest questions as well as complete nonsense, for our shared love of German and stand-up comedy. I realized we were going to become great friends during the Eleazaro phase, and I was not wrong.

This journey would not have been the same without Cecilia Sgalletta's gentle smile, Alfredo Stanzione's gossip, Marco Pasini's generosity, Francesco Gabrielli's jokes (and impressions), Clemente Smarra's confidence boosts, Marina Moleti's hospitality, the conversations about travelling, India, and supervisors with Amartya Harsh Singh, Anant Shri, Shreyansh Agrawal, and Minahil Butt, the post-lunch slackline attempts and table-football tournaments with Michele Fossati (pizza-gang forever), and Davide Bason's infectious laughter.

Thank you also to Anna Delmonte, Cristiano Mucci, Vincenzo Naso, and Francesco Gentile for the lunchtime conversations in the cafeteria and for all the wonderful evenings spent together in Barcola, at Dopolavoro, at Marino's, and along Viale. You all had a significant impact on the life of my work-life balance.

Speaking of the cafeteria, it would be impossible not to mention Patrizia, Debora, Francesca, Barbara, and Lucio. People say you do not realize how much something means to you until it is gone, and your cuisine and coffee-break conversations are the perfect example of that. Together with Cristina's welcoming smile at reception, you helped make SISSA feel like home.

Last, but certainly not least, the people of SISSA for Schools, who introduced me to the joy of science communication. In particular, thank you to Francesca, Elvira, and Marta. Thanks for including me in your activities and for trusting me enough to involve me in projects such as SISSA Open Day and Trieste Next. Thursday mornings with you and the children were often exactly the boost I needed and the answer to the question "Why are you doing this?" at times when my compass had been completely demagnetized by the anxiety of the countless projects and deadlines that so often come with academic life.

To you and all the others who made these four years truly special: this achievement would not have been possible without the support and contributions of each of you. Thank you.

Ringraziamenti

Vorrei innanzitutto ringraziare il mio relatore, il Prof. Davide Crepaldi. Grazie per avermi guidata in questi anni lasciandomi sempre la libertà di scoprire ciò che davvero mi appassionava e di aver lasciato che fosse la curiosità ad orientare la mia ricerca. Questa libertà è stata uno degli aspetti più preziosi del mio percorso di dottorato. Una delle cose che ho apprezzato maggiormente è stata la fiducia che hai riposto in me: lo spazio per esplorare, sbagliare, seguire direzioni inaspettate e, infine, crescere come ricercatrice. La ricerca raramente segue un percorso lineare e questo dottorato non ha fatto eccezione. Tra battute d'arresto, incertezze e sfide che a volte sembravano insormontabili, ho imparato l'importanza della perseveranza e della capacità di adattarsi, e che anche gli ostacoli più intimidatori (e Dio sa quanti ne abbiamo incontrati, considerando che ho intrapreso un progetto EEG provenendo da una formazione umanistica) possono essere affrontati con pazienza, costanza e una sana dose di testardaggine. Ti sono profondamente grata per la tua fiducia e per tutte le opportunità che mi hai dato.

Un ringraziamento va anche agli altri professori della SISSA — Alessandro Treves, Domenica Bueti, Eugenio Piasini, Davide Zoccolan, Mathew Diamond e Raffaella Rumiati — per il loro insegnamento, la loro guida e per avermi dato l'opportunità di intraprendere questo percorso.

Quando ho iniziato questo viaggio, non avrei mai immaginato che un giorno avrei avuto il privilegio di imparare da e di essere guidata dal Prof. Dick Aslin e dalla Dott.ssa Erin Isbilen.

Dick, lavorare con te durante il mio periodo a Yale è stata un'opportunità che non avrei mai potuto immaginare quando ho iniziato il mio percorso accademico. Come molti studenti che si avvicinano a questo campo, conoscevo il tuo nome molto prima di avere la possibilità di incontrarti, attraverso articoli, libri e le innumerevoli citazioni del tuo lavoro che hanno contribuito a formare la mia (e di tanti altri) comprensione delle scienze cognitive. Ciò che mi ha colpito di più non è stata soltanto la tua straordinaria competenza scientifica, ma il modo in cui ti relazioni alle idee e alle persone. Nonostante i tuoi risultati eccezionali, mi hai sempre fatta sentire una pari. Ricordo ancora il mio stupore nel vedere con quanta naturalezza trattassi le mie idee come degne di essere ascoltate. Quella fiducia ha avuto su di me un impatto molto più grande di quanto probabilmente immagini. Oltre ai tuoi successi

scientifici, ti sono grata per la tua gentilezza, il tuo umorismo, la tua disponibilità e i tuoi aneddoti. Durante il mio periodo a Yale, tu e David Lewkowicz avete contribuito a creare il LLAMB lab un ambiente accogliente, stimolante e profondamente umano. L'espressione "once a Haskins, always a Haskins" non è soltanto un motto: è qualcosa che ho sperimentato in prima persona. Grazie per la fiducia che mi hai accordato. È stato un onore e privilegio poter imparare da te.

Erin, averti incontrata è stato uno dei regali più grandi che il mio percorso di dottorato mi abbia offerto. Ricordo ancora la prima volta che ho visto il tuo lavoro alla conferenza a San Sebastian, poco dopo il mio passaggio dalla linguistica alle neuroscienze cognitive e allo statistical learning. All'epoca ammiravo la tua ricerca da lontano e l'idea di lavorare con te a Yale era ben oltre qualsiasi cosa potessi realisticamente immaginare. Grazie per aver creduto in me e per aver contribuito a rendere possibile quell'opportunità. Hai reso possibile il sogno di una piccola ragazza Gilmore. La tua guida ha avuto un impatto profondo sulla mia crescita come ricercatrice. Nonostante i tuoi numerosi successi, mi hai sempre trattata con rispetto e gentilezza, facendomi sentire che le mie idee e il mio contributo avevano valore. Al di là delle tue competenze scientifiche, ciò che ricorderò più di tutto è il calore con cui mi hai accolta. Durante il mio periodo a Yale mi hai fatta sentire non solo parte di un laboratorio, ma parte di una comunità. Condividere festività, celebrazioni e momenti quotidiani con te e la tua famiglia ha reso molto più facile essere lontana da casa e mi ha dato un senso di appartenenza che conserverò sempre con me. Ti sono profondamente grata per la tua guida, la tua amicizia e la tua fiducia. Mi considero incredibilmente fortunata ad averti avuta al mio fianco sia come mentore sia come amica.

A proposito di San Sebastián, vorrei ringraziare anche Louisa Bogaerts e Noam Siegelman per aver organizzato il Consensus Workshop e guidato la stesura del Consensus Paper. Prendendo in prestito le parole che avete usato nel vostro stimolante e toccante Early Career Talk all'IASL di quest'anno, questa esperienza è stata senza dubbio uno dei momenti più significativi del mio percorso di dottorato. Probabilmente molti dei miei lavori futuri nasceranno proprio dagli scambi avuti durante quell'evento (alcuni lo hanno già fatto), quindi grazie di cuore!

Questa tesi non sarebbe stata possibile senza le reti di persone che hanno reso speciali questi anni, sia alla SISSA sia a Yale. Il laboratorio L2R non sarebbe stato lo stesso senza i miei compagni. Ognuno di voi, a modo proprio, mi ha insegnato qualcosa su come affrontare le sfide che le neuroscienze ci pongono davanti e, insieme,

avete contribuito ad arricchire il mio personale bagaglio di strumenti scientifici e umani.

Maria Ktori, grazie per aver condiviso con me la tua esperienza con l'EEG. Una donna saggia una volta mi disse di non dimenticare mai la prima persona che ti ha dato gli strumenti per fare ciò che fai oggi. Per me quella persona sei tu. Grazie per aver condiviso il tuo sapere con una gentilezza rara e, insieme a Mara De Rosa e Yamil Vidal, per avermi affidato il vostro cucciolo "XinX".

Elena Greatti, grazie per avermi fatto compagnia durante le interminabili sessioni di EEG. Sei una delle lavoratrici più instancabili che conosca; ho sempre ammirato la tua perseveranza. Non vedo l'ora di scoprire cosa riuscirai a realizzare in futuro.

Olga Solaja e Jon Carr, grazie per aver condiviso con me negli anni la vostra esperienza, le vostre idee e per avermi permesso di ereditare il votro lavoro sul VSL. Vi sono particolarmente grata per avermi accolta nel gruppo a San Sebastián quando ero ancora una spaesata studentessa del primo anno di dottorato che muoveva i suoi primi passi nel mondo dello Statistical Learning. Il vostro incoraggiamento ha fatto una differenza enorme.

Anna Marie Liebelt, grazie per i consigli sull'uncinetto (e per i regali), le ricette per le perfette crêpes, le guide su croissant e brioche e i consigli sulla corsa. Hai una straordinaria capacità di portare calore e positività ovunque tu vada, ed è stato un piacere conoscerti. Sei una delle persone più solari che abbia incontrato e ti auguro davvero il meglio. Continua a correre verso i tuoi obiettivi: non ho alcun dubbio che li raggiungerai.

Giulio Severijnen, grazie per esserci stato sia nei momenti difficili sia in quelli felici, e per avermi sempre fatta sentire accolta e compresa. Rimarrai sempre il mio collega preferito con cui giocare a carte durante i viaggi in treno e fare piani per serate con cibo da asporto sul divano guardando Harry Potter.

Timon Kunze, grazie per il tuo sguardo sereno, per i tuoi consigli sempre intelligenti e per la tua capacità di farmi vedere le cose da una diversa prospettiva. Elsa De Simone, grazie per il tuo sorriso contagioso e per il tuo ottimismo, che riuscivano a migliorare anche le giornate più lunghe.

Dora Stamenkovic, grazie per aver sempre saputo stare al passo con il mio entusiasmo, la mia rabbia o la mia frustrazione ogni volta che succedeva qualcosa in laboratorio. Soprattutto, grazie per aver sempre avuto una stima così alta di me e del mio lavoro, spesso più alta della mia. Persone gentili e di supporto come te sono davvero rare.

Un ringraziamento speciale va ai nostri meravigliosi studenti magistrali, oggi brillanti dottorandi e professionisti.

Sasha Kenjeeva, lì fin dai primi giorni miei e di Francesco alla SISSA, grazie per avermi accolta e per aver condiviso con me momenti di risate, tristezza, frustrazione e felicità. Averti accanto nei miei primi anni di dottorato è stato un vero regalo. Hai una mente brillante e sensibile (combinazione rara in accademia), e non vedo l'ora di vedere dove e cosa ti porterà il futuro.

Emma Peri ed Emma Coba, grazie per tutto il vostro aiuto in laboratorio. Le già infinite sessioni di testing sarebbero state molto più infinite senza le nostre sessioni di gossip. Un grazie anche a tutti gli altri nostri ex studenti magistrali: Ilaria Appel, Emma Montecchiari, Nare Haroutynian, Irmak Ergin e Paola Sartori.

Alcuni dei momenti più significativi di questo dottorato hanno avuto luogo durante il mio periodo a Yale. Quella che era iniziata come un'opportunità accademica è diventata molto di più.

Grazie a tutte le persone del LLAMB Lab e dello Yale Child Study Center — Sara Sánchez-Alonso, Virginia Chambers, Nabin Koirala, Anna Püsök, Jinglei Ren e Aditya Chander — per avermi coinvolta non solo nei dibattiti scientifici, ma anche nelle celebrazioni, nelle avventure e nella vita quotidiana.

Un ringraziamento va anche a Roeland Hancock, William Mennie e Alexander Forrence per il supporto e la formazione sul sistema OP-MEG presso il Wu Tsai Institute. Grazie inoltre a Mitchell Kim, Melissa Vergara e Mary Shamon Ayala per il prezioso aiuto nella raccolta e codifica dei dati.

Sebbene possa sembrare frutto di collaborazioni scientifiche, questa tesi è stata resa possibile anche da una rete molto più ampia e spesso sottovalutata nel mondo accademico perchè al di fuori di esso. Negli ultimi anni ho imparato che il futuro è spesso imprevedibile. La vita mi ha insegnato che non sempre abbiamo il tempo, le energie, la forza che immaginiamo, e che l'affetto non dovrebbe mai rimanere implicito. Per questo, quando si parla delle persone importanti, preferisco spendere una parola in più (questa tesi è già lunga, perciò qualche pagina in più non guasterà).

Secondo i dati Istat 2026, il tasso di abbandono scolastico precoce in Italia è dell'8,2%. Dietro questa media, però, si nasconde un enorme divario: il tasso sale al 20,7% tra i figli di genitori con al massimo la licenza media e scende all'1,1% tra chi ha almeno un genitore laureato. Io sono la prima laureata della mia famiglia, anche allargata. Le statistiche direbbero che le probabilità non erano esattamente dalla mia parte. Eppure ci sono arrivata, anche (e soprattutto), perché voi, mamma e papà, avete creduto nell'istruzione molto prima di poterla misurare con un titolo

di studio. Mi avete spinta a migliorare, sostenuta, dedicato tempo ed energie. Mi avete permesso di cogliere opportunità, spesso rinunciando a qualcosa per voi stessi. Se oggi posso scrivere queste righe sulla mia tesi di dottorato, lo devo a voi. Questo traguardo è tanto mio quanto vostro.

Grazie a mia nonna Sandra, la persona con cui ho passato gran parte della mia infanzia e adolescenza, prima che i suoi ricordi iniziassero a sbiadire. Classe 1936, ci si aspetterebbe di trovare una donna con valori di un secolo fa, dove nella società una donna veniva misurata come moglie e madre. Non lei. Fiera della sua quinta elementare e instancabile lavoratrice, ci ha sempre visto lungo, spingendomi sepmpre ad impegnarmi nello studio e ad inseguire i miei sogni. Non passa settimana in cui non mi chieda se tu sia fiera di quello che ho raggiunto. Grazie al nonno Mario, che mi ha insegnato da bambina che possiamo fare cose che gli altri non si aspettano e che il limite è solo l'immaginazione. In questa occasione, la cosa che manca di più è il tuo "Sandra, va tore e paste che ea tosa se laurea".

Grazie anche a Nonna Rosetta, zii Sonia e Johnny, per esserci, e al gruppo *Monarin* – zia Teresa, zii Daniela e Lino, zii Berto e Novella – per avere avuto un abbaccio sempre pronto per me, quando quello della nonna è venuto meno.

Alessandro e Francesco, oggi tocca a me ma i prossimi siete voi! Scherzi a parte, siete la cosa più vicina che io abbia a dei fratelli. E' stato meraviglioso vedervi crescere e un onore essere stata parte di questo processo. La mia presenza è stata incostante in questi ultimi anni ma voi non mi avere mai giudicata e siete sempre stati pronti per un Melograno quando ne avevo bisogno. Vi voglio un bene dell'anima.

Oltre a quella di sangue, c'è anche la famiglia che ti scegli. Primo tra questi Alessio, all'inizio di questo percorso mio compagno, ora mio marito (o come lo definiscono i miei genitori, fioeo de anema). Mi hai spronata e supportata in ogni momento e modo possibile, dall'aiutarmi con python che non risponde (anche se ero sempre io a non rispondere), ad illuminare le giornate facendo sì che l'Ombra non mi inghoittisse. Non credo ci siano parole per descrivere quanto questo traguardo sia anche tuo.

Di famiglie che ti scegli, ce ne sono tante. E' raro poter dire di avere nella propria vita delle persone a cui non c'è nulla che non si possa dire. Io, oltre ad Alessio, ne ho tre.

Marta, una mente brillante e fuoriclasse in tanti aspetti, che ci ha guidat* (e, a tratti, trascinat*) nel nostro percorso in magistrale (e oltre). Non so se lo sai, ma non credo di parlare solo a nome mio se dico che sei stata il nostro modello in tante cose.

Gaia, un'anima divertente e ricca, di aneddoti, di colori, di emozioni. Una persona tra le più rinfrescanti nella mia vita, con cui ci si può aprire alla risata e al pianto con la stessa naturalezza. Difficile trovarne di simili.

E Silvia, una personalità creativa e sagace. Lei è musica, parola, immagine; un'animo che appartiene a tante generazioni e tra gli sguardi più compassionevoli e amorevoli a cui mi piace tornare.

Alcune famiglie son più datate, altre, invece, più nuove. La SISSA me ne ha regalate alcune.

Primo tra le SISSA Very Important People, c'è Francesco. Con me dal mio primo giorno in SISSA e da lì compagno inseparabile. Come hai detto tu, lo stesso vale per me, ma con un piccolo cambiamento: non ci sarebbe stato un primo, né un secondo, un terzo, né un quarto Natale senza di te (ma non ti ho insegnato nulla? Il tempo si misura in Natali). Risate, lamentele, confidenze e abbracci non saranno mai abbastanza con te (come anche cene a base di fast-food). Non avrei potuto chiedere di meglio, in quanto ad amicizie. Parlando di Fra, impossibile non menzionare Kendall, tra gli animi più altruisti e gentili che conosca. Poche persone hanno la cura ed attenzione disinteresata che hai tu.

Giulio, probabilmente la persona più diversa da me che abbia scelto di tenere così vicina. In questi anni ci siamo scontrati più volte di quante riesca a ricordare, ma ogni discussione mi ha lasciato qualcosa: una prospettiva nuova, una domanda in più, un modo diverso di guardare le cose. Quando ci siamo conosciuti non avrei immaginato il ruolo che avresti avuto nella mia vita. Dietro un carattere a volte spigoloso e un modo tutto tuo di interpretare il mondo, ho scoperto una persona capace di una sensibilità e di una profondità rare. La tua curiosità non si limita alla ricerca o al lavoro: è il motore con cui affronti la vita, mettendo continuamente in discussione ciò che ti circonda e, spesso, anche te stesso. Nel tempo ho avuto la fortuna di vederti cambiare, crescere e mostrarti per quello che sei davvero. E più ti conosco, più mi rendo conto di quanto ci sia ancora da imparare da te. Grazie per aver reso questi anni più stimolanti, più complessi e, senza dubbio, più interessanti.

Ben, una delle menti più intelligenti, e allo stesso tempo, inspiegabilmente insicure, che io conosca (combinazione piuttosto affascinante per una neuroscienziata). Sono onorata che, nel corso di questi anni, tu mi abbia considerata un'amica e una persona con cui confidarti. Porterò sempre con me i ricordi delle nostre colazioni

da Mapo, dei tè condivisi e delle infinite conversazioni serali sulla fisica (continuo a sospettare che la tua fiducia sulle competenze in materia sia a dir poco mal riposta), libri, film, gatti e, naturalmente, pettegolezzi di ogni genere. Tra tutte le tue qualità, però, quella che preferisco è la tua straordinaria capacità di essere, senza alcuno sforzo, la persona più adorabilmente "patata" che abbia mai conosciuto. Mi mancheranno i tuoi abbracci (e baci in testa).

Meriem, la Rory della mia Paris (e vice versa). Non saprei nemmeno da dove iniziare. In questi anni sei stata una presenza costante, una di quelle persone che riescono a rendere più leggere le giornate difficili e ancora più belle quelle felici. È difficile immaginare questo percorso senza le nostre conversazioni, i nostri aggiornamenti quotidiani e un'amicizia che, con il tempo, è diventata una delle cose più preziose che porto con me di questi anni. Con te ho condiviso riferimenti culutali (aka TikTok trends), risate, tristezze, riflessioni, gioie, ansie, frustrazioni, drammi. Cose che arrivano solo con chi ti conosce davvero. Grazie per esserci stata quando le cose andavano bene, quando andavano male, quando andavano così e così e quando non andavano proprio. C'è qualcosa di profondamente rassicurante nel sapere che qualcuno prova le tue stesse emozioni e comprende perfettamente ciò che stai vivendo. Sei una delle menti più meravigliosamente complesse e luminose che io conosca. Sono fortunata a poterti chiamare amica.

Il Trio Indrio non sarebbe completo senza Lodovico Capuano, l'altra metà veneta del gruppo. Grazie per le sessioni di gossip e le passeggiate serali al Molo, per le discussioni sui massimi sistemi e su baggianate, per l'amore condiviso per il tedesco e la stand-up comedy (ho capito che saremmo diventati grandi amici con Eleazaro, non mi sbagliavo).

Questo percorso non sarebbe stato lo stesso senza il sorriso gentile di Cecilia Sgalletta, il gossip di Alfredo Stanzione, la generosità di Marco Pasini, lo sguardo gentile di, le battute (e imitazioni) di Francesco Gabrielli, i boost di autostima di Clemente Smarra, l'ospitalità di Marina Moleti, le chiacchierate sui viaggi, sull'India e sui supervisor con Amartya Harsh Singh, Anant Shri, Shreyansh Agrawal e Minahil Butt, i tentativi di slack-line post-pranzo e I tornei di biliardino con Michele Fossati (pizza-gang forever), la risata contagiosa di Davide Bason.

Grazie anche a Anna Delmonte, Cristiano Mucci, Vincenzo Naso, Francesco Gentile. Per le chicchierate in mensa a pranzo, e per tutte le belle serate passate assieme a Barcola, al Dopolavoro, da Marino e in Viale. Avete significativamente impattato la parte "vita" del mio bilanciamento "vita-lavoro".

Parlando di mensa, impossibile non menzionare Patrizia, Debora, Francesca, Barbara e Lucio. Quando si dice "non sai di avere qualcosa finché non lo hai perso", e i pranzi preparati da voi e le chiacchiere alle pause caffé ne sono il perfetto esempio.

Assieme al sorriso accogliente di Cristina alla reception, avete contribuito a rendere la SISSA una casa.

Ultimi, ma non per importanza, le persone di SISSA for Schools, coloro che mi hanno fatto scoprire l'amore per la divulgazione. In particplar modo grazie a Francesca, Elvira Marta. Grazie per avermi inclusa nelle vostre attività e per avermi dato così fiducia da includermi in progetti quali il SISSA Open Day e Trieste Next. I giovedì mattina con voi e con i bambini sono stati spesso la carica che mi serviva e la risposta al "perchè lo fai" in momenti in cui la mia bussola era smagnetizzata dall'ansia dei mille progetti e scadenze tanto comuni in accademia.

A voi e a tutti coloro che hanno reso questi anni così significativi: questo traguardo porta il mio nome, ma non sarebbe stato possibile senza il sostegno, la fiducia e il contributo di ciascuno di voi. Grazie di cuore.

A page for notes and memories

Abstract

Statistical learning (SL) refers to the ability to extract regularities from structured sensory input and has been proposed as a fundamental mechanism supporting language acquisition, reading, and adaptive behavior. Despite extensive research, important questions remain regarding the nature of SL, its organization across modalities, its development across the lifespan, and the relationship between neural and behavioral expressions of learning. The present thesis addressed these questions through a series of behavioral and imaging studies examining SL in children, young adults, and older adults.

Across the empirical chapters, SL was investigated using both neural frequency-tagging paradigms and behavioral measures, with a particular focus on modality-specificity, cross-modal transfer, individual variability, and lifespan changes. Results revealed robust neural sensitivity to structured regularities in both young and older adults, even when explicit behavioral evidence of learning was weak or absent. Cross-modal studies further demonstrated that learning and transfer varied across auditory and visual domains, suggesting that SL is shaped by both shared computational principles and modality-specific representational constraints. Developmental and aging findings indicated that some forms of sensitivity to regularities remain relatively preserved across the lifespan, although the behavioral expression of learning and the integration of information across modalities may differ as a function of age.

Across studies, substantial inter-individual variability emerged, and correlations between neural and behavioral measures were often limited. These findings challenge strongly unitary conceptions of SL and instead support the view that learning is best understood as a multicomponential phenomenon. Building on these results, the thesis proposes a multi-level framework in which different paradigms capture distinct expressions of learning that vary in representational format, processing demands, and accessibility to conscious report. Within this framework, neural entrainment primarily reflects early online sensitivity to statistical structure, whereas behavioral measures additionally recruit processes related to prediction, memory retrieval, decision-making, and metacognitive reflection.

Overall, the present work argues that SL is neither fully domain-general nor entirely modality-specific, but rather emerges from dynamic interactions among perceptual systems, prior knowledge, developmental factors, and individual cognitive differences. By integrating neural and behavioral methodologies across modalities and age groups, the thesis contributes to a more comprehensive understanding of how humans acquire and generalize statistical structure in complex sensory environments.

Forewords

This thesis brings together the work carried out during my doctoral studies at SISSA. When I began this journey, my interests revolved around language, learning, and the mechanisms that allow humans to extract structure from the sensory world. Over the years, these interests developed through a series of projects that, at first glance, appeared only loosely connected. Looking back, however, they all converged on a common question: how do humans acquire, represent, and generalize statistical regularities across modalities, developmental stages, and levels of processing?

The chapters of this thesis weave together findings from the following projects:

- **Ruzza, C.**, & Crepaldi, D. *Neural Sensitivity without Behavioural Evidence: Frequency-Based Statistical Learning in Fast Periodic Visual Stimulation.* Manuscript currently under revision for publication in *Journal of Memory and Language*. Pre-print available at PsyArXiv: https://osf.io/preprints/psyarxiv/n4fjr_v1.
- **Ruzza, C.**, & Crepaldi, D. *Neural and Behavioural Signatures of Statistical Learning in Aging: Evidence from Fast Periodic Visual Stimulation.* Manuscript currently under revision for publication in *Psychology of Aging*. Pre-print available at PsyArXiv: https://osf.io/preprints/psyarxiv/u9xj5_v1.
- **Ruzza, C.**, Isbilen, E. S., & Aslin, R. N. *Cross-Modal Transfer of Statistical Learning: Evidence from Auditory and Visual Sequence Learning.* Manuscript currently under revision for publication in *Journal of Experimental Psychology: General*. Pre-print available at PsyArXiv: https://osf.io/preprints/psyarxiv/fevds_v1.

A fourth project, closely related to the themes developed throughout this thesis, is still ongoing and therefore could not be fully included here:

- **Ruzza, C.**, Isbilen, E. S., & Aslin, R. N. *Listening to Read: The Relationship between Auditory and Visual Statistical Learning and Literacy.* Data collection and analyses are currently in progress.

Several additional collaborations also shaped the ideas developed in these pages, even when their results were not directly incorporated into the thesis. In particular, discussions arising from the collaborative project greatly influenced my thinking about measurement, individual differences, and the theoretical interpretation of statistical learning:

- Bogaerts, L., Crepaldi, D., Zhou, H., **Ruzza, C.**, et al., & Siegelman, N. *Time for Consensus: Open Challenges and Recommendations for Research on Individual Differences in Statistical Learning.* Manuscript currently under revision for publication in *Perspectives on Psychological Science*. Pre-print available at PsyArXiv: https://osf.io/preprints/psyarxiv/eh7fm_v1.

Other projects to which I contributed during these years did not find a place within the present thesis, either because they addressed different research questions or because they remain ongoing:

- Marjanović, K., Vidal, Y., **Ruzza, C.**, & Crepaldi, D. *Oculomotor and Neural Correlates of Semantic and Morphological Priming in Natural Sentence Reading.* Manuscript currently under revision. Pre-print available at: https://doi.org/10.1101/2022.08.31.506138.

- Isbilen, E. S., **Ruzza, C.**, & Aslin, R. N. *Cross-Modal Statistical Learning and Neural Entrainment in Young Readers: Links to Phonological Skills and Literacy.* Currently in data collection.

- **Ruzza, C.**, Isbilen, E. S., & Aslin, R. N. *Spatiotemporal Mapping of Cross-Modal Learning: Evidence from EEG and OP-MEG.* Currently in the piloting stage.

Although only some of these projects appear in the chapters that follow, each contributed in some way to the development of the questions, methods, and theoretical perspectives presented here. Together, they reflect my attempt to understand how learning unfolds across modalities, across the lifespan, and across different levels of neural and behavioural expression.

This thesis is therefore not only a collection of studies, but also a record of an intellectual journey shaped by mentors, collaborators, and participants who generously contributed their time and expertise. I am deeply grateful to all of them.

Contents

1. General Introduction **29**
1.1 Statistical Learning as a Unitary or Multi-Componential Ability? 30
1.2 Cross-Modal Statistical Learning and Multisensory Integration 33
1.3 Individual Differences and the Measurement Problem in Statistical Learning 35
1.4 Neural Entrainment and Frequency-Tagging Approaches to Statistical Learning 37
1.5 Statistical Learning Across the Lifespan 40
1.6 The Present Thesis 43

2. Neural Sensitivity without Behavioural Evidence: Frequency-Based Statistical Learning in Fast Periodic Visual Stimulation **47**

2.1 Introduction 49

2.2 Experiment 1 55
2.2.1 Materials and Methods 55
2.2.1.1 Participants 55
2.2.1.2 Stimuli 55
2.2.1.3 Experimental Paradigm and Procedure 56
2.2.1.4 EEG Data Acquisition and Pre-processing 59
2.2.1.5 Data Analysis 60
2.2.2 Results 61
2.2.2.1 FPVS Control Task 61
2.2.2.2 FPVS Experimental Task 62
2.2.2.3 Memory Task 63
2.2.2.4 Discrimination Task 64
2.2.2.5 Familiarity Task 66
2.2.3 Discussion 67

2.3 Experiment 2 69
2.3.1 Materials and Methods 69

2.3.1.1 Participants . . . 69
2.3.1.2 Stimuli and Analyses . . . 69
2.3.2 Results . . . 70
2.3.2.1 FPVS Control Task . . . 70
2.3.2.2 FPVS Experimental Task . . . 70
2.3.2.3 Memory Task . . . 72
2.3.2.4 Discrimination Task . . . 72
2.3.2.5 Familiarity Task . . . 73
2.3.3 Discussion . . . 74

2.4 General Discussion . . . 77

2.5 Conclusion . . . 81

3. Neural and Behavioural Signatures of Statistical Learning in Aging: Evidence from Fast Periodic Visual Stimulation . . . 83

3.1 Introduction . . . 85

3.2 Materials and Methods . . . 93
3.2.1 Transparency and Openness . . . 93
3.2.2 Participants . . . 93
3.2.3 Stimuli . . . 94
3.2.4 Experimental Paradigm and Procedure . . . 94
3.2.5 EEG Data Acquisition and Pre-processing . . . 99
3.2.6 Data Analysis . . . 99

3.3 Results . . . 101
3.3.1 Group-Level Analyses . . . 101
3.3.1.1 FPVS Control Task . . . 102
3.3.1.2 FPVS Experimental Task . . . 103
3.3.1.3 Discrimination Task . . . 106
3.3.1.4 Familiarity Task . . . 108
3.3.1.5 VSL Task . . . 109
3.3.2 Reliability Analyses . . . 110
3.3.2.1 FPVS Tasks . . . 110
3.3.2.2 Behavioural Tasks . . . 111
3.3.3 Correlation Analyses . . . 111

3.3.3.1 Age-Related Effects . . . 112
3.3.3.2 Neural–Behavioural Relations . . . 114
3.3.3.3 Behavioural–Task Relations . . . 115

3.4 General Discussion . . . 117

3.5 Conclusion . . . 123

4. Cross-Modal Transfer of Statistical Learning: Evidence from auditory and visual sequence learning . . . 124

4.1 Introduction . . . 126

4.2 Materials and Methods . . . 133
4.2.1 Participants . . . 133
4.2.2 Stimuli . . . 134
4.2.3 Experimental Paradigm and Procedure . . . 134
4.2.4 Behavioral Data Analysis . . . 136

4.3 Results . . . 138
4.3.1 Transfer Effects . . . 138
4.3.2 Modality Directionality Effects . . . 140
4.3.3 Exposure Order Effects . . . 142
4.3.4 Interaction Effects . . . 144

4.4 General Discussion . . . 148

4.5 Conclusion . . . 157

5. Listening to Read: the relationship between Auditory and Visual Statistical Learning and Literacy . . . 162

5.1 Introduction . . . 164

5.2 Materials and Methods . . . 172
5.2.1 Participants . . . 172
5.2.2 Stimuli . . . 173
5.2.3 Experimental Paradigm and Procedure . . . 174

5.2.4 Behavioral Data Analysis . . . 176

5.3 Results . . . 177
5.3.1 Practice Task . . . 177
5.3.2 Letter Identification . . . 177
5.3.3 Rating Task . . . 178
5.3.4 TOWRE-2 Performance . . . 181
5.3.5 Relations Between Age, Reading Ability, and Statistical Learning 183

5.4 Discussion . . . 188

5.5 Conclusion . . . 195

6. General Discussion . . . 198

6.1 Empirical Patterns Across Modalities, Measures, and Age . . . 200

6.2 Theoretical and Methodological Implications . . . 208

6.3 Toward a Multi-Level Framework of Statistical Learning . . . 217
6.3.1 A Multi-Level Framework of Statistical Learning . . . 222

6.4 Limitations and Future Directions . . . 227

7. Conclusions . . . 230

Supplementary Materials . . . 258

Chapter 1

General Introduction

Statistical learning (SL) refers to the ability to extract regularities from structured sensory input through mere exposure (Aslin, 2017), enabling learners to detect patterns such as transitional probabilities, co-occurrence statistics, and distributional relationships among stimuli (Frost et al., 2019). This ability is considered a fundamental mechanism through which humans acquire knowledge about the structure of their environment, often incidentally and without explicit instruction. Across development, individuals continuously encounter highly structured sensory streams containing regularities that unfold across time and space, and SL has been proposed as one of the core processes supporting adaptation to such environments.

Research on SL emerged primarily from studies of language acquisition. Seminal work demonstrated that infants can segment continuous streams of speech into word-like units by tracking transitional probabilities between adjacent syllables (Saffran et al., 1996a). These findings suggested that learners can acquire meaningful structure directly from the statistical properties of sensory input, even in the absence of explicit cues or instruction. Since then, SL has been widely investigated as a mechanism contributing to early language acquisition and cognitive development. Subsequent work extended these findings beyond speech to a broad range of sensory modalities and stimulus types. Evidence for SL has been observed across auditory, visual, tactile, and motor domains using both linguistic and non-linguistic materials (Frost et al., 2019). Such findings led to the proposal that SL may reflect a domain-general computational mechanism capable of operating across multiple forms of structured input. At the same time, accumulating evidence suggests that learning outcomes may differ substantially depending on the modality, representational format, and task demands involved.

Some of the theoretical issues discussed in this chapter were also addressed during a workshop that led to *"Time for Consensus: Open Challenges and Recommendations for Research on Individual Differences in Statistical Learning"* manuscript, currently under revision for publication in *Perspectives on Psychological Science*. The pre-print can be found on PsyArXiv at https://osf.io/preprints/psyarxiv/eh7fm_v1.

Within language research, SL has been linked to multiple aspects of linguistic development and processing, including speech segmentation, vocabulary acquisition, phonological learning, syntactic processing, and reading development (Kidd and Arciuli, 2015; Pavlidou and Bogaerts, 2019). Written language itself contains rich statistical structure, including regularities in letter co-occurrences, grapheme–phoneme correspondences, and orthographic patterns, which learners gradually acquire through exposure and experience. From this perspective, reading acquisition can be understood partly as the extraction and integration of probabilistic regularities across auditory and visual modalities.

Importantly, SL appears early in development and remains observable throughout the lifespan. Infants, children, and adults all demonstrate sensitivity to structured regularities, although the magnitude and nature of learning may vary depending on developmental stage, sensory modality, and cognitive demands (Saffran and Kirkham, 2017; Frost et al., 2019). Furthermore, growing evidence suggests that prior linguistic knowledge and experience shape SL performance itself, indicating that learning is influenced not only by incoming statistical structure but also by pre-existing representations and expectations (Siegelman et al., 2018).

At the same time, the apparent ubiquity of SL has generated important theoretical questions regarding the nature of the construct itself. Although robust learning effects have been repeatedly demonstrated at the group level, increasing evidence suggests substantial variability across individuals, tasks, and modalities (Frost et al., 2019). Moreover, findings across paradigms often show weak convergence, raising questions about whether SL reflects a single domain-general ability or instead a collection of partially distinct learning mechanisms shaped by modality-specific and task-specific constraints. Understanding how SL varies across individuals, modalities, and developmental stages has therefore become a central challenge in contemporary SL research.

1.1 Statistical Learning as a Unitary or Multi-Componential Ability?

Although SL has often been described as a domain-general mechanism for extracting regularities from structured input, growing evidence suggests that learning performance varies substantially across modalities, tasks, and individuals (Frost et al., 2019; Siegelman et al., 2018). This variability has generated an important theoretical debate regarding the nature of SL itself: namely, whether SL reflects a single unified cognitive ability or instead a collection of partially distinct learning mechanisms shaped by modality-specific and task-specific constraints.

Early demonstrations of SL across auditory, visual, and motor domains supported the idea of a domain-general learning system capable of operating over different forms of sensory input (Conway and Christiansen, 2005; Kirkham et al., 2002; Saffran et al., 1999). According to this view, learners rely on shared computational principles to extract regularities regardless of stimulus modality or representational format. Such proposals were theoretically appealing because they offered a parsimonious account of how humans adapt to the structured nature of their environments across multiple domains of cognition.

However, more recent work has challenged the assumption that SL constitutes a single coherent ability. Studies examining correlations across SL tasks frequently report weak or inconsistent associations, even when tasks are designed to measure similar forms of learning (Siegelman et al., 2017a; Frost et al., 2019). In many cases, individuals who perform well on one SL task do not necessarily perform well on another, particularly when tasks differ in modality or stimulus characteristics. These findings have raised concerns regarding both the construct validity of SL tasks and the possibility that different paradigms rely on partially distinct cognitive and perceptual processes.

One influential account proposed by Frost et al. (2015) argues that SL should not be conceptualized as a single domain-general mechanism, but rather as a family of learning processes that share common computational principles while remaining constrained by modality-specific representational systems. According to this framework, statistical computations may operate differently depending on the sensory modality, temporal structure, attentional demands, and prior experience associated with the input. Auditory processing, for example, is particularly suited for extracting temporal and sequential regularities, whereas visual processing may preferentially support spatial or simultaneous structure (Conway and Christiansen, 2005; Zhou et al., 2024). From this perspective, learning outcomes emerge through interactions between general learning computations and modality-specific perceptual constraints.

Evidence for modality-specificity has been observed across both developmental and adult populations. Studies comparing auditory and visual SL often report limited convergence between modalities despite above-chance learning within each domain (Pavlidou and Bogaerts, 2019). Similarly, developmental studies suggest that auditory and visual learning may follow distinct trajectories across childhood. For example, Raviv and Arnon (2018) reported relatively stable auditory learning across development but age-related improvements in visual learning, suggesting that modality-specific factors may influence how SL evolves over time. Other work further indicates that visual and auditory SL may relate differently to language and literacy outcomes (Qi et al., 2019; Pavlidou and Bogaerts, 2019).

An additional challenge to unitary accounts of SL comes from research examining the influence of prior knowledge and linguistic experience on learning performance. Rather than reflecting purely implicit extraction of environmental statistics, SL appears strongly shaped by existing representations and expectations (Siegelman et al., 2018). Learners often show enhanced performance for statistically structured input that aligns with previously acquired linguistic or perceptual knowledge, suggesting that SL depends not only on online statistical computations but also on interactions with long-term memory and experience-dependent representations. These findings complicate traditional views of SL as a fully domain-general and context-independent mechanism.

At the same time, evidence supporting some degree of shared processing across modalities continues to exist. Neuroimaging and behavioral studies have identified partially overlapping neural networks involved in extracting structured regularities across sensory domains (Conway and Christiansen, 2005; Frost et al., 2019). Moreover, learners are often capable of transferring statistical knowledge across tasks and modalities under at least some conditions, suggesting that common learning principles may still operate despite modality-specific constraints. Consequently, contemporary perspectives increasingly propose hybrid accounts in which SL reflects interactions between shared computational mechanisms and modality-dependent representational systems rather than a strictly unitary or fully fragmented construct.

Importantly, these theoretical debates have substantial implications for interpreting developmental differences, individual variability, and cross-modal learning effects. If SL reflects multiple partially distinct processes, then variability across modalities or tasks may be theoretically meaningful rather than simply reflecting measurement noise. Understanding whether statistical regularities generalize across sensory modalities therefore provides an important avenue for investigating the structure and organization of SL itself. These questions become particularly relevant in contexts requiring integration between modalities, such as reading acquisition, where learners must connect auditory phonological representations with visual orthographic forms.

1.2 Cross-Modal Statistical Learning and Multisensory Integration

Natural environments are inherently multimodal. Speech is typically accompanied by visual articulatory information, written language requires mappings between orthographic and phonological representations, and everyday perception depends on the integration of information across sensory systems. Consequently, understanding SL requires not only examining learning within isolated modalities, but also investigating how statistical regularities are integrated and generalized across modalities (Conway and Christiansen, 2005; Frost et al., 2015).

A substantial body of research has demonstrated that learners can extract regularities independently within both auditory and visual domains. In the auditory modality, SL has been linked to speech segmentation, phonological learning, and language acquisition (Saffran and Kirkham, 2017; Conway and Pisoni, 2008). Auditory input unfolds sequentially over time, making sensitivity to temporal transitional probabilities particularly important for identifying recurring linguistic units. In contrast, visual SL has often been associated with the extraction of spatial and sequential regularities from streams of shapes, symbols, or orthographic stimuli (Kirkham et al., 2002; Fiser and Aslin, 2002). Although both modalities support statistical extraction, accumulating evidence suggests that learning may be shaped by modality-specific perceptual and representational constraints.

Comparisons between auditory and visual SL have produced mixed findings regarding the extent to which learning mechanisms generalize across modalities. Some studies report evidence consistent with shared computational principles operating across sensory systems (Conway and Christiansen, 2005), whereas others find weak correlations between auditory and visual learning performance, suggesting partial modality specificity (Pavlidou and Bogaerts, 2019; Zhou et al., 2024). For example, Pavlidou and Bogaerts (2019) observed above-chance learning in auditory, visual, and tactile modalities in children, but found limited correlations across modalities despite significant relations with literacy measures. Such findings suggest that successful learning within one modality does not necessarily predict learning in another, challenging strongly domain-general accounts of SL.

Cross-modal transfer paradigms provide a particularly informative framework for investigating the structure of SL because they directly assess whether regularities acquired in one sensory modality can generalize to another. This issue is especially relevant for reading acquisition, which requires learners to connect auditory phonological representations with visual orthographic forms. Efficient reading therefore depends not only on learning regularities within modalities independently, but also

on integrating information across modalities (Qi et al., 2019; Garcia et al., 2019).

Several studies suggest that children and adults can generalize statistical structure across modalities under some conditions. Research on orthographic learning indicates that exposure to spoken forms can shape expectations regarding written representations (Beyersmann et al., 2023). Similarly, audiovisual associative learning paradigms demonstrate that successful reading development depends on increasingly efficient mappings between phonological and orthographic information (Garcia et al., 2019). More broadly, studies examining multimodal learning environments suggest that learners can integrate statistical information across modalities when regularities are temporally or structurally aligned (Conway and Christiansen, 2005).

At the same time, evidence for robust cross-modal transfer remains inconsistent. Some studies report relatively weak or absent transfer effects, particularly when tasks involve different representational formats or increased cognitive demands (Mitchel and Weiss, 2011b; Pavlidou and Bogaerts, 2019). Evidence further suggests that transfer may be asymmetrical across modalities. For example, visual-to-auditory transfer may emerge more readily than auditory-to-visual transfer in some contexts, potentially reflecting differences in sensory dominance, temporal processing, or representational complexity (Batikh et al., 2025). These findings indicate that successful learning within individual modalities does not necessarily guarantee efficient cross-modal generalization.

One possible explanation for these mixed findings is that cross-modal SL depends on interactions between shared learning computations and modality-specific representational systems (Frost et al., 2015). From this perspective, learners may extract abstract regularities that can partially generalize across modalities, while still remaining constrained by sensory-specific processing characteristics. Auditory learning, for instance, may preferentially encode temporal and phonological structure, whereas visual learning may rely more strongly on spatial or orthographic representations. Cross-modal transfer would therefore require not only statistical extraction itself, but also successful transformation or alignment between modality-specific representations.

Cross-modal SL may also be influenced by the hierarchical and overlapping nature of environmental regularities. Real-world sensory input often contains multiple simultaneous structures operating at different levels of abstraction, requiring learners to integrate competing or complementary cues across modalities (Frost et al., 2019). In language, for example, learners must simultaneously process phonological, orthographic, semantic, and syntactic regularities. Such environments may place substantial demands on attention, working memory, and representational flexibility, particularly during development or aging.

Importantly, methodological factors may contribute substantially to inconsistencies in cross-modal SL literature. Transfer paradigms often differ in stimulus complexity, exposure duration, response requirements, and measurement sensitivity, making findings difficult to compare directly (Siegelman et al., 2017a; Frost et al., 2019). In developmental populations, these challenges may be amplified by variability in attention, executive functioning, and metacognitive abilities. Furthermore, many traditional behavioral paradigms rely on offline familiarity judgments or forced-choice procedures that may underestimate implicit learning processes (Isbilen and Christiansen, 2022; Batterink et al., 2019). These limitations have motivated increasing interest in multimethod approaches capable of capturing online learning dynamics more directly.

Overall, current evidence suggests that SL involves both modality-general and modality-specific components, but the extent to which statistical knowledge generalizes across sensory systems remains incompletely understood. Investigating cross-modal transfer may therefore provide important insight into the architecture of SL itself, while also clarifying how learners integrate auditory and visual information during complex cognitive processes such as reading and language acquisition.

1.3 Individual Differences and the Measurement Problem in Statistical Learning

Although SL has traditionally been studied through robust group-level learning effects, increasing attention has been devoted to understanding variability across individuals. Early SL research primarily focused on demonstrating that learners could reliably extract regularities from structured input under controlled laboratory conditions (Saffran and Kirkham, 2017; Kirkham et al., 2002). However, more recent work has emphasized that individuals differ substantially in their learning performance across tasks, modalities, and contexts (Frost et al., 2019; Bogaerts et al., 2022a). These individual differences have become theoretically important because they may provide insight into the cognitive architecture of SL, its developmental trajectories, and its relation to language and literacy abilities.

Research examining individual variability has produced mixed findings. Several studies report significant associations between SL and language-related outcomes, including vocabulary acquisition, syntactic processing, reading ability, and phonological awareness (Kidd and Arciuli, 2015; Arciuli and Simpson, 2012; Qi et al., 2019; Pavlidou and Bogaerts, 2019). For example, Kidd and Arciuli (2015) found that children with stronger SL abilities demonstrated better syntactic comprehension, suggesting that sensitivity to sequential regularities may contribute to language

development. Similarly, studies examining literacy have linked both auditory and visual SL to reading performance and orthographic processing (Arciuli and Simpson, 2012; Pavlidou and Bogaerts, 2019). Such findings support the idea that SL may contribute meaningfully to individual differences in cognitive and linguistic abilities.

At the same time, evidence linking SL measures across tasks and modalities has often been surprisingly weak. Correlations between different SL paradigms are frequently small or inconsistent, even when tasks are intended to measure similar underlying computations (Siegelman et al., 2017a; Frost et al., 2019). In some cases, individuals who exhibit strong learning in one modality show limited learning in another, while different tasks within the same modality may also fail to converge reliably. These findings have raised concerns regarding whether commonly used SL tasks measure a coherent underlying construct or instead capture multiple partially distinct processes influenced by modality-specific and task-specific demands.

One major challenge concerns the psychometric reliability of SL measures. Many canonical SL paradigms were originally designed to demonstrate learning effects at the group level rather than to measure stable individual differences (Hedge et al., 2018). As a consequence, tasks that produce robust average effects may nevertheless show poor reliability when used to assess variability across participants. This issue is particularly problematic for correlational and developmental research because unreliable measures substantially reduce the ability to detect meaningful associations with other cognitive abilities (Siegelman et al., 2017a; Bogaerts et al., 2022a).

Additional concerns involve the substantial heterogeneity across SL paradigms. Tasks vary widely in stimulus modality, complexity, exposure duration, attentional demands, memory requirements, and response procedures (Frost et al., 2019). Some paradigms rely on offline familiarity judgments or forced-choice tests administered after exposure, whereas others use online measures such as reaction times, eye movements, or neural entrainment during learning itself. These methodological differences may strongly influence observed learning effects and may engage partially distinct cognitive mechanisms, complicating interpretation and comparison across studies.

The distinction between implicit and explicit learning processes further contributes to the measurement problem in SL research. Although SL is often conceptualized as an implicit process, behavioral tasks may capture a mixture of implicit sensitivity, explicit awareness, decision strategies, and metacognitive factors (Batterink et al., 2019). For example, learners may acquire statistical regularities without being able to explicitly report them, while offline judgment tasks may depend heavily on memory retrieval, response bias, or conscious reflection. Consequently, poor performance on behavioral measures does not necessarily imply absence of learning, and conversely, successful task performance may reflect processes extending beyond

implicit statistical extraction itself.

These issues may be especially pronounced in developmental and aging populations. Children and older adults often show increased variability in attention, working memory, processing speed, and executive functioning, all of which may influence performance on behavioral SL tasks independently of learning ability (Arnon, 2020; Raviv and Arnon, 2018). Developmental differences observed across studies may therefore reflect not only changes in statistical extraction itself, but also broader cognitive and perceptual factors associated with task performance. Similarly, sensory decline and age-related cognitive changes may differentially affect auditory and visual learning measures in older adulthood.

In response to these challenges, researchers have increasingly advocated for multimethod approaches combining behavioral and neural measures of SL (Frost et al., 2019; Batterink et al., 2019). Neural indices may provide more direct access to online learning dynamics while reducing some of the decisional and metacognitive demands associated with traditional behavioral paradigms. In particular, electrophysiological approaches such as neural entrainment and frequency-tagging paradigms have emerged as promising tools for investigating implicit statistical processing across modalities and populations.

Overall, the growing focus on individual differences has highlighted important conceptual and methodological limitations within SL research. Rather than representing mere measurement noise, variability across individuals, tasks, and modalities may reflect meaningful differences in learning mechanisms, representational systems, and cognitive constraints. Understanding these sources of variability therefore remains essential for clarifying the nature of SL and its role in language, reading, and cognition across the lifespan.

1.4 Neural Entrainment and Frequency-Tagging Approaches to Statistical Learning

Traditional behavioral paradigms have played a central role in establishing the existence of SL across modalities and developmental stages. However, growing concerns regarding the reliability, sensitivity, and interpretability of behavioral measures have motivated increasing interest in neural approaches capable of capturing learning processes more directly (Siegelman et al., 2017a; Batterink et al., 2019; Frost et al., 2019). In particular, electrophysiological methods provide the opportunity to investigate online learning dynamics as statistical regularities are extracted from continuous sensory input, potentially offering insight into mechanisms that remain difficult to observe through offline behavioral judgments alone.

One influential approach involves neural entrainment, broadly defined as the synchronization of neural activity to rhythmic or temporally structured sensory input (Ding et al., 2016). During exposure to structured auditory or visual streams, neural oscillations can become aligned not only with low-level sensory rhythms, but also with higher-order regularities embedded within the input. In the context of SL, this phenomenon has been interpreted as reflecting the emergence of internal representations corresponding to statistically coherent units such as words, sequences, or chunks (Buiatti et al., 2009; Batterink and Paller, 2017).

Frequency-tagging paradigms, including Fast Periodic Visual Stimulation (FPVS) approaches, have become particularly important tools for investigating SL and perceptual learning. In these paradigms, stimuli are presented at fixed temporal frequencies, allowing researchers to isolate neural responses at frequencies corresponding to both sensory presentation rates and higher-order statistical structure (Rossion, 2014). Because frequency-tagged responses can be measured objectively in the electroencephalography (EEG) signal, these approaches provide highly sensitive and implicit indices of learning that do not depend on explicit reports or metacognitive awareness.

Evidence from auditory SL studies suggests that neural entrainment can track the online segmentation of continuous speech streams into statistically coherent units. For example, Batterink and Paller (2017) demonstrated that neural activity synchronizes not only to syllable presentation rates but also to frequencies corresponding to embedded word structure, indicating rapid extraction of transitional regularities during exposure. Similar findings have been reported in infants and adults, suggesting that neural synchronization to statistical structure emerges early and may reflect a core mechanism supporting language acquisition (Buiatti et al., 2009; Choi et al., 2020).

Neural entrainment approaches have also been extended to the visual modality. FPVS paradigms allow researchers to investigate sensitivity to visual regularities by embedding structured patterns within rapidly presented streams of visual stimuli (Rossion, 2014). Such paradigms have been used to examine orthographic processing, visual word recognition, and category learning while minimizing task demands and explicit response requirements. Importantly, visual frequency-tagging approaches may provide particularly valuable insight into orthographic and reading-related processing because they enable the investigation of automatic neural discrimination responses to structured visual input.

A major advantage of neural entrainment methods is their ability to capture implicit and online aspects of learning that may not be reflected in behavioral performance. Behavioral SL tasks often rely on delayed familiarity judgments, forced-

choice decisions, or explicit recognition measures, all of which may be influenced by attention, memory retrieval, decision strategies, and response bias (Batterink et al., 2019). In contrast, neural frequency-tagged responses can reveal sensitivity to statistical structure even when learners are unable to explicitly report acquired regularities. This distinction may be especially important in populations such as children or older adults, where metacognitive demands and response variability can complicate interpretation of behavioral measures.

Neural approaches may also help clarify theoretical debates regarding the nature of SL itself. For example, neural entrainment paradigms can be used to examine whether statistical regularities are represented similarly across modalities or whether modality-specific neural dynamics emerge during learning (Ding et al., 2016). Likewise, frequency-tagging approaches provide opportunities to investigate cross-modal transfer and hierarchical structure processing while dissociating sensory-driven responses from higher-order statistical representations.

Importantly, neural and behavioral measures do not always converge strongly (Batterink et al., 2019). Some studies report neural evidence of learning in the absence of robust behavioral effects, whereas others observe behavioral discrimination without clear neural signatures. These discrepancies suggest that behavioral and neural measures may capture partially distinct aspects of learning, including implicit sensitivity, explicit awareness, consolidation processes, or decisional components. Consequently, combining neural and behavioral approaches may provide a more comprehensive characterization of SL than either method alone.

Neural entrainment paradigms additionally offer several methodological advantages for lifespan research. Because frequency-tagged responses can be obtained passively with relatively limited task demands, these methods are particularly well suited for developmental and aging populations. In children, neural measures may provide sensitive indices of emerging statistical representations before stable behavioral responses can be reliably obtained. In older adults, neural entrainment approaches may help dissociate preserved implicit learning mechanisms from age-related declines in explicit memory or executive functioning.

Despite these advantages, important questions remain regarding the interpretation and specificity of neural entrainment effects. Ongoing debates concern the extent to which frequency-tagged responses reflect genuine statistical learning as opposed to lower-level sensory adaptation, attentional modulation, or acoustic and perceptual regularities inherent in the stimulus stream (Moreau et al., 2022). Furthermore, the relationship between neural synchronization and long-term learning outcomes remains incompletely understood. Continued integration of behavioral, developmental, and neural approaches will therefore be essential for clarifying the

mechanisms underlying entrainment-based measures of SL.

Overall, neural entrainment and FPVS paradigms provide powerful complementary tools for investigating SL across modalities and across the lifespan. By capturing online neural sensitivity to structured regularities, these approaches offer important opportunities to address longstanding theoretical and methodological challenges within SL research while advancing understanding of how statistical structure is encoded, generalized, and represented in the human brain.

1.5 Statistical Learning Across the Lifespan

SL emerges early in development and remains observable throughout adulthood, suggesting that sensitivity to environmental regularities represents a fundamental property of human cognition (Saffran and Kirkham, 2017; Frost et al., 2019). At the same time, growing evidence indicates that the mechanisms supporting SL may change across the lifespan as a function of perceptual maturation, cognitive development, accumulated experience, and age-related decline. Understanding how SL evolves across different stages of life is therefore essential for clarifying both the architecture of learning systems and their role in language, reading, and cognition.

Research in infancy and childhood has provided some of the strongest evidence for early-emerging SL abilities. Infants demonstrate sensitivity to statistical regularities in speech streams within the first year of life, suggesting that SL may contribute to early language acquisition before the emergence of extensive lexical knowledge (Saffran and Kirkham, 2017). Subsequent work extended these findings to visual and non-linguistic auditory domains, indicating that children can extract structured regularities across multiple forms of sensory input (Kirkham et al., 2002; Fiser and Aslin, 2002). Such findings have led to proposals that SL constitutes one of the foundational mechanisms supporting cognitive and linguistic development.

Developmental studies further suggest that SL may play an important role in literacy acquisition and language learning during childhood. Sensitivity to sequential regularities has been associated with phonological processing, syntactic comprehension, vocabulary development, and reading ability (Kidd and Arciuli, 2015; Arciuli and Simpson, 2012; Qi et al., 2019). In particular, reading acquisition requires children to integrate auditory phonological representations with visual orthographic forms, making cross-modal learning especially relevant during this developmental period. Evidence suggests that both auditory and visual SL contribute to literacy-related skills, although the relative contribution of each modality may differ depending on task demands and developmental stage (Pavlidou and Bogaerts, 2019).

At the same time, developmental trajectories appear to vary across modalities.

Several studies suggest that auditory SL remains relatively stable across childhood, whereas visual SL may show more pronounced developmental improvement (Raviv and Arnon, 2018). These modality-specific trajectories may reflect differences in perceptual maturation, attentional control, reading experience, or the increasing integration of phonological and orthographic systems during literacy acquisition. However, developmental findings remain mixed, and concerns regarding task reliability and measurement validity complicate interpretation of age-related differences (Arnon, 2020; Frost et al., 2019).

Importantly, developmental changes in SL may not solely reflect maturation of statistical extraction mechanisms themselves. Age-related improvements in attention, working memory, executive functioning, processing speed, and metacognitive abilities may substantially influence performance on behavioral SL tasks (Arnon, 2020). Consequently, apparent developmental differences may partly arise from broader cognitive changes affecting task engagement, memory retrieval, or response strategies rather than from changes in implicit learning capacity alone. These issues have motivated increasing interest in neural and online measures capable of assessing learning more directly across development.

In adulthood, SL is generally considered relatively robust and stable, although substantial individual variability continues to be observed (Frost et al., 2019). Adults demonstrate reliable learning across auditory and visual modalities as well as sensitivity to complex sequential and hierarchical regularities. At the same time, adult learning is strongly influenced by prior linguistic knowledge and accumulated experience (Siegelman et al., 2018). Existing representations and expectations can facilitate or constrain statistical extraction processes, indicating that SL in adulthood operates within the context of extensive long-term knowledge structures.

Research in older adulthood has produced more mixed findings regarding the preservation of SL abilities. Some studies suggest relatively preserved implicit learning despite age-related declines in explicit memory and executive functioning, whereas others report reduced sensitivity to complex sequential regularities, particularly under conditions involving increased attentional or perceptual demands (Neger et al., 2014). Age-related sensory decline may further differentially affect auditory and visual learning, potentially altering how statistical structure is encoded and integrated across modalities.

Neural evidence additionally suggests that aging may influence the dynamics of entrainment and synchronization to structured sensory input. Changes in neural oscillatory activity, processing speed, attentional control, and sensory precision may affect the efficiency with which older adults track statistical regularities within continuous streams (Henry et al., 2017). At the same time, older adults may rely more

heavily on prior knowledge and predictive processing to compensate for declines in sensory or cognitive functioning. Such compensatory mechanisms may produce complex patterns of preserved and impaired learning depending on task structure and modality.

Across the lifespan, modality-specific constraints appear to interact with broader developmental and cognitive factors. Children may rely more strongly on auditory sequential information during language acquisition, whereas visual learning becomes increasingly important with literacy development and reading experience. In older adulthood, sensory decline and cognitive slowing may differentially influence auditory and visual processing, potentially altering patterns of cross-modal integration and transfer. Understanding these lifespan changes therefore requires considering not only learning mechanisms themselves, but also the evolving perceptual, cognitive, and experiential systems within which learning occurs.

Overall, current evidence suggests that SL reflects a dynamic set of processes that interact with development, aging, prior experience, and modality-specific constraints across the lifespan. However, many important questions remain unresolved regarding how learning mechanisms evolve across different stages of life and how these changes influence cross-modal integration, literacy, and neural processing. Integrating behavioral and neural approaches across developmental and aging populations may therefore provide important insight into the organization and adaptability of SL systems throughout the human lifespan.

1.6 The Present Thesis

The research reviewed thus far suggests that SL represents a fundamental mechanism through which humans extract regularities from structured sensory input. At the same time, substantial theoretical and methodological questions remain unresolved regarding the nature of SL, its organization across modalities, its variability across individuals, and its development across the lifespan. Although robust learning effects have been repeatedly demonstrated at the group level, evidence increasingly suggests that SL may not reflect a single unified cognitive ability. Instead, learning appears shaped by interactions between modality-specific perceptual systems, prior knowledge, developmental constraints, and task-related factors (Frost et al., 2015, 2019; Siegelman et al., 2017a).

One major unresolved issue concerns the extent to which statistical knowledge generalizes across sensory modalities. While auditory and visual SL have each been extensively investigated independently, relatively few studies have directly examined cross-modal transfer and multimodal integration processes. This question is particularly relevant for language and literacy acquisition, which require learners to integrate auditory phonological representations with visual orthographic information. Existing findings regarding cross-modal transfer remain mixed, with some studies suggesting partial generalization across modalities and others reporting weak or inconsistent transfer effects (Pavlidou and Bogaerts, 2019; Qi et al., 2019; Conway and Christiansen, 2005). Clarifying the extent to which learning mechanisms are shared or modality-specific therefore remains central for understanding the architecture of SL itself.

A second challenge concerns the measurement of SL and the interpretation of individual variability. Behavioral SL paradigms often show limited reliability and weak convergence across tasks, raising concerns regarding construct validity and the extent to which different paradigms measure common underlying processes (Siegelman et al., 2017a; Hedge et al., 2018). Furthermore, traditional offline behavioral measures may capture not only implicit statistical sensitivity but also attentional, mnemonic, decisional, and metacognitive processes (Batterink et al., 2019). These methodological limitations complicate interpretation of developmental differences, modality effects, and correlations between SL and language-related abilities.

Recent advances in neural approaches, particularly neural entrainment and frequency-tagging paradigms, provide new opportunities to investigate SL more directly. Electrophysiological measures allow researchers to examine online synchronization to structured regularities during learning itself while minimizing some of the decisional demands associated with behavioral tasks (Ding et al., 2016; Batterink and Paller,

2017). Such approaches may be particularly valuable for developmental and aging populations, where behavioral variability and task demands can complicate assessment of implicit learning processes. At the same time, the relationship between behavioral and neural measures of SL remains incompletely understood, and relatively few studies have combined these approaches within multimodal and lifespan frameworks.

Several important questions remain unresolved within the SL literature. First, it is still unclear to what extent SL reflects shared domain-general mechanisms versus modality-specific learning systems. Although auditory and visual learning have each been robustly demonstrated, evidence regarding cross-modal transfer and convergence across tasks remains inconsistent (Frost et al., 2015; Siegelman et al., 2017a). Second, relatively little is known about how behavioral and neural measures of SL relate to one another, particularly across different sensory modalities and populations. Third, developmental and aging perspectives have rarely been integrated within the same research framework, limiting understanding of how SL changes across the lifespan and how these changes interact with language, literacy, and sensory processing. Finally, substantial individual variability is consistently observed across SL paradigms, yet the sources and implications of this variability remain poorly understood.

The present thesis addresses these questions through a series of behavioral and EEG studies investigating SL across auditory and visual modalities in children, young adults, and older adults. Across chapters, the thesis examines modality-specific and cross-modal learning, neural entrainment to structured input, developmental and aging effects, and relations between behavioral and neural indices of learning.

Chapter 2 investigates auditory SL in young adults using an EEG frequency-tagging paradigm. Participants were exposed to continuous artificial speech streams while neural entrainment at syllable and word frequencies was measured. Behavioral familiarity judgments were additionally collected following familiarization. Results demonstrated reliable neural tracking of statistically defined word structure together with behavioral evidence of learning, while also revealing substantial individual variability and only limited correspondence between neural and behavioral indices. These findings highlight both the sensitivity of frequency-tagging approaches and the complexity of interpreting individual differences in SL.

Chapter 3 extends this approach to healthy older adults in order to examine how aging influences auditory SL and neural entrainment. Using the same frequency-tagging paradigm, the study investigated whether older adults show preserved sensitivity to statistical structure despite age-related changes in cognition and sensory processing. Results indicated that older adults continued to exhibit neural entrain-

ment to structured auditory input, although behavioral learning effects and brain–behavior relations differed from those typically observed in younger adults. These findings suggest partial preservation of SL mechanisms in aging while also pointing to age-related changes in the efficiency or expression of learning processes.

Chapter 4 examines cross-modal SL and transfer in young adults using both auditory and visual behavioral paradigms. Participants were first familiarized with structured auditory streams and subsequently tested on both auditory and visually presented orthographic stimuli. The study investigated whether statistical regularities learned in the auditory modality generalized to visual representations. Results revealed reliable auditory learning together with weaker but significant auditory-to-visual transfer effects, suggesting partial cross-modal generalization of learned regularities. At the same time, substantial variability across participants supported the view that cross-modal transfer may depend on both shared and modality-specific mechanisms.

Chapter 5 investigates cross-modal SL and literacy-related processes in children during early reading development. Children completed auditory familiarization and auditory and visual judgment tasks together with standardized reading assessments. The study examined whether auditory regularities generalized to visual orthographic forms and whether learning performance related to literacy abilities. Results showed reliable auditory learning and weaker but directionally consistent visual transfer effects. Importantly, individual differences in auditory learning and visual transfer were positively associated with standardized reading ability, particularly word reading performance. These findings suggest that SL and cross-modal integration processes may contribute to literacy acquisition during childhood.

Chapter 6 integrates the findings from the empirical studies within broader theoretical and methodological debates surrounding SL. The discussion evaluates the implications of the present findings for questions concerning modality specificity, multicomponentiality, individual variability, neural–behavioral dissociations, and developmental change across the lifespan. Building on contemporary theoretical accounts, the chapter proposes a multi-level framework of SL in which different paradigms may access partially distinct representational and processing levels differing in accessibility to conscious report, processing demands, and neural implementation. More specifically, the discussion argues that neural entrainment measures may primarily index early online sensitivity to statistical regularities, whereas behavioral paradigms additionally recruit processes related to prediction, decision-making, memory retrieval, and metacognitive reflection. The chapter further addresses methodological challenges in SL research, including issues of task reliability, construct validity, and weak cross-task convergence, while outlining future directions for integrating neural

and behavioral approaches within multimodal and lifespan frameworks.

Together, the empirical chapters provide converging evidence that SL reflects an interaction between shared computational principles and modality-specific representational constraints. Across behavioral and neural paradigms, substantial individual variability emerged consistently, while developmental and aging findings suggested that SL remains relatively robust across the lifespan despite changes in sensory, cognitive, and linguistic processing. Moreover, the findings demonstrate that neural frequency-tagging approaches provide complementary insight into SL beyond traditional behavioral measures, particularly in populations where explicit behavioral responses may underestimate learning.

Overall, the thesis adopts the perspective that SL is neither fully domain-general nor entirely modality-specific, but instead emerges through dynamic interactions between perceptual systems, developmental processes, prior knowledge, and individual cognitive differences. By integrating multimodal behavioral paradigms with neural measures across developmental and aging populations, the present work aims to contribute to a more comprehensive understanding of how humans acquire, represent, and generalize statistical structure in complex sensory environments.

Chapter 2

Neural Sensitivity without Behavioural Evidence: Frequency-Based Statistical Learning in Fast Periodic Visual Stimulation

Claudia Ruzza[1]
Davide Crepaldi[2]

[1]Department of Cognitive Neuroscience
Scuola Internazionale Superiore di Studi Avanzati (SISSA)
Trieste, Italy

[2]Dipartimento di Scienze del Sistema Nervoso e del Comportamento
Università di Pavia
Pavia, Italy

This chapter is adapted from a manuscript currently under revision for publication in *Journal of Memory and Language* The pre-print can be found on PsyArXiv at https://osf.io/preprints/psyarxiv/n4fjr_v1.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant Agreement No 679010 STATLEARN ERC-2015-STG/ERC-2015-STG.

Abstract

Humans are sensitive to statistical cues in their perceptual input, including associations between stimuli (transitional probability) and their absolute frequency (unimodal information). To investigate the mechanisms behind the acquisition of this latter cue, both at an implicit and explicit level, we combined Fast Periodic Visual Stimulation (FPVS) with EEG recordings and two-alternative forced choice (2AFC) tasks. Skilled readers viewed rapid sequences (6 Hz) of words and pseudofonts, in which low-frequency items appeared regularly at a slower periodic rate (1.2 Hz) among high-frequency items. EEG frequency-domain analyses revealed robust neural responses at the oddball frequency, indicating sensitivity to frequency-based regularities and implicit grouping. Despite these clear signs of neural learning, no discrimination between low- and high-frequency items emerged in our participants' behaviour. This pattern of results was confirmed in a replication experiment. Together, these findings suggest that frequency-based regularities are robustly encoded at the neural level, even when behavioural expression remains limited. They also suggest that neural entrainment and behavioural learning might track different aspects of the complex neural and cognitive machinery that makes us sensitive to the statistics of our environment.

I would like to thank Maria Ktori for guidance in EEG testing; Mara De Rosa and Yamil Vidal for their help with experiment implementation and analysis coding; and Elena Greatti, Emma Peri and Emma Coba for their assistance with data collection.

2.1 Introduction

Statistical learning (SL) refers to the ability to implicitly and automatically extract regularities from structured input, such as the co-occurrence and distribution of elements over time or space.

This mechanism enables the discovery of meaningful units in continuous sensory streams without explicit instruction (Aslin, 2017; Frost et al., 2019). A classic demonstration comes from language research, showing that learners can rely solely on transitional probabilities (TPs) between syllables to segment continuous speech and identify word-like units (Saffran et al., 1996a). Similar abilities have been observed in the visual domain, where participants extract regularities from sequences of shapes or objects (Kirkham et al., 2002; Turk-Browne et al., 2005; Siegelman and Frost, 2015).

Beyond sensitivity to sequential co-occurrence, SL also supports the extraction of distributional structure from the input. In particular, learners can use frequency distributions to form categorical representations and establish category boundaries. For example, exposure to bimodal distributions of phonetic tokens leads to the formation of distinct phoneme categories, whereas unimodal distributions promote category merging (Maye et al., 2002). Analogous findings in the visual domain demonstrate that observers can derive categories from the statistical distribution of low-level features, such as orientation or spatial frequency (Rosenthal et al., 2001). These findings suggest that SL mechanisms operate not only over sequential dependencies, but also over distributional properties, supporting the emergence of structured representations. In the current study, we investigate whether such distributional information can elicit implicit SL within a Fast Periodic Visual Stimulation (FPVS) paradigm combined with electrophysiological recording.

Statistical Learning in Visual Word Processing

Visual word processing is increasingly understood as relying on statistical learning (SL). According to the Local Combination Detector (LCD) model, for example, visual recognition operates through a hierarchical organization of the ventral visual stream, in which neurons are progressively tuned to increasingly complex feature combinations based on exposure to environmental regularities (Dehaene et al., 2005). Within this framework, sensitivity to orthographic structure emerges from experience with the statistical co-occurrence of visual features (e.g., letters and letter clusters), leading to graded neural selectivity for familiar patterns along the occipitotemporal cortex (Vinckier et al., 2007).

Crucially, evidence from both human and non-human studies demonstrates that such sensitivity to orthographic regularities reflects a broader capacity for visual SL. Implicit learning of co-occurrence patterns is observed not only for letter-like symbols, but also for non-linguistic visual stimuli, such as artificial objects and sinusoidal gratings, where regularities are defined by feature pairings (Vidal et al., 2021). Similar sensitivities have also been documented in developing readers (Cassar and Treiman, 1997; Pacton et al., 2001), and in connection with phenomena that were always interpreted as genuinely linguistic in nature (e.g., morphology; Lelonkiewicz et al., 2020, 2023; Ktori et al., 2026).

Together, these findings support the view that visual word processing exploits a general-purpose capacity to extract and represent statistical regularities in the visual environment, positioning SL as a core mechanism underlying sensitivity to orthographic structure.

Fast Periodic Visual Stimulation as a Tool to Measure Statistical Learning

Building on the view that visual word processing relies on SL mechanisms, the Fast Periodic Visual Stimulation (FPVS) approach offers a powerful method to track such learning directly at the neural level. FPVS capitalizes on neural entrainment to periodic visual input (Norcia et al., 2015) and uses frequency tagging to isolate discrimination responses in the brain (Rossion, 2014). In a typical paradigm, base stimuli are presented at a constant rapid frequency (F), while oddball stimuli—differing along a dimension of interest are inserted at regular intervals, appearing at a lower frequency (F/n). This stimulation elicits steady-state visual evoked potentials (SSVEPs; Regan 1966), with distinct responses at the base and oddball frequencies and their harmonics. Crucially, the presence of a response at the oddball frequency reflects the brain's ability to discriminate between stimulus classes, providing an objective and implicit neural index of sensitivity to the relevant regularity, independent of task demands.

Thanks to its high sensitivity and task-free nature, FPVS has been widely used to investigate category-selective processing at the perceptual level, including faces, objects, and words, in both typical and atypical populations (Liu-Shuang et al., 2014; Lochy et al., 2015; Rossion et al., 2015; Jacques et al., 2016; Retter and Rossion, 2016; Lochy et al., 2016, 2018; Stacchi et al., 2019; Coll et al., 2019; Stothart and Smith, 2020; Van der Donck et al., 2020; Pescuma et al., 2022; De Rosa et al., 2022). Importantly, beyond capturing pre-existing category distinctions, the FPVS oddball design closely parallels paradigms used to study automatic novelty detection, such as mismatch negativity (Näätänen et al., 2007; Kimura et al., 2011; Stefanics et al.,

2015; Kremláček et al., 2016), where rare deviants are embedded within streams of standard stimuli. This structure provides a natural context for probing the emergence of sensitivity to statistical regularities, making FPVS particularly well suited to investigate SL in visual processing.

Taking advantage of this unique feature of the FPVS paradigm, scholars have started to use it to explicitly track Statistical Learning. In these designs, the slower frequency would not be marked by stimuli of a different sort (e.g., faces within a stream of objects, words within a stream of nonwords), thus assessing the brain's sensitivity to pre-existing categories. Rather, oddballs would be items of the same exact nature of any other stimulus in the stream (e.g., words in words), only marked statistically – either via a drop in transitional probability, like in the classic Saffran paradigm (e.g., Batterink and Paller, 2017), or simply by a reduced item frequency (e.g., De Rosa et al., 2022).

Neural Entrainment and Behavioural Measures: Convergence and Dissociation in Statistical Learning

Early entrainment studies would typically not include a behavioural validation; they mostly treated frequency-tagged responses as sufficient evidence of learning (Sun et al., 2017; De Rosa et al., 2022; David et al., 2023a; Lochy et al., 2024; De Rosa et al., 2024; Rutkowski et al., 2025; Ling et al., 2025). While methodologically powerful, this approach limits interpretation, as it remains unclear to what extent neural entrainment reflects learning that is behaviourally accessible. In response, an increasing number of studies have begun to directly examine the relationship between neural entrainment and behaviour by combining neural measures with independent behavioural tasks. In several cases, these measures converge.

Studies of face individuation and motor sequence learning report robust neural responses alongside strong behavioural performance, indicating successful learning of structured input (Dzhelyova et al., 2020; Moser et al., 2021b; Lum et al., 2023). Similarly, in FPVS adaptations targeting recognition memory, neural responses to previously encoded stimuli emerge during passive viewing, and participants subsequently show high accuracy in explicit recognition tasks (Stothart and Smith, 2020). Convergent findings have also been reported across domains, including face perception, object processing, and visual word recognition, where frequency-tagged responses index discrimination or categorisation processes that are mirrored in behavioural performance (Liu-Shuang et al., 2014; Lochy et al., 2015; Rossion et al., 2015; Liu-Shuang et al., 2016; Stothart et al., 2017; Sznabel et al., 2023; Stothart et al., 2025).

Despite these findings, other studies report weaker or no behavioural outcomes.

Some studies report neural responses to embedded regularities to emerge rapidly, whereas behavioural evidence of learning appears only after extended exposure or remains modest (Sáringer et al., 2024; Barillon et al., 2025); while other studies report clear dissociations, with robust neural entrainment in the absence of reliable behavioural learning, or substantial individual variability in the correspondence between the two measures (Rawal and Tseng, 2021; Chiarella et al., 2026).

Consistent with these mixed findings, correlations between neural and behavioural measures are highly variable. Strong correlations are observed in well-aligned paradigms—such as FPVS studies of faces and words, where neural responses closely predict behavioural performance (Dzhelyova et al., 2020; Moser et al., 2021b; Barillon et al., 2025). However, correlations are often moderate or weak when behavioural tasks rely on explicit judgments or when awareness varies (van der Wulp et al., 2025). For example, Batterink and Paller (2017) found that neural entrainment predicted performance on an implicit target detection task, but showed no reliable relationship with explicit familiarity or recognition measures. Importantly, several studies report little or no correlation between neural sensitivity to regularities and behavioural performance (Sznabel et al., 2023; Stothart and Smith, 2020; Morris et al., 2025a; Stothart et al., 2025).

Overall, whether neural and behavioural measures align depends critically on task alignment, learning stage, and the specific component of SL being assessed. These discrepancies likely reflect multiple factors. Behavioural tasks may lack sensitivity or may not be well matched to the implicit nature of the learning process, requiring explicit access to knowledge that is primarily encoded at a neural level. Additionally, neural entrainment may capture early, online, or transient representations of regularities that are not yet consolidated or accessible at the time of behavioural testing. More fundamentally, a growing body of work suggests that SL is not a unitary process, but a multicomponential ability (Frost et al., 2015). From this perspective, neural entrainment measures (e.g., FPVS, RSVP paradigms) and behavioural tasks may tap into partially distinct components of learning, explaining why convergence is sometimes observed but often incomplete. A final important factor is surely methodological in nature; some of these inconsistencies might depend on the fact that not all of these studies were explicitly designed to address individual variability. There is now more awareness that assessing a cognitive skill at the level of the single participant requires considerably more precision and statistical power than previously thought (Siegelman and Frost, 2015; Hedge et al., 2018; Brysbaert, 2019).

The role of Individual Differences

A growing body of research highlights substantial variability in SL abilities across individuals, challenging the traditional assumption that SL operates uniformly across learners. Empirical evidence shows that sensitivity to regularities varies widely, with some individuals exhibiting robust implicit pattern learning while others perform at or near chance levels (Kaufman et al., 2010; Siegelman and Frost, 2015). This variability has been increasingly recognized not as noise, but as a meaningful source of theoretical insight into the mechanisms underlying SL (Boogert et al., 2018; Kidd et al., 2018). Importantly, individual differences in SL have been linked to a range of cognitive outcomes, particularly in the language domain. Better performance in SL tasks has been associated with larger vocabularies, more advanced syntactic processing, and improved literacy skills in both first and second languages (Mainela-Arnold and Evans, 2014; Singh et al., 2012; Kidd and Arciuli, 2016; Misyak et al., 2010; Arciuli and Simpson, 2012; Frost et al., 2013a; Tong et al., 2019; von Koss Torkildsen et al., 2019a). Longitudinal studies further suggest that early SL abilities can predict later language development (Ellis et al., 2014; Shafto et al., 2012). However, the nature of these relationships remains debated, as SL performance may also reflect contributions from domain-general cognitive factors such as attention, memory, or processing speed (Conway et al., 2010; Misyak and Christiansen, 2012). Moreover, recent work points to inconsistencies in correlational findings and emphasizes the need for more reliable and theoretically grounded measures (Schmalz et al., 2019; Siegelman, 2020a). Within this context, examining individual variability provides a critical avenue for understanding how different components of SL contribute to learning and how neural and behavioural measures may differentially capture these processes.

Research Questions

Building on the view that SL can be indexed through neural entrainment, the present study investigates how frequency-based regularities drive implicit learning, and whether such learning extends beyond the neural level to behavioural expression. Building on the paradigm of De Rosa et al. (2022)—thus focusing on item frequency as a statistical cue—we specifically ask whether the observation of a strong and reliable neural entrainment necessarily results in behavioural learning that is equally robust, or whether neural responses can emerge in the absence of a clear behavioural counterpart. Previous work has shown that differences in frequency of occurrence are sufficient to elicit neural discrimination responses, suggesting that the brain can rapidly and implicitly extract such regularities (De Rosa et al., 2022). However, it remains unclear whether this neural sensitivity has a measurable behavioural counterpart. These issues bear on ongoing debates about the nature of SL, including whether it reflects a unitary mechanism or a multi-component process involving partially dissociable implicit and explicit systems.

2.2 Experiment 1

This experiment was designed to investigate whether frequency-based statistical regularities can be learned implicitly through exposure, and whether such learning is reflected at both neural and behavioural levels. Using an FPVS paradigm combined with behavioural tasks, we aimed to assess sensitivity to frequency differences, the retention of this information, and the relationship between neural responses and explicit performance.

2.2.1 Materials and Methods

All data, experimental materials, and analysis code are openly available on the Open Science Framework (OSF) at https://osf.io/tr2ku.

2.2.1.1 Participants

Forty-five paid volunteers (11 males; age: M = 24.29, SD = 4.43) took part in the experiment after giving written informed consent. Four subjects were rejected due to recording technical problems. All participants were Italian native speakers, right-handed, reporting having normal or corrected-to-normal vision and no history of linguistic, cognitive or neurological impairment. Participants remained unaware of the goals of the experiment until the end of the testing session. The present study was approved by SISSA Ethics Committee (Protocol Nr. 9001-III/13).

2.2.1.2 Stimuli

Stimuli comprised two categories: words and pseudofont strings. Each category included 16 items, all consisting of five-letter strings. Word stimuli (e.g., *conte*) were frequent singular nouns ($\log_{10}$ frequency: M = 3.31, SD = 0.12; orthographic neighbourhood density: OLD20 M = 1.16, SD = 0.23). Pseudofont strings (e.g., *abqsu*; $\log_{10}$ frequency = 3.24, SD = 0.34) were constructed by combining characters from the BACS-2 serif artificial script, designed to closely match the visual properties of Latin letters (Vidal et al., 2017).

A parallel set of foil stimuli were created for the behavioral tasks. Word foils were matched to the word stimulus set at the group level, controlling for length (all items were five letters long), lexical frequency ($\log_{10}$ frequency: M = 3.30, SD = 0.13), and orthographic neighbourhood density (OLD20 M = 1.13, SD = 0.22). We also made sure that no foil item had more than 2 letters in common with any target word. In addition, semantic constraints were applied to ensure sufficient distance from target

items. Semantic relatedness between each target and candidate foil was quantified using multilingual transformer-based word embeddings (cosine similarity). Foils showing high similarity to any target (cosine similarity $\geq .35$) were excluded, and only items with low maximum similarity across all targets were retained (acceptable set: $M = .18$, $SD = .07$, range = .02–.34), thereby minimizing close semantic relationships, idiomatic expressions, and fixed lexical associations. Pseudofont foils were generated using the same criteria, except of course frequency of occurrence ($\log_{10}$ frequency M = 3.11, SD = 0.25). Examples for the two experimental conditions are provided in Fig. 2.2. All stimuli are reported in the Supplementary Materials Section (see Table S1 for word stimuli and Table S2 for pseudofont stimuli).

2.2.1.3 Experimental Paradigm and Procedure

We implemented an FPVS-EEG study composed of a control and an experimental condition, paired with three two-alternative forced choice (2AFC) tasks designed to assess learning behaviourally: a Memory task performed after the FPVS control condition, and a Discrimination task and a Familiarity task performed after the FPVS experimental condition (see Fig. 2.1).

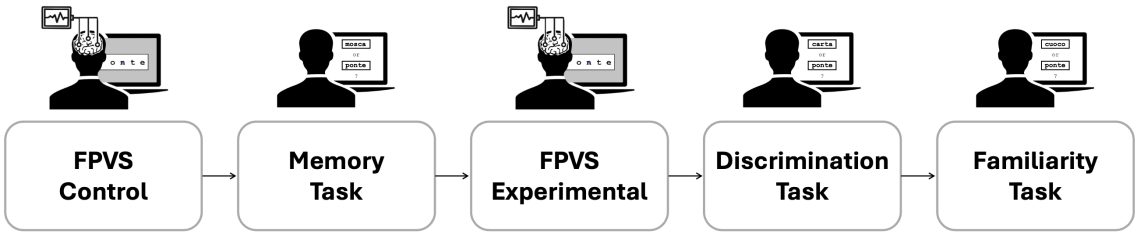


Figure 2.1: **Experimental procedure and task sequence.** Participants first completed the FPVS Control Task, followed by the behavioural Memory Task. They then completed the FPVS Experimental Task, after which two behavioural assessments were administered: the Discrimination Task and the Familiarity Task. During the FPVS tasks, participants were instructed to attend to the screen and perform a cross colour–change detection task to ensure sustained visual attention while stimuli were presented and EEG activity was recorded. During the behavioural tasks, participants made forced-choice responses between stimuli presented on screen. The order of tasks was fixed across participants.

The FPVS experimental task was designed using the oddball paradigm from De Rosa et al. (2022). Stimuli (N = 16) were randomly divided into two arbitrary sets: oddball stimuli (N = 8) and base stimuli (N = 8). This assignment was performed independently for each stimulus type and participant, and remained constant throughout the experiment. Contrary to classical oddball paradigms, where base and

oddball stimuli differ in category (e.g., words among pseudowords; Lochy et al. 2016), here both belonged to the same category and differed only in presentation frequency, with oddball stimuli presented four times less frequently than base stimuli. Stimuli were delivered in sequences of 360 items (6 items/s), including 72 oddballs (9 repetitions per item) and 288 base stimuli (36 repetitions per item), corresponding to an oddball frequency of 1.2 Hz (i.e., one oddball every five base stimuli).

In the control condition, no base/oddball distinction was present, and all stimuli were presented with equal probability (mean repetitions = 22.5, SD $<$ 2) – that is, we disrupted any statistical structure in the stimulus train.

The Memory task was designed to simply assess whether participants remembered the stimuli they were presented during the FPVS control run (both words and pseudofont strings). Stimuli were paired one-to-one with two distinct sets of foils (one per block), such that each stimulus appeared once per block, resulting in 32 stimulus–vs-foil trials. In addition, a foil–vs-foil condition was constructed using the same procedure, pairing items from the foil sets, yielding an additional 32 trials; The Memory task therefore comprised 64 word trials and 64 pseudofont trials overall (128 trials in total). The inclusion of the foil–vs-foil condition ensured that above-chance performance in the stimulus–vs-foil condition reflected prior familiarization during the FPVS phase rather than task strategies. Presentation location of the target item was counterbalanced across trials.

The Discrimination task was similar to the Memory task – we asked participants to identify items that were presented during the FPVS stream vs. unfamiliar foils. However, this task came after the FPVS experimental condition, so we were now interested in seeing whether participants had a better discrimination ability for higher-frequency/base stimuli vs. lower-frequency/oddball stimuli. The target timuli were paired one-to-one with three distinct sets of foils (one per block), resulting in 24 base–vs-foil trials and 24 oddball–vs-foil trials, with the presentation location of the target item counterbalanced across trials.

In the Familiarity task, we wanted to compare directly base and oddball stimuli. Therefore, base stimuli were randomly paired with oddball stimuli, and participants were asked which of the two was more familiar. Each stimulus was repeated three times, with the constraint that the same base–oddball pairing did not occur more than once, and presentation location was counterbalanced across trials.

In all three behavioural tasks, participants were encouraged to respond as fast as they could. A 2-seconds timeout was inserted and participants could select items by pressing 2 predefined keys on a keyboard.

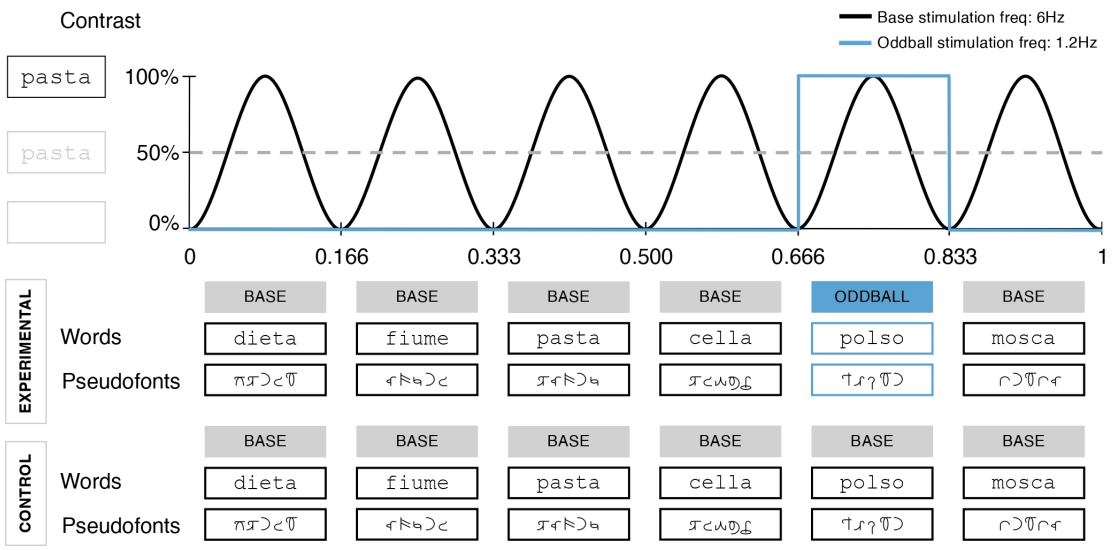


Figure 2.2: **FPVS Paradigm Structure.** In each sequence, stimuli were presented at a frequency of 6 Hz (over 60-seconds period trials) by sinusoidal contrast modulation. Each stimulation sequence consisted of 360 stimuli. Each stimulus reached maximum contrast 83 ms after its onset and completely faded out after 166 ms, corresponding to one full cycle, ranging from a 94% light grey background to full contrast and back. In the Control task, all stimuli were presented at 6Hz and roughly the same number of times per trial. In the Experimental task, oddball stimuli were presented every fifth item, thus resulting in a presentation frequency of 1.2 Hz (6/5 Hz). Unlike the traditional FPVS-oddball design, where oddballs are mixed within a sequence of base stimuli from a different category, oddballs and base stimuli drawn randomly from the same category (oddballs highlighted in blue for clarity), differing from bases only by their token frequency. Examples illustrate the two different conditions used in the experiment, i.e., words and pseudofont strings. Adapted from De Rosa et al. (2022).

In both the experimental and the control FPVS tasks, a trial comprised a 60s stimulation sequence, where stimuli were delivered through sinusoidal contrast modulation at a frequency of 6Hz. The presentation order of the stimuli was pseudo-randomized to eliminate consecutive item repetitions: stimuli repetitions in the Control task and base tokens in the Experimental task were separated by a minimum of 4 items, whereas oddball repetitions in the Experimental task were separated by $\leq$2 oddball presentations (i.e., minimum distance 14 items). There were 5 experimental trials per condition such that the total stimulation time was 10 minutes: 60s (trial duration) X 2 conditions (Words and Pseudofonts) X 5 trials. The trial presentation order was randomized, and trials were separated by a 30s break, ending with a 15s

countdown indicating the beginning of the next trial.

In the FPVS tasks, participants were instructed to fixate a cross displayed at the center of the screen and to press a button as soon as they detected a color change (from red to blue or vice versa), happening randomly six times during each trial, independent of the oddball-base manipulation. The purpose of this color-change detection task was to ensure that participants paid constant attention throughout the stimulation period.

Overall, participants' performance in the color-change detection task was highly accurate in both the Control task (Accuracy: M = 97.7%, SD = 0.13; Reaction Time: M = 445ms, SD = 0.1) and the Experimental task (Accuracy: M = 97.1%, SD = 0.15; Reaction Time: M = 472ms, SD = 0.12) and was comparable across the different stimuli conditions (Control task: Words - Accuracy = 97.4%, SD = 0.13, Reaction Time: M = 449ms, SD = 0.11; Pseudofonts - Accuracy = 98%, SD = 0.12, Reaction Time: M = 441ms, SD = 0.11; Experimental task: Words - Accuracy = 96.7%, SD = 0.17, Reaction Time: M = 475ms, SD = 0.12; Pseudofonts - Accuracy = 98%, SD = 0.13, Reaction Time: M = 474ms, SD = 0.13).

Participants were tested individually in a soundproof room, seated about 80 cm away from a 27" BenQ XL2720Z monitor (screen resolution of 1920 x 1080 pixels and a refresh rate of 60Hz). Stimuli were displayed using PsychToolbox-3 (Brainard and Vision, 1997) on MATLAB R2021b in the FPSV tasks and using PsychoPy (Peirce, 2007) in the 2AFC tasks. All items were shown at the center of the screen. Word stimuli were displayed in lowercase using the fixed-width Courier New font, while pseudofont strings were rendered in the BACS-2 serif font. Both fonts were bolded to 70% of their original weight to enhance visibility.

2.2.1.4 EEG Data Acquisition and Pre-processing

During the FPVS tasks, electroencephalographic (EEG) activity was recorded using an Active Two Biosemi system with 128 active electrodes positioned on an elastic cap, according to the radial-ABC system. We recorded at a sampling rate of 256Hz. Additional electrodes were employed for reference (Common Mode Sense active electrode, CMS) and grounding (Driven Right Leg passive electrode, DRL), with all electrode offset magnitude below 20mV. EEG data pre-processing was carried out in MATLAB using a combination of custom scripts and the EEGLAB toolbox (Delorme and Makeig, 2004). The pre-processing pipeline included band-pass filtering (0.1–100Hz), visual inspection of channels for artifacts through time series and spectra, and Independent Component Analysis to remove artifacts associated with eye

blinks and movements. Removed channels were interpolated and the data were segmented into 60s epochs. The data were then re-referenced to the average of all electrodes, resulting in one dataset consisting of five 60s trials per participant for each experimental condition.

2.2.1.5 Data Analysis

Data analyses were conducted to assess participants' sensitivity to frequency-based regularities at both the neural and behavioural levels, as well as the relationship between these two indices of learning. Neural analyses focused on the presence and magnitude of frequency-tagged responses elicited by the FPVS paradigm, providing an implicit measure of sensitivity to differences in item frequency. Behavioural analyses, in contrast, aimed to evaluate whether such regularities were retained and accessible in explicit tasks, and whether performance varied as a function of stimulus type and experimental condition. Together, these complementary approaches allowed us to characterize both the implicit encoding of statistical structure and its potential expression in overt behaviour.

EEG Data Analysis

The EEG data analysis follows the methodology employed in previous studies (De Rosa et al., 2022). Trials were averaged in the time domain separately for each condition and participant to reduce EEG activity not phase-locked to the stimulation. Subsequently, the data underwent Fast Fourier Transform (FFT), from which normalized amplitude spectra were derived for each channel by calculating the absolute value of the FFT and dividing it by the number of data points. Due to the 60s epoch length, the frequency resolution was notably high ($1/60 = 0.0167$Hz).

To address noise variations across the frequency spectrum, we employed a local baseline-correction method subtracting the mean of the surrounding 20 bins (10 from each side) from each frequency bin, as in previous literature (Dzhelyova and Rossion, 2014; De Rosa et al., 2022). Base and oddball responses were then defined as the sum of the baseline-corrected amplitudes at the base (6Hz) and oddball (1.2Hz) stimulation frequencies, along with their corresponding first three harmonics (base: 12, 18, 24 Hz; oddball: 2.4, 3.6, 4.8 Hz).

The EEG activity corresponding to the base stimulation frequency was analyzed within an occipito-parietal region consisting of twenty electrodes centered around and encompassing Oz (labeled as A23 according to Biosemi's ABC labeling), known to exhibit maximal SSVEP responses (Norcia et al., 2015). For the EEG response to the oddball stimulation frequency, two predefined regions-of-interest (ROIs) were

investigated. These ROIs were selected based on prior research employing similar visual stimuli (such as pseudofonts and letter strings, as in De Rosa et al. 2022 and Lochy et al. 2015). The primary ROIs for oddball responses comprised a cluster of five occipito-temporal sites, including and surrounding electrode PO7 on the left and PO8 on the right hemisphere (respectively labeled as A10 and B7 in Biosemi's ABC labeling).

Finally, in order to test the significance of the oddball response in each condition, the response of interest was averaged across the channels of each ROI for each participant, and then subjected to a one sample t-test.

Behavioral Data Analysis

Performance on the 2AFC tasks was assessed using accuracy as primary metric. To analyze the data, generalized linear mixed models (LMMs) were employed using R (R Core Team, 2023) lme4 package (Bates et al., 2015). For both the Memory and Discrimination tasks, accuracy was calculated as the proportion of correct responses. In the Familiarity task, we measured the rate of preference for one of two frequency categories (base or oddball).

2.2.2 Results

We first report the neural results obtained from the FPVS tasks, aimed at assessing participants' sensitivity to frequency-based regularities at the implicit level. We then turn to behavioural performance to evaluate whether such regularities were retained and accessible in explicit tasks.

2.2.2.1 FPVS Control Task

In the FPVS Control task, a response was observed at the base fundamental stimulation frequency of 6Hz. It was mainly localized in the occipito-temporal area and was significantly above the level of noise with both types of stimuli (Words: $t(41) = 16.75$, $p < 0.001$, $g = 2.62$ [95% CI: 1.473 - 2.418]; Pseudofonts: $t(41) = 14.25$, $p < 0.001$, $g = 2.23$ [95% CI: 1.462 - 2.529]), which indicates the effective synchronization of the visual system with the periodic stimulation, hence a constant allocation of attention resources.

No significantly relevant response was observed at the oddball stimulation frequency for both words (Left ROI: $t(41) = 0.256$, $p = 0.799$, $g = 0.04$ [95% CI: -0.267 - 0.346]; Right ROI: $t(41) = 2.062$, $p = 0.046$, $g = 0.322$ [95% CI: 0.006 - 0.634])

and pseudofonts (Left ROI: t(41) = -0.437, p = 0.664, g = -0.0683 [95% CI: -0.374 - 0.239]; Right ROI: t(41) = 0.734, p = 0.467, g = 0.115 95% CI: -0.193 to 0.421]), hence confirming that the activation we might observe at the oddball frequency in the Experimental task can only be the result of the item frequency manipulation.

2.2.2.2 FPVS Experimental Task

In the FPVS Experimental task (see Fig. 2.3), a significant response at the base fundamental stimulation frequency was again observed in the occipito-temporal area with both types of stimuli (Words: t(41) = 17.81, p < 0.001, g = 2.78 [95% CI: 2.097 - 3.457]; Pseudofonts: t(41) = 16.54, p < 0.001, g = 2.58 [95% CI: 1.938 - 3.221]). We also observed a significant response to the oddball stimulation frequency of 1.2Hz, in the predefined ROIs for both words (Left ROI: t(41) = 5.73, p < 0.0001, g = 0.895 [95% CI: 0.528 - 1.254]; Right ROI: t(41) = 3.58, p = 0.0001, g = 0.558 [95% CI: 0.226 - 0.885]) and pseudofonts (Left ROI: t(41) = 4.15, p = 0.0002, g = 0.648 [95% CI: 0.307 - 0.982]; Right ROI: t(41) = 4.16, p = 0.0002, g = 0.649 [95% CI: 0.308 - 0.983]). The scalp topography of these effects is similar to the pattern seen in previous studies that used the same types of stimuli at similar frequency rates (Liu-Shuang et al., 2014; De Rosa et al., 2022).

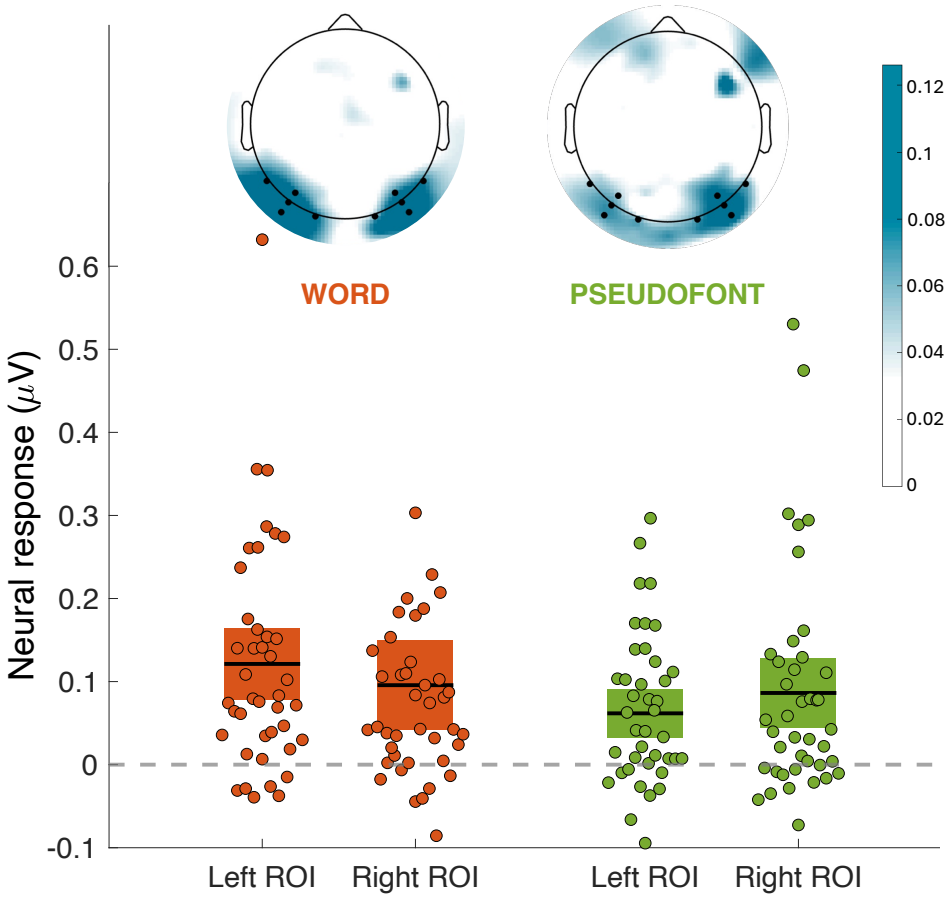


Figure 2.3: **FPVS grand-averaged token frequency response** Grand-averaged token frequency response estimated as the summed baseline corrected amplitudes for the oddball frequency (1.2Hz) and its three harmonics (2.4, 3.6, 4.8 Hz) for the two different conditions (Words in orange, Pseudofonts in green) in the left and right occipito-temporal ROIs, marked in the corresponding scalp topographies. Boxes illustrate the 95% t-interval around the mean. Dots stand for single participants. The dashed line represents the level of noise.

2.2.2.3 Memory Task

Participants performed above chance in the word condition; they could recognize stimuli they were previously familiarized to from new items (M = 79.91%, SD = 11.37, t(41) = 17.06, p < 0.0001, d = 17.06 [95% CI: 43.68% - 48.88%]). This did not occur for pseudofonts, where performance did not exceed chance level (M = 46.43%, SD = 8.41, t(41) = -2.75, p = 0.995, one-tailed, d=−0.42).

Since the stimuli were repeated twice (one appearance per block), we checked for differences in performance between blocks. No significant difference was detected (words: t(41) = -0.8, p = 0.43; pseudofonts: t(41) = 0.95, p = 0.34).

As for the foil vs. foil condition, participants performed at chance as expected in both conditions (Words: M = 57.21%, SD = 9.17, t(41) = 5.1, p < 0.0001, d = 0.79 [95% CI: 54.36% - 60.07%]; Pseudofonts: M = 50.9%, SD = 10.29, t(41) = 17.06, p = 0.58, d = 0.087 [95% CI: 47.69% - 54.1%]), confirming that the discrimination observed in the stimuli vs. foil condition reflects learning during the FPVS Control Task.

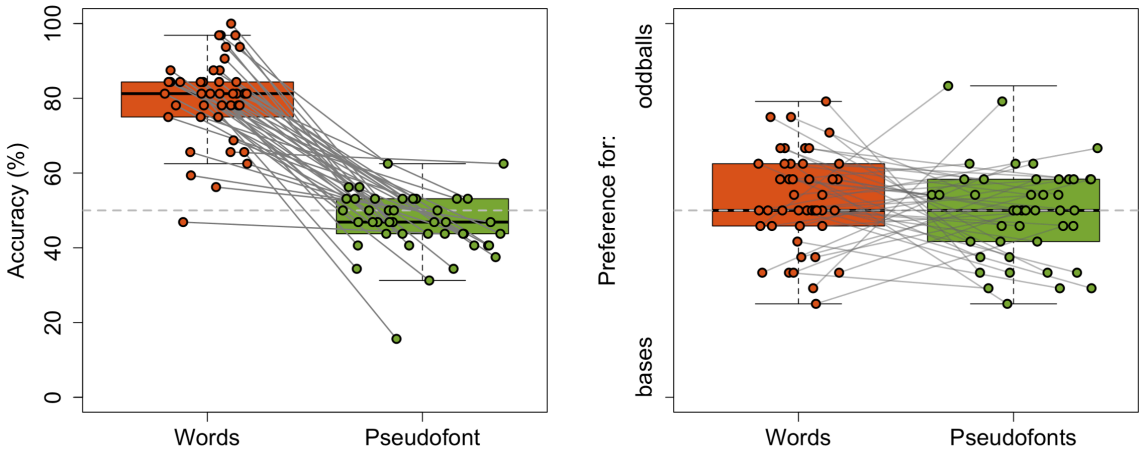


Figure 2.4: **Performance in Memory Task (left) and Familiarity Task (right)** (a) Accuracy in the memory task. Boxes depict the 95% t-interval around the mean and dots represent individual participants. The dashed line indicates the noise level. (b) Preference for oddballs or bases as being more frequent/familiar. Each dot represents one participant and the dashed line indicates chance level.

2.2.2.4 Discrimination Task

In the discrimination task, we compared participants' explicit memory for bases and oddballs, in order to assess whether there is an asymmetry in performance that can be ascribed to the frequency with which participants were exposed to the stimuli during the FPVS task. Participants were almost at ceiling with words, both with base and oddballs (bases: M = 89.13%, SD = 9.5, t(41) = 26.38, p < 0.0001, d = 4.12 [95% CI: 86.13% - 92.12%]; oddballs: M = 88.21%, SD = 13.21, t(41) = 18.53,

$p < 0.0001$, d = 2.89 [95% CI: 84.04% - 92.13%]). Pairwise comparisons revealed no significant differences in performance across blocks for both words (bases: $|t| \leq 0.90$, $ps = 1.00$; oddballs: $|t| \leq 1.56$, $ps \geq .36$) and pseudofonts (bases: $|t| \leq 0.45$, $ps = 1.00$; oddballs: all $|t| \leq 1.81$, $ps \geq .21$), Bonferroni-corrected. Collapsing across blocks, there was no evidence of differential performance between base and oddball items for either pseudofonts ($\beta = -0.004$, SE = 0.020, $t = -0.21$, $p = .84$) or words ($\beta = 0.005$, SE = 0.020, $t = 0.26$, $p = .80$), indicating that frequency differences did not translate into explicit behavioural sensitivity.

As one might expect based on the results of the Memory task, participants' performance was clearly much worse with pseudofonts, but still significantly above chance (bases: M = 54.88%, SD = 13.91, t(41) = 2.25, p = 0.03, d = 0.35 [95% CI: 50.49% - 59.27%]; oddballs: M = 55.08%, SD = 14.89, t(41) = 2.18, p = 0.035, d = 0.34 [95% CI: 50.38% - 59.78%]; see Fig. 2.5). Here again, there was no significant difference in performance between bases and oddballs ($\beta = -0.004$, SE = 0.020, $t = -0.21$, $p = .84$), and no reliable differences emerged across blocks (all $|t| \leq 1.81$, all $ps \geq .21$, Bonferroni-corrected).

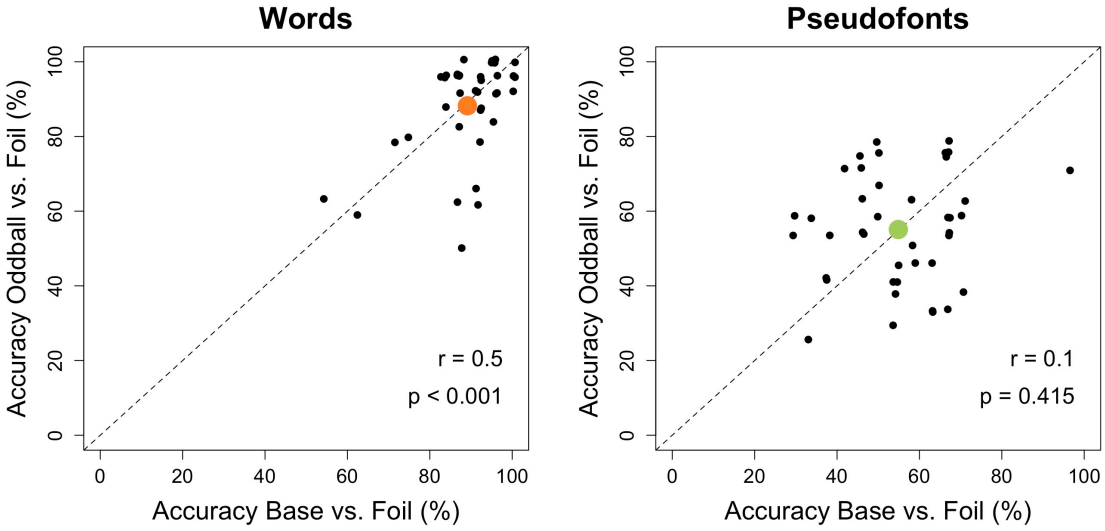


Figure 2.5: **Discrimination Task performance** The left panel illustrates accuracy in discriminating bases from foils and oddballs from foils for words. The right panel displays the same comparison for pseudofonts. Each black dot represents the performance of an individual participant, with the larger colored dot representing the group mean. The diagonal dashed line indicates equal performance across both tasks.

Interestingly, in the Word condition there was a fairly clear correlation between the performance on the bases and the performance on the oddballs (see Fig. 2.5). This suggests that an overall memory factor subtends these results; participants were either good or bad at remembering the FPVS items, but this applies symmetrically to both low- and high-frequency items. On the contrary, there was no meaningful correlation in the pseudofonts condition, which clearly argues against the memory factor account proposed for the word items. In fact, the stark difference between words and pseudofonts might even suggest that the discrimination data on these latter is essentially just noise – our participants didn't really have a memory trace for pseudofont strings that they were showing during the FPVS runs (at least at the group level).

2.2.2.5 Familiarity Task

The Familiarity task was meant to pit bases and oddballs directly against each other; we explicitly asked our participants which of the two types of stimuli they considered to be more familiar.

As visible in Figure 2.4b, participants showed no clear preference for either of the two frequency categories (rate of preference for oddballs; Words: M = 52.84% SD = 12.89, t(41) = 1.46, p = 0.51, d = 0.22 [95% CI: 48.92% - 56.76%]; Pseudofonts: M = 49.71% SD = 12.75, t(41) = -0.15, p = 0.88, d = -0.22 [95% CI: 45.84% - 53.59%]). No significant relation was observed in the rate of preference for oddballs between the word and the pseudofont conditions (r = 0.23, t(42) = 1.54, p = 0.13, [95% CI: -0.494, 0.071]).

A great deal of individual variability was present in both the word and the pseudofont condition, which might indicate that at the individual level some participants did develope a preference for one of the two categories. We will come back to this point later in the manuscript.

2.2.3 Discussion

Experiment 1 aimed at testing whether frequency-based statistical regularities generate a reliable neural entrainment. To this purpose, we used De Rosa et al. (2022)'s paradigm, and extended it to include a control condition with no statistical cues – to effectively check that no neural entrainment emerges when no cues are available. Moreover, Experiment 1 asks whether neural entrainment translates into some sign of behavioural learning.

At the neural level, the results provide clear evidence that the brain is sensitive to frequency-based regularities. Robust oddball responses were observed for both words and pseudofonts, replicating previous FPVS findings (De Rosa et al., 2022). We also extended De Rosa et al.'s results by showing that the signal disappears, as it should, when the statistical structure in the train of stimuli is disrupted (while the purely visual entrainment at 6 Hz is perfectly preserved). This provides further confirmation that the 1.2 Hz entrainment is genuinely created by the item frequency distribution, and demonstrates that the visual system can track statistical structure in the input even in the absence of categorical differences, supporting the idea that FPVS captures rapid and automatic learning mechanisms.

In contrast, behavioural results revealed a markedly different pattern. Even in the absence of any statistical cues (that is, in the control FPVS condition), participants showed clear signs of recognition only for words; with pseudofont strings, their recognition performance was very weak, just barely above chance. This suggests that even remembering items from an FPVS stimulus stream is quite hard, especially if one does not have prior familiarity or existing representations to rely on (words vs. pseudofonts). This shows a clear interaction between the dynamic and fast encoding of visual stimuli and the content of our (episodic and semantic) memory. Crucially, these modulations of our recognition performance occur despite similarly strong neural responses for both stimulus types, providing a first clear indication of a dissociation between neural sensitivity and the behavioural expression of learning.

A further dissociation emerges when considering sensitivity to the core manipulation in our design – frequency differences. Behaviourally, participants showed no reliable distinction between bases and oddballs, neither in discrimination nor in familiarity judgments. This contrasts sharply with the robust oddball responses observed in the EEG signal, indicating that the brain effectively encodes frequency differences even when such information is not accessible to conscious report or decision-making.

Together, these findings suggest that the statistical learning mechanisms indexed by FPVS operate at an implicit level that does not necessarily translate into explicit behavioural knowledge.

Two limitations should be noted, however. First, the use of identical stimuli in the control and the experimental FPVS runs may have led to over-familiarization; participants might have come to the exposure of the statistically-structured stream after having learned the stimuli too well during the exposure to the unstructured condition. This might have potentially reduced the sensitivity of the behavioural measures. Second, our design was such that the target items in the Memory and Discrimination tasks were repeated across trials, while foils were not. Although the lack of any block effect is reassuring in this respect, we cannot entirely exclude that this target repetition might have worked as a cue for the participants. We then ran a second experiment where we tried to replicate the pattern of results described here while fixing these limitations. Moreover, running a second batch of participants yielded a much improved sample size, which gave us the chance to assess the correlation between neural entrainment and behavioural learning at the individual level.

2.3 Experiment 2

Experiment 2 was designed to replicate and extend the findings of Experiment 1 while addressing some of its methodological limitations. In particular, we aimed to further investigate the dissociation between neural sensitivity to frequency-based statistical regularities and their behavioural expression, while minimizing potential over-familiarization effects and improving the sensitivity of the behavioural measures.

2.3.1 Materials and Methods

All data, experimental materials, and analysis code are openly available on the Open Science Framework (OSF) at https://osf.io/tr2ku.

2.3.1.1 Participants

Forty-four paid volunteers (11 males; age: M = 23.18, SD = 2.68) took part in the experiment after giving written informed consent. One subject was rejected due technical problems. Participants were recruited with the same criteria as for Experiment 1 and were unaware of the goals of the experiment during the testing session.

2.3.1.2 Stimuli and Analyses

The stimuli, procedures, and analyses employed in this experiment were identical to those used in Experiment 1, with the following exceptions.

In contrast to Experiment 1, two distinct stimulus sets were created to be used in the control and experimental FPVS conditions. Each set included both words and pseudofont strings constructed according to the same criteria described previously (see Experiment 1, Stimuli Section). Word stimuli were frequent singular nouns ($\log_{10}$ frequency: M = 3.28, SD = 0.16; orthographic similarity: OLD20 M = 1.13, SD = 0.19), while pseudofont strings were matched in orthographic properties ($\log_{10}$ frequency = M = 3.21, SD = 0.35). All stimuli are reported in the Supplementary Materials Section (see Table S3 for word stimuli and Table S4 for pseudofont stimuli).

Participants were assigned to one of two counterbalanced groups: for half of the participants, Set 1 was used in the FPVS Control condition and Set 2 in the FPVS Experimental condition; for the other half, the assignment was reversed. Thus, stimuli presented in the Control and Experimental conditions were always non-overlapping within participants.

A second novelty was that we now generated a unique set of foils for the behavioral tasks. This set was matched to the target items at the group level for length (five letters), lexical frequency ($\log_{10}$ frequency: M = 3.27, SD = 0.14), and orthographic neighbourhood density (OLD20 M = 1.12, SD = 0.18). Also, as in Experiment 1, we made sure that there was never more than two letters in common with the target items, nor any close semantic relationship (for the words condition). Critically, however, this was now guaranteed between each individual target item and each individual foil, so that we could randomly select a new foil for each target within each of the two (for the Memory task) or three (for the Discimination task) testing blocks. That is, the foils were now repeated across different testing blocks just like the target items, thus removing the asymmetry that might have worked as a cue in Experiment 1.

2.3.2 Results

As in Experiment 1, we first report the neural results from the FPVS tasks, assessing implicit sensitivity to frequency-based regularities, before turning to behavioral performance to examine explicit retention of the learned structure.

2.3.2.1 FPVS Control Task

Just as in Experiment 1, we observed a strong neural response at the base fundamental stimulation frequency of 6Hz in the FPVS Control task (words = t(43) = 15.99, p < 0.0001, g = 2.44 [95% CI: 1.835 - 3.034]; pseudofonts = t(43) = 16.7, p < 0.0001, g= 2.55 [95% CI: 1.923 - 3.162]).

Although a significant response was observed at the oddball stimulation frequency of 1.2 Hz for words in the left ROI (t(43) = 0.61, p = 0.571, g = 0.09 [95% CI: -0.208 - 0.392]), the effect size suggests that this finding might reflect individual variability rather than a robust effect. No significant responses were observed in the right ROI for words (t(43) = 3.745, p = 0.0005, g = 0.571 [95% CI: 0.256 - 0.892]) or in either ROI for pseudofonts (Left ROI: t(43) = 0.39, p = 0.7, g = 0.056 [95% CI: -0.241 - 0.358]; Right ROI: t(43) = -0.858, p = 0.4, g = -0.131 95% CI: -0.43 to 0.17]).

2.3.2.2 FPVS Experimental Task

In the FPVS Experimental task (see Fig.2.6), a significant response was observed at the base stimulation frequency of 6HZ in the occipito-temporal area, for both

types of stimuli (words: $t(43) = 14.47$, $p < 0.0001$, $g = 2.21$ [95% CI: 1.646 - 2.761]; pseudofonts: $t(43) = 15.05$, $p < 0.0001$, $g = 2.29$ [95% CI: 1.718 - 2.864]). We also observed a significant response to the oddball stimulation frequency of 1.2Hz in both the predefined ROIs, again for both words (Left ROI: $t(43) = 3.58$, $p < 0.0009$, $g = 0.546$ [95% CI: 0.222 - 0.863]; Right ROI: $t(43) = 3.193$, $p = 0.003$, $g = 0.487$ [95% CI: 0.168 - 0.801]) and pseudofonts (Left ROI: $t(43) = 2.9$, $p = 0.006$, $g = 0.442$ [95% CI: 0.126 - 0.753]; Right ROI: $t(43) = 3.646$, $p = 0.0007$, $g = 0.556$ [95% CI: 0.232 - 0.875]). Scalp topographies patterns were similar to those in Experiment 1.

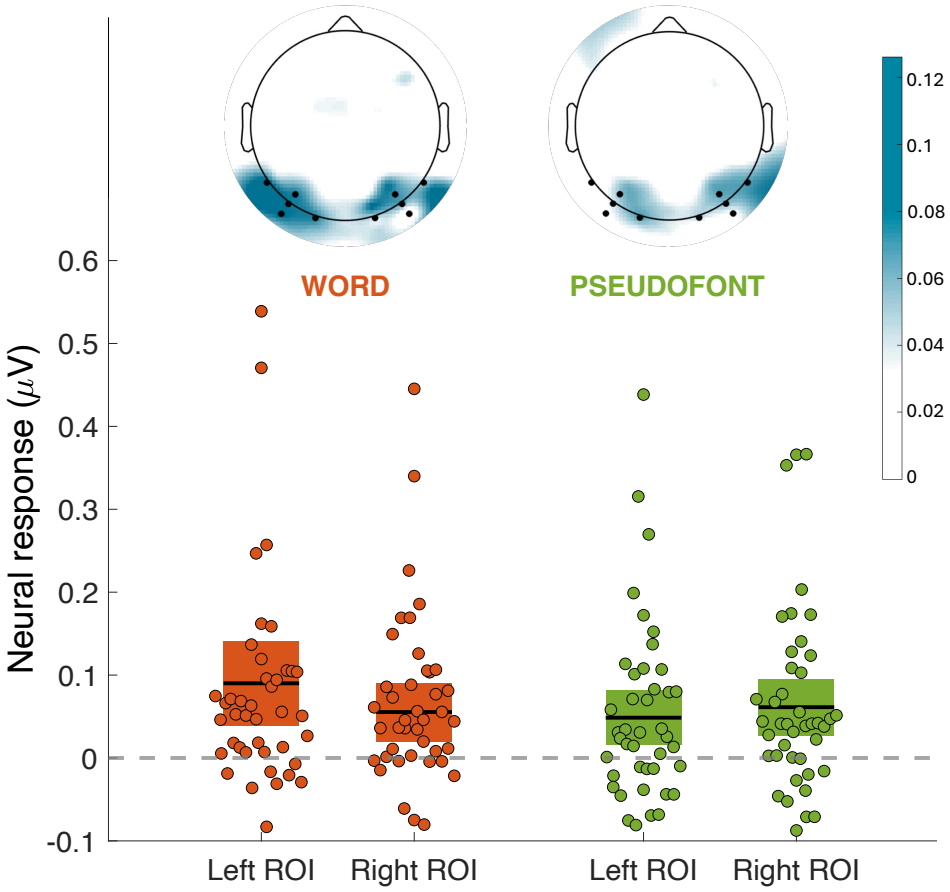


Figure 2.6: **FPVS grand-averaged token frequency response** Grand-averaged token frequency response estimated as the summed baseline corrected amplitudes for the oddball frequency (1.2 Hz) and its three harmonics for the two different conditions (Words and Pseudofonts) in the left and right occipito-temporal ROIs, marked in the corresponding scalp topographies. Boxes depict the 95% t-interval around the mean, dots stand for single participants and the dashed line illustrates the level of noise.

2.3.2.3 Memory Task

In the Memory Task, participants performed again above chance in the word condition (M = 67.73% SD = 10.36, t(43) = 11.23, p < 0.0001, d = 1.71 [95% CI: 64.54% - 70.92%]) and at chance in the pseudofont condition (M = 51.24% SD = 11.47, t(43) = 0.71, p = 0.48, d = 0.11 [95% CI: 47.70% - 54.77%]). As for the foil vs. foil condition, participants performed at chance, as expected, in both conditions (Words: M = 50.58% SD = 9.6, t(43) = 0.06, p = 0.69, d = 0.79 [95% CI: 47.63% - 53.54%]; Pseudofonts: M = 52.33% SD = 10.02, t(43) = 1.52, p = 0.14, d = 0.23 [95% CI: 49.24% - 55.41%]).

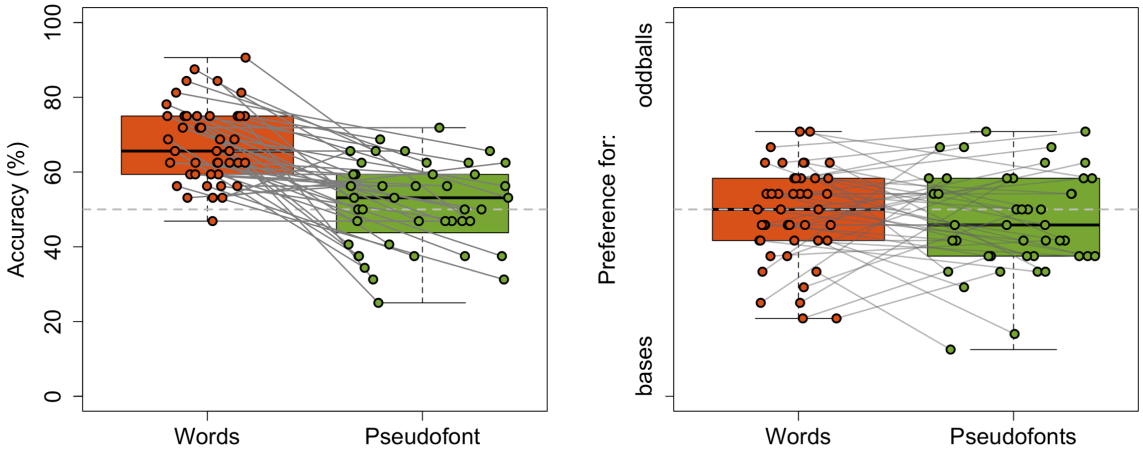


Figure 2.7: **Performance in Memory Task (left) and Familiarity Task (right)** (a) Accuracy in the memory task. Boxes depict the 95% t-interval around the mean and dots represent individual participants. The dashed line indicates the noise level. (b) Preference for oddballs or bases as being more frequent/familiar. Each dot represents one participant and the dashed line indicates chance level.

2.3.2.4 Discrimination Task

In the discrimination task, participants were again almost at ceiling with words, in both the base and the oddball condition (Bases: M = 78.59% SD = 16.48, t(43) = 11.37, p < 0.0001, d = 1.73 [95% CI: 73.51% - 83.66%]; Oddballs: M = 76.07% SD = 15.62, t(43) = 10.94, p < 0.0001, d = 1.67 [95% CI: 71.26% - 80.87%]), while they performed at chance with pseudofonts (Bases: M = 52.42% SD = 14.06, t(43)

= 1.13, p = 0.27, d = 0.17 [95% CI: 48.09% - 56.75%]; Oddballs: M = 51.55% SD = 12.79, t(43) = 0.79, p = 0.43, d = 0.12 [95% CI: 47.61% - 55.49%]).

Considerable individual variability was observed in both conditions (see Fig. 2.8), indicating differences across participants. Analyses revealed the same positive correlation observed in Experiment 1 between performance in the word base vs. foil and oddball vs. foil conditions (r = 0.67, t(43) = 5.79, p < 0.0001), which was not present for pseudofonts (r = 0.079, t(43) = 0.51, p = 0.61).

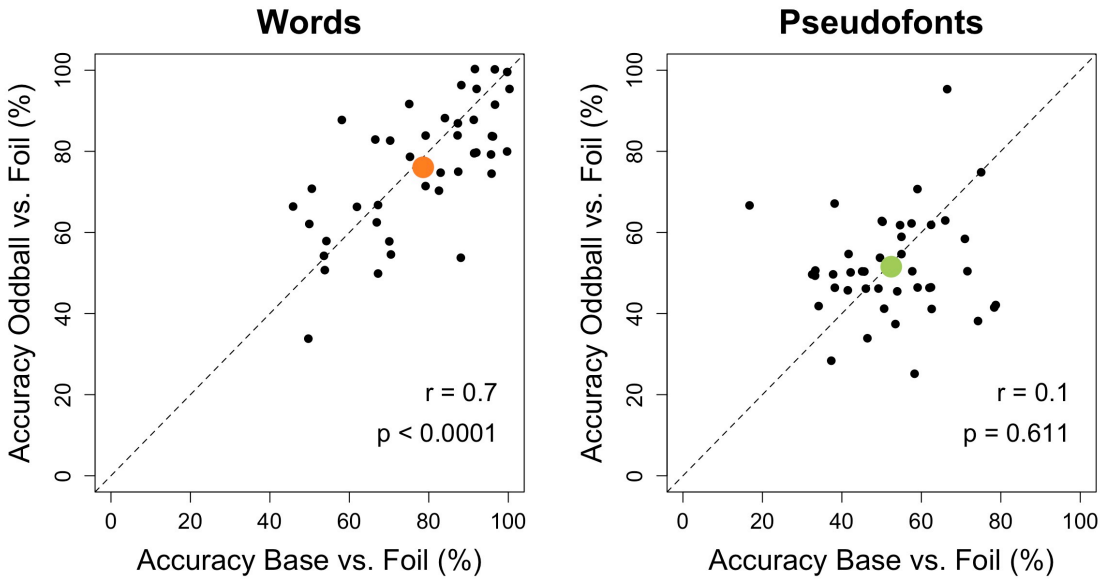


Figure 2.8: **Discrimination Task performance** Discrimination performance accuracy in the base vs. foil and oddball vs. foil conditions, for words and pseudofonts. Each black dot represents the performance of an individual participant, with the larger colored dot representing the group mean. The diagonal dashed line indicates equal performance across both tasks.

2.3.2.5 Familiarity Task

As in Experiment 1, also in Experiment 2 participants showed at group level no preference for either of the two frequency category groups of items (rate of preference for oddballs; Words: M = 47.16% SD = 12.77, t(43) = -0.79, p = 0.44, d = -0.12 [95% CI: 44.6% - 52.37%]; Pseudofonts: M = 49.71% SD = 13.44, t(43) = -1.4, p = 0.17, d = -0.21 [95% CI: 43.07% - 51.25%]), though again substantial individual variability was present in both the word and the pseudofont conditions (see Fig.

2.7b). No significant relation was observed in the rate of preference for oddballs between the word and the pseudofont conditions (r = 0.24, t(43) = 1.6, p = 0.12, [95% CI: -0.061, 0.506]).

2.3.3 Discussion

Experiment 2 perfectly replicated the results of Experiment 1, with an improved design that took care of a couple of potentially harmful issues. This fully confirms the pattern of results that we described above: differences in item frequency within a relatively short and quick stimulus train determine a strong and reliable neural entrainment, both for familiar and unfamiliar stimuli. This can be interpreted in terms of item grouping – the brain used the difference in frequency of occurrence to put items in different bins, thus yielding an oddball effect that is similar, at least qualitatively, to what we observe with items of different sorts (e.g., pictures of faces vs. pictures of objects; words vs. nonwords; De Rosa et al., 2022; Lochy et al., 2015).

Crucially, we did not find any behavioural trace of this neural learning. Of course, this does not mean that it is generally impossible to find a behavioural counterpart of frequency-based SL neural entrainment; we can never be sure that, e.g., using other tasks or other stimuli would deliver those signs. Yet, we did use multiple behavioural tasks that are commonly adopted to address SL, so we can at least hold that we obtained a very clear dissociation between neural entrainment and behaviour in a relevant experimental setup.

Thus far, we investigated the relationship between neural entrainment to statistical regularities and behavioural learning at the group level. The results showed robust and reliable neural entrainment in the absence of clear behavioural learning effects when assessed with standard statistical learning tasks. A complementary approach is to examine these processes from an individual variability perspective, by assessing the extent to which neural and behavioural measures covary across participants. However, this approach is considerably more demanding. Individual differences research typically requires large samples (e.g., Brysbaert, 2019), and tasks that are sensitive at the group level are not necessarily suitable as psychometric tools for capturing stable individual differences (Siegelman and Frost, 2015; Hedge et al., 2018).

To address these challenges, we combined data from Experiment 1 and Experiment 2 (total $N = 84$). Although the two experiments differed slightly in their design, they were conceptually identical and yielded highly similar patterns of results at the group level, making it appropriate to consider them jointly. This approach allows us to provide a more robust characterization of individual variability while minimizing

the influence of sample-specific noise. In addition, we explicitly assessed the reliability of each measure as a prerequisite for any meaningful investigation of individual differences.

Reliability

We assessed the reliability of all behavioural and neural measures across the combined sample using internal consistency (Cronbach's α).

The neural response to the oddball stimulation frequency in the FPVS task showed low reliability overall. Internal consistency was low for both words ($\alpha = .21$) and pseudofonts ($\alpha = .39$), indicating that this neural measure does not provide a stable estimate of individual differences under the present conditions.

Behavioural measures showed higher, but variable, reliability across tasks. Reliability was highest for the discrimination task with words (base: $\alpha = .76$; oddball: $\alpha = .78$), suggesting that this task captures relatively stable individual differences in performance. In contrast, reliability was substantially lower for pseudofonts ($\alpha \approx .46$), and the familiarity task also showed relatively low internal consistency overall (words: $\alpha = .45$; pseudofonts: $\alpha = .51$).

Taken together, these results reveal substantial variability in the psychometric quality of the measures. While some behavioural tasks—particularly discrimination with words—approach acceptable levels of reliability, others do not, and the FPVS neural measure shows consistently low reliability.

Implications for Individual Differences Research

These findings highlight important challenges for the study of individual variability in statistical learning. First, the tasks used here are widely adopted in the literature, yet they were not originally designed with psychometric purposes in mind. As a result, their ability to capture stable individual differences is limited, despite their effectiveness at detecting group-level effects. This dissociation between experimental sensitivity and psychometric reliability has been emphasized in recent work (Siegelman and Frost, 2015; Hedge et al., 2018), and is clearly evident in the present data.

Second, to our knowledge, this is one of the first studies to explicitly assess the reliability of neural entrainment measures in this context. Despite using a paradigm closely aligned with previous FPVS designs, reliability was low. One likely explanation concerns the number of trials. FPVS paradigms are optimized to yield strong group-level effects with relatively few stimulus sequences due to their high signal-to-noise ratio. However, this strength may come at the cost of reduced reliability at

the individual level. Increasing the number of runs may improve reliability, although this remains an open empirical question and may introduce additional factors such as neural adaptation across repeated exposures.

A further consideration concerns the nature of the statistical cues manipulated in the present study. Unlike most statistical learning paradigms, which focus on transitional probabilities, we manipulated item frequency (i.e., univariate statistics). Although there is clear evidence that such information is encoded both in natural language (e.g., word frequency effects) and in experimental settings (e.g., Zhou et al., 2024), its suitability for capturing individual differences has not been systematically investigated. It is therefore possible that sensitivity to this type of statistical information is more difficult to characterize at the individual level, although current evidence does not allow strong conclusions on this point.

Overall, the present findings suggest that investigating individual variability in statistical learning requires careful consideration of both measurement reliability and task design. While the paradigms used here are well suited to detect group-level effects, they do not consistently provide reliable estimates at the individual level—particularly for neural entrainment measures. More broadly, these results highlight the need for methodological work aimed at developing tasks that can robustly capture both neural and behavioural sensitivity to distributional regularities at the level of the individual participant.

2.4 General Discussion

The present study investigated the mechanisms behind the learning of frequency-based distinctions in a rapid stream of visual stimuli. Across two experiments that provided perfectly consistent results, we confirmed previous work showing that frequency of occurrence is sufficient to elicit neural discrimination responses (De Rosa et al., 2022). Crucially, we extend this work by showing that this solid and significant neural discrimination does not necessarily determine behavioural learning, therefore establishing a dissociation between neural sensitivity and behavioural expression of Statistical Learning (SL), which has deep implications for its nature and the neurocognitive mechanisms behind it.

Neural Entrainment and Behavioural Measures of Learning

The first key finding is that frequency-based regularities are robustly encoded at the neural level. In both experiments, clear responses at the oddball frequency were observed for both words and pseudofonts, despite the absence of pre-existing perceptual or categorical differences between base and oddball stimuli. This demonstrates that the visual system is highly sensitive to differences in frequency of occurrence, replicating and extending previous FPVS findings (Rossion, 2014; Norcia et al., 2015) and recent work showing that frequency alone can drive neural discrimination (De Rosa et al., 2022). These results are consistent with the broader view that SL reflects sensitivity to distributional properties of the input (Frost et al., 2019), and with hierarchical models of visual processing in which repeated exposure shapes neural tuning (Dehaene et al., 2005; Vinckier et al., 2007).

In contrast, behavioural evidence for learning was limited, if not entirely absent. Across both experiments, participants were only able to discriminate between words that were presented in the FPVS stream and comparable foils; they were not able to produce the same performance with the pseudofont strings. Note that this is mere memory for the overall set of stimuli, not a sign of learning of the frequency difference; even this first step was only possible for stimuli that were highly familiar (mid-frequency words), but not for items with no previous representations in memory (pseudofont strings). Critically, even for words, behavioural responses did not reflect sensitivity to frequency differences *per se*. Similar discrepancies between neural sensitivity and behavioural expression have been reported in previous work (Sáringer et al., 2024; Barillon et al., 2025). Such dissociations are increasingly documented in the literature. While some studies report neural-behavioural correlations under well-aligned conditions (Dzhelyova et al., 2020; Moser et al., 2021b),

others report weak or even absent relationships, particularly when behavioural tasks rely on explicit judgments (Chiarella et al., 2026; Rawal and Tseng, 2021; van der Wulp et al., 2025). Several factors likely contribute to this dissociation. First, neural entrainment reflects implicit, online processing of structured input, whereas behavioural tasks arguably require some access to the learned information (and even explicit, fully-conscious access sometimes). Second, neural entrainment may capture early and potentially transient representations that are not yet consolidated; therefore, it is only natural that they do not easily emerge in behavioural measures. Third, behavioural expression appears to depend on the interaction between newly learned regularities and prior knowledge, as reflected in the advantage observed for words over pseudofonts; this is in stark contrast with the neural response, which is largely independent of the specific stimuli used in the perceptual stream (see also, e.g., De Rosa et al., 2022). Finally, behavioural tasks introduce additional cognitive demands—such as attention, working memory, and decision processes—which may obscure or dilute the contribution of statistical learning.

Individual Differences and the Issue of Reliability

Ideally, we would have liked to complement the group-level dissociation – we can observe very solid neural entrainment without signs of behavioural learning *across subjects* – with an individual-subject dissociation – a low correlation between neural entrainment and behavioural learning *in each individual*. However, this was not the primary goal of this study, so we did not explicitly designed our tasks as psychometric tools. We checked ex-post whether we obtained enough precision at the level of the individual subject, and we discovered that this was not the case: in particular, the FPVS neural measures showed low internal consistency, indicating that while the paradigm is highly sensitive at the group level, it is unsuitable for assessing variability across individuals (at least under the present conditions; we discussed in the Discussion Section ways in which the current paradigm could be improved as a measure of individual variability).

These findings highlight a broader methodological challenge. Tasks that are highly effective at detecting group-level effects do not necessarily function as reliable psychometric tools. This dissociation – often referred to as the "reliability paradox" (Hedge et al., 2018) – is clearly evident in the present data. This calls for more research on individual variability, both on the theoretical and methodological side; we clearly need paradigms and tasks *ad hoc* to measure individual differences reliably, which is a necessary prerequisite to then investigate correlations and relationships between different aspects, or component perhaps, of the human SL machine.

Statistical Learning as a Multicomponential Process

Taken together, the present findings sit well with a multicomponential framework of SL (Frost et al., 2015; Siegelman and Frost, 2015; Aslin, 2017; Bogaerts et al., 2022a). According to these accounts, SL reflects a set of domain-general computational principles instantiated within modality-specific systems, with learning outcomes shaped by stimulus properties, prior knowledge, and task demands.

Within this framework, neural entrainment and behavioural performance may reflect partially distinct aspects of learning. Neural entrainment may index the online extraction of statistical regularities during perceptual processing, whereas behavioural performance may depend more strongly on the extent to which this information is encoded, integrated with memory representations, and made accessible for explicit task demands. Our data do not directly adjudicate the representational format of these memory traces (e.g., chunk-like representations vs. stored transitional probabilities; see Perruchet and Vinter, 1998; Thiessen et al., 2013; Slone and Johnson, 2018). Crucially, however, our findings suggest that these components are not necessarily tightly coupled: learning-related neural sensitivity may emerge without corresponding behavioural expression, particularly when acquired knowledge remains implicit or when behavioural tasks are poorly aligned with the learned structure.

This dichotomy between neural entrainment and behavioural learning resonates with previous theoretical accounts that also stipulated the existence of separate SL mechanisms. The extraction and integration framework of Thiessen et al. 2013 – if one imagines that neural entrainment might track the extraction of the relevant statistical information, but perhaps integration is necessary in order for that information to generate behavioural changes. We do not seem to read such implicit-explicit distinction in the original theoretical proposal, but we potentially see a connection between the extraction of information from the input as a more genuinely perceptual process, which might or might not leave behavioural traces; and the integration of that information in memory as a more obvious stage where representations change (or are built) in such a way that they affect behaviour observably.

Perhaps, an even more obvious connection can be established between the current results and views that hold the existence of a separate perceptual binding/perceptual stage of SL on the one hand, and a memory stage where the new information gets integrated into memory on the other (e.g., Batterink and Paller, 2017). Quite naturally, one could ascribe the kind of neural entrainment we observed here, without accompanying signs of behavioural learning, to a process that reaches the former stage, but not the latter. Another interesting note is that the literature now reports a few examples of entrainment without behavioural learning, but no accounts of the opposite dissociation – behavioural learning without entrainment. Of course,

absence of evidence is not directly evidence of absence, so we should definitely take this note with quite some caution; however, the current data do support the idea that SL is not only based on separate mechanisms, but also that these mechanisms are hierarchically organized, such that perceptual learning/extraction, as tracked by neural entrainment, is necessary for memory integration, and therefore behavioural learning, to occur – but not vice versa.

The possible role of the distributional, univariate cues

One important factor to consider while interpreting these results was already mentioned above – the fact that we did not observe behavioural learning in this study does not necessarily imply that it is impossible to find it, and that neural entrainment necessarily dissociates from SL as it is typically assessed through behavioural learning. The specificities of our tasks may not have been optimally aligned to capture the learned regularities behaviourally. Tasks requiring explicit judgments may underestimate learning that remains implicit. Developing behavioural measures that better capture online or implicit processing will be crucial for bridging neural and behavioural findings.

The present study focused on frequency-based regularities – that is, a univariate, distributional cue that stands in rather stark contrast with the more classic SL approach based on associative, bivariate statistics as indexed by transitional probability (e.g., Saffran et al., 1996a; Batterink and Paller, 2017; Frost et al., 2019). Distributional cues are increasingly becoming a subject of interest in the SL community (e.g., De Rosa et al., 2022), but are still quite understudied (especially compared to associative statistics). They are very much at the core of several fundamental distinctions in important domains of cognition (e.g., the difference between function words and content words in language), and are clearly probabilistic signals that could possibly fall under the scope of the same learning system that takes care of bivariate statistics (or even *any* statistic, perhaps? See, e.g., Thiessen et al., 2013; Lavi-Rotbain and Arnon, 2021, 2022; Zhou et al., 2024).

While it is quite clear now that differences in item frequency yield neural entrainment, and therefore are encoded at the neural level, they may not be sufficient on their own to support robust behavioural learning. In other words, it is possible that the dissociation highlighted here between neural and behavioural learning is specifically related to the kind of univariate, distributional cue that we used here. Indeed, Batterink and Paller (2017) investigated the relationship between neural entrainment and SL as assessed behaviourally using the classic transitional probability design. They were able to find evidence of behavioural learning, along with

significant neural entrainment. They also found a correlation between the amount of neural entrainment and the size of the SL behavioural effect across participants. This would suggest a much tighter coupling between neural and behavioural SL effects with bivariate, associative cues. However, the correlation between entrainment and behavioural learning was only statistically significant for an RT-based, target-detection task – not for the 2AFC task that is most often used to quantify SL and that we used here. Also, and perhaps more importantly, Batterink and Paller (2017)'s design was not optimal for an individual variability analysis; the participant sample was fairly small, and there was information on the reliablity of the metrics to be correlated. So, more work is in order here, perhaps with a specific emphasis on the comparison between univariate and bivarate statistical cues.

2.5 Conclusion

In conclusion, the present study demonstrates that frequency-based regularities are robustly and rapidly encoded at the neural level, as indexed by FPVS. We also showed that, at least under the conditions tested here, neural entrainment does not automatically translate into behavioural learning. This establishes a crucial dissociation between implicit, processing based, neural encoding of statistical regularities on the one hand; and explicit, reflection based access to the learned information. These findings support a multicomponential account of statistical learning, in which different measures capture distinct stages or components of the learning process. More broadly, they underscore the importance of combining neural and behavioural approaches to fully characterise learning, and caution against strong interpretations of neural entrainment as direct evidence of behaviourally relevant knowledge.

Chapter 3

Neural and Behavioural Signatures of Statistical Learning in Aging: Evidence from Fast Periodic Visual Stimulation

Claudia Ruzza[1]
Davide Crepaldi[2]

[1]Department of Cognitive Neuroscience
Scuola Internazionale Superiore di Studi Avanzati (SISSA)
Trieste, Italy

[2]Dipartimento di Scienze del Sistema Nervoso e del Comportamento
Università di Pavia
Pavia, Italy

This chapter is adapted from a manuscript currently under revision for publication in *Psychology of Aging*. The pre-print can be found on PsyArXiv at https://osf.io/preprints/psyarxiv/u9xj5_v1.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant Agreement No 679010 STATLEARN ERC-2015-STG/ERC-2015-STG.

Abstract

Statistical learning (SL) enables the extraction of regularities from structured input and is thought to support visual word processing. Using a Fast Periodic Visual Stimulation (FPVS) paradigm, we investigated whether sensitivity to frequency-based regularities is preserved in healthy aging and how such sensitivity is expressed at the neural and behavioural level. Fifty-one older adults and a comparison group of young adults completed EEG and behavioural tasks, alongside a classic visual statistical learning (VSL) task. Results revealed robust neural discrimination of frequency-based categories in older adults, indicating preserved sensitivity to statistical regularities, although responses were reduced in amplitude compared to young adults. Behaviourally, evidence of learning was weaker and more variable across tasks, and no reliable relationship emerged between neural and behavioural indices. In contrast, performance in the VSL task declined with age and was positively associated with behavioural discrimination performance. Together, these findings suggest a dissociation between preserved neural sensitivity to regularities and their behavioural expression in healthy aging.

I would like to thank Maria Ktori for guidance in EEG testing; Mara De Rosa and Yamil Vidal for their help with experiment implementation and analysis coding; and Elena Greatti, Emma Peri and Emma Coba for their assistance with data collection.

3.1 Introduction

Statistical learning (SL) refers to the ability to implicitly extract regularities from structured input (Bogaerts et al., 2021), such as the co-occurrence and distribution of elements over time or space. This mechanism enables learners to segment continuous sensory streams and identify meaningful units without explicit instruction (Aslin, 2017; Frost et al., 2019). A seminal demonstration comes from language research, showing that learners can rely on transitional probabilities (TPs) between syllables to segment continuous speech into word-like units (Saffran et al., 1996a). Comparable abilities have been observed in the visual domain, where participants extract regularities from sequences of shapes or objects (Kirkham et al., 2002; Turk-Browne et al., 2005; Siegelman and Frost, 2015).

Statistical Learning as a foundation for structured representations in Vision and Language

SL extends beyond sensitivity to sequential co-occurrence and also supports the extraction of distributional structure. Learners can use frequency distributions to form categorical representations and establish category boundaries; for example, exposure to bimodal versus unimodal distributions of phonetic tokens leads to the formation or merging of phoneme categories (Maye et al., 2002). Similar effects have been demonstrated in vision, where observers derive categories from statistical distributions of low-level features (Rosenthal et al., 2001). Together, these findings indicate that SL operates over both univariate and bivariate statistics in the input, supporting the emergence of structured representations in different modalities. Given its central role in learning and perception, SL provides a key framework for investigating how cognitive processes are maintained or altered across the lifespan.

Visual word processing is increasingly understood as relying on domain-general mechanisms akin to statistical learning, alongside specialized, language-specific systems. According to the Local Combination Detector (LCD) model, visual recognition is supported by a hierarchical organization of the ventral visual stream, in which neurons become progressively tuned to increasingly complex feature combinations through exposure to environmental regularities (Dehaene et al., 2005). Within this framework, sensitivity to orthographic structure emerges from experience with the statistical co-occurrence of visual features, leading to graded neural selectivity for familiar patterns along the occipitotemporal cortex (Vinckier et al., 2007). Crucially, sensitivity to such regularities reflects a broader capacity for visual SL that extends beyond linguistic stimuli. Implicit learning of co-occurrence patterns has

been observed for artificial symbols, objects, and low-level visual features, indicating that the same mechanisms operate across different types of visual input (Vidal et al., 2021). Similar sensitivities are also present in developing readers (Cassar and Treiman, 1997; Pacton et al., 2001), further supporting the view that visual word recognition builds on general-purpose learning mechanisms. Together, these findings position SL as a core mechanism underlying sensitivity to orthographic structure and highlight its relevance for understanding how visual and linguistic processing are shaped by experience across the lifespan.

Statistical Learning in Aging: Preservation, Constraints, and Variability

Research on SL in healthy older adults provides a mixed picture, with several studies suggesting a substantial preservation of core SL mechanisms alongside evidence for age-related changes. On the one hand, multiple studies report that older adults retain the ability to extract regularities from the environment, performing comparably to younger adults in tasks such as cross-situational word learning (Bulgarelli et al., 2021), implicit motor sequence learning (Johannsen et al., 2022), artificial grammar learning (Schevenels et al., 2023), and contextual cueing paradigms (Nagy et al., 2024). Similarly, speech segmentation based on statistical cues appears largely preserved under low-demand conditions (Palmer et al., 2018), suggesting that fundamental mechanisms for tracking co-occurrence and distributional information remain functional across the lifespan.

On the other hand, evidence from other studies observes age-related differences, especially under specific conditions. Older adults often show slower learning rates and lower asymptotic performance (Ge et al., 2025), reduced ability to exploit more fine-grained or complex statistical structures (e.g., discriminating words from part-words in speech segmentation; Palmer et al., 2018), and impairments in learning task-irrelevant regularities such as distractor suppression (Lega et al., 2023). Additionally, deficits have been reported in offline consolidation of learned sequences (Murray and Prashad, 2025), and in modality-specific contexts, particularly in the auditory domain (Schevenels et al., 2023).

These findings suggest that while basic SL abilities may be preserved, their efficiency is modulated by cognitive demands, task complexity, and domain-general factors such as working memory and processing speed. Overall, the literature does not provide a clear consensus: statistical learning in healthy aging appears to be neither uniformly preserved nor uniformly impaired, but rather dependent on task characteristics and the cognitive resources required to support learning.

FPVS as an implicit tool to study Statistical Learning in Aging

One possible reason for the heterogeneous findings observed in the statistical learning literature on aging is that many traditional SL tasks rely not only on sensitivity to regularities, but also on additional cognitive functions such as explicit decision-making, working memory, attentional control, metacognitive judgments, and strategic responding, all of which are known to be affected by aging. In this context, the Fast Periodic Visual Stimulation (FPVS) approach provides a powerful tool to investigate statistical learning more directly at the neural level, while minimizing the contribution of these additional cognitive demands; it leverages neural entrainment to periodic visual input (Norcia et al., 2015) and employs frequency tagging to isolate category-selective responses (Rossion, 2014).

In this paradigm, stimuli are presented at a rapid base frequency, while deviant (oddball) items are inserted at regular intervals at a lower frequency. This design elicits steady-state visual evoked potentials (SSVEPs; Regan 1966), with distinct responses at both base and oddball frequencies and their harmonics. Critically, responses at the oddball frequency provide an implicit and objective neural index of discrimination between stimulus categories, independent of explicit task demands. Owing to its high sensitivity and task-free nature, FPVS has been widely used to investigate perceptual categorization across domains such as faces, objects, and written words in both typical (Liu-Shuang et al., 2014; Lochy et al., 2015; Rossion et al., 2015; Retter and Rossion, 2016; Stacchi et al., 2019; Coll et al., 2019; Volfart et al., 2021; Bánki et al., 2022; Pescuma et al., 2022) and atypical populations (Lochy et al., 2016; Stothart et al., 2017; Vettori et al., 2019; Stothart and Smith, 2020; Van der Donck et al., 2020; Rossion, 2020; David et al., 2023b; Nijhof et al., 2024; Stothart et al., 2025; Davidson, 2025; Hermann et al., 2025; Rutkowski et al., 2025). Moreover, the FPVS oddball design shares key properties with paradigms probing automatic novelty detection, such as mismatch negativity (Näätänen et al., 2007; Kimura et al., 2011; Stefanics et al., 2015; Kremláček et al., 2016), making it particularly suitable for studying sensitivity to statistical regularities. These non-invasive neural entrainment techniques are increasingly being applied to the study of special populations, including developmental and clinical groups, where they offer a promising avenue for investigating changes in perceptual and learning mechanisms across the lifespan (Nurdal, 2022).

Beyond discrimination between pre-existing stimulus categories, however, an important and largely unexplored question is whether the FPVS oddball paradigm can also capture the emergence of a novel distinction based purely on the statistical structure of the input. Indeed, the FPVS oddball design provides a context that may naturally foster implicit statistical learning. Similar to paradigms used to investi-

gate automatic novelty detection and mismatch responses (Näätänen et al., 2007; Kimura et al., 2011; Kremláček et al., 2016; Stefanics et al., 2015), oddball stimuli are presented as infrequent events embedded within a stream of more frequent stimuli. Differences in the frequency with which oddball and base stimuli occur are therefore inherent to the structure of the paradigm itself. Importantly, this difference in frequency of occurrence constitutes a source of distributional information on the basis of which oddball stimuli (i.e., a group of infrequent items) could potentially be differentiated from base stimuli (i.e., a group of frequent items), independently of any pre-existing categorical distinction between them. This raises a fundamental question from a statistical learning perspective: can the brain rapidly and implicitly learn such frequency-based regularities, and is a difference in frequency of occurrence alone sufficient to elicit a neural discrimination response?

To address these questions, De Rosa et al. (2022) introduced a modification of the classic FPVS oddball paradigm in which all stimuli within a sequence belonged to the same category (e.g., words embedded among other words). Under these conditions, oddball and base stimuli could not be distinguished on the basis of any pre-existing categorical, semantic, or perceptual difference. Instead, the only systematic distinction between the two groups was their relative frequency of occurrence within the stimulation stream: oddball items appeared four times less frequently than base items. Assignment of individual stimuli to the oddball and base categories was randomized across participants, ensuring that any neural discrimination response could only emerge from the statistical structure of the input. Remarkably, this manipulation was sufficient to elicit a reliable neural response at the oddball frequency, demonstrating that the brain can rapidly and implicitly differentiate between groups of stimuli solely on the basis of their frequency of occurrence. Neural entrainment at the oddball frequency therefore indexed the emergence of a novel distinction generated by exposure to the distributional properties of the stimulation stream, providing direct evidence that FPVS can be used to investigate statistical learning mechanisms in the absence of pre-existing category differences.

Within the context of healthy aging, FPVS has begun to provide important insights into how neural sensitivity to statistical regularities is preserved or altered across the lifespan (Nurdal, 2022). Because it offers an implicit and task-independent measure of processing, FPVS is particularly well suited to older populations, in whom behavioural measures are often confounded by general slowing, increased variability, and reliance on explicit decision-making. Evidence from FPVS and related neural entrainment paradigms indicates that older adults can exhibit preserved neural discrimination of visual and linguistic categories, particularly when stimulation parameters (e.g., slower presentation rates) are adapted to age-related changes in temporal

processing (Sáringer et al., 2024; Peykarjou et al., 2024; Marchive et al., 2025). At the same time, converging findings point to qualitative changes in neural responses, including reduced amplitude, decreased selectivity, and increased neural noise in frequency-tagged signals (Mao et al., 2025; Huizeling et al., 2021). These findings highlight the value of FPVS as a sensitive and non-invasive approach to disentangle preserved versus age-dependent cognitive functions. However, none of these results refer to the FVPS paradigm specifically adapted to isolate statistical learning, and therefore they carry little information on how this cognitive skill changes with age.

Together, these findings highlight the value of FPVS as a sensitive and non-invasive approach to disentangle preserved versus altered components of statistical learning in aging, and to better characterize neural processing across the lifespan.

Neural and Behavioural Indices of Statistical Learning: Convergence, Dissociation, and Aging

A substantial body of research combining neural entrainment paradigms (e.g., FPVS and RSVP) with behavioural measures indicates that the relationship between neural sensitivity to structured input and overt behavioural performance is complex and not always straightforward. While frequency-tagged responses provide a powerful and objective index of learning-related neural processes (Rossion, 2014; Norcia et al., 2015), many studies do not include behavioural validation, and when both measures are assessed, findings are mixed: research combining neural entrainment paradigms (e.g., FPVS and RSVP) with behavioural measures reveals a heterogeneous pattern of results.

The unclear relationship between neural entrainment and behavioural learning also affects the research on older adults. On the one hand, several studies report converging preservation at both neural and behavioural levels, showing that the elderly can exhibit comparable neural entrainment responses and intact performance in perceptual categorization or implicit learning tasks, particularly under low-demand conditions or at slower stimulation rates (Sáringer et al., 2024; Marchive et al., 2025). In contrast, other studies combining neural entrainment with behavioural measures often reveal dissociations between neural and behavioural indices. For instance, older adults may show preserved or even robust neural sensitivity to regularities (e.g., SSVEPs or frequency-tagged signals) despite reduced behavioural performance, such as slower reaction times, lower accuracy, or increased attentional blink effects in RSVP paradigms, suggesting that implicit learning mechanisms remain relatively preserved while their behavioural expression becomes less efficient (Sklenar and Mienaltowski, 2019; Donohue et al., 2019; Huizeling et al., 2021). Conversely, cases have

also been reported in which behavioural performance appears preserved despite altered neural signatures, including reduced neural selectivity, increased neural noise, or less efficient oscillatory modulation, indicating age-related changes in underlying processing mechanisms without overt behavioural deficits (Walker et al., 2022). Finally, some studies document age-related differences at both neural and behavioural levels, with older adults showing reduced entrainment amplitude, altered frequency tuning, and less efficient neural modulation, alongside behavioural slowing, reduced accuracy under high cognitive load, and impairments in attentional control and working memory (Mao et al., 2025). Such variability raises the possibility that neural entrainment and behavioural performance rely on partially dissociable mechanisms, particularly in aging populations.

Crucially, regardless of whether neural entrainment and behavioural performance converge or diverge, studies specifically examining their relationship in aging populations similarly report mixed and inconsistent findings. Some studies report clear convergence, with robust neural responses accompanied by strong behavioural performance across domains such as face perception, object processing, and word recognition (Liu-Shuang et al., 2014; Lochy et al., 2015; Batterink and Paller, 2017; Dzhelyova et al., 2020; Moser et al., 2021b).

However, other work highlights dissociations, with neural entrainment emerging rapidly or reliably even when behavioural evidence is weak, delayed, or absent, and correlations between neural and behavioural measures ranging from strong to negligible depending on task alignment and learning stage (Stothart and Smith, 2020; Rawal and Tseng, 2021; Sznabel et al., 2023; Sáringer et al., 2024).

These discrepancies suggest that neural and behavioural tasks may tap into partially distinct components of statistical learning, with neural measures capturing early or implicit representations that are not always accessible at the behavioural level (Frost et al., 2015).

Heterogeneous Aging Trajectories in Statistical Learning

Research on statistical learning increasingly suggests that sensitivity to regularities does not decline uniformly across the lifespan, but instead shows substantial variability across older adults (Bogaerts et al., 2026). While some individuals exhibit relatively preserved implicit learning abilities, others show reduced sensitivity to regularities or weaker behavioural expression of learning (Kaufman et al., 2010; Siegelman and Frost, 2015). This heterogeneity has become an important consideration in the study of SL and aging, as it suggests that age-related changes may reflect multiple trajectories rather than a single pattern of decline.

More generally, variability in SL performance has been linked to a range of cognitive and linguistic outcomes, including vocabulary size, syntactic processing, and literacy skills (Arciuli and Simpson, 2012; Kidd and Arciuli, 2016; Frost et al., 2013a), and may even predict later language development (Shafto et al., 2012; Ellis et al., 2014). At the same time, SL performance is known to depend partly on broader cognitive factors such as attention, working memory, and processing speed (Conway et al., 2010; Misyak and Christiansen, 2012), which themselves are subject to substantial age-related changes. Moreover, correlational findings across SL tasks are often inconsistent (Schmalz et al., 2019; Siegelman, 2020a), suggesting that different paradigms may capture partially distinct components of learning.

Within aging populations specifically, behavioural studies report considerable variability in learning rate, asymptotic performance, and sensitivity to different types of regularities (Bulgarelli et al., 2021; Ge et al., 2025). At the neural level, age-related variability has been associated with differences in oscillatory dynamics, including theta and alpha activity, as well as in more complex measures such as phase–amplitude coupling (Strzelczyk et al., 2025; Zając-Lamparska et al., 2025). In addition, studies combining neural and behavioural measures indicate that older adults may differ in the extent to which neural sensitivity to regularities is reflected in overt behavioural performance (Donohue et al., 2019; Morris et al., 2025a). Some individuals show relatively preserved neural responses despite weaker behavioural performance, whereas others exhibit parallel changes across both levels of measurement (Orpella et al., 2022; Verty et al., 2024; Huizeling et al., 2021).

Together, these findings suggest that statistical learning in aging is characterized by heterogeneous trajectories across individuals and measures. Examining both neural and behavioural indices may therefore provide a more comprehensive understanding of how sensitivity to regularities is preserved, altered, or behaviourally expressed across the aging process.

Research Questions

The present study investigates whether such sensitivity is preserved in healthy aging and how it is expressed across neural and behavioural levels.

Specifically, we address four main questions: (1) whether neural sensitivity to frequency-based regularities is preserved in older adults and differs from that observed in younger adults; (2) whether behavioural tasks reveal evidence of learning or memory for the frequency manipulation in aging; (3) whether neural and behavioural measures provide converging or dissociable evidence of learning; and (4) how FPVS-based measures relate to performance in a standardized Visual Statistical Learning (VSL) task. Finally, we examine whether chronological age is associated with variability in neural and behavioural measures within the older adult group.

Together, these questions aim to clarify how sensitivity to statistical regularities is preserved and behaviourally expressed across healthy aging.

3.2 Materials and Methods

The present study examined statistical learning in aging by combining electrophysiological and behavioural approaches. Building on De Rosa et al. (2022) FPVS paradigm, we measured neural sensitivity to frequency-based regularities in streams of words and pseudofonts. Behavioural performance was assessed through 2AFC tasks, alongside an independent Visual Statistical Learning (VSL) task. Because traditional SL paradigms often show substantial measurement error and limited reliability in individual-differences research, we employed the VSL task developed by Siegelman et al. (2017b), which was specifically designed to provide improved psychometric properties and more stable estimates of visual statistical learning performance. This multimethod design allowed us to examine the relationship between neural and behavioural indices of learning and to investigate how these measures are modulated by healthy aging.

3.2.1 Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures used in this study. The study hypotheses and analytic plan were not preregistered.

De-identified data, analysis code, and study materials are publicly available through the Open Science Framework (OSF) at `https://osf.io/n8fzu`.

Analyses were conducted using MATLAB R2021b (The MathWorks Inc., 2021), EEGLAB (Delorme and Makeig, 2004), and R (R Core Team, 2023), including the `lme4` package for mixed-effects modelling.

3.2.2 Participants

Fifty-seven paid volunteers (22 males; age: M = 68.65, range: 55–83, SD = 6.81) participated in the experiment after providing written informed consent. Six participants were excluded due to technical issues during recording. The sample of younger participants (N = 38; 13 males; age: M = 23.16, range: 19–33, SD = 3.07) comes from a previous study of ours (Ruzza and Crepaldi, 2026); they allowed us to use their prior experimental data and agreed on completing the VSL task.

Data collection took place between 2023 and 2025 in Trieste, Italy. All participants were right-handed White Italian native speakers with normal or corrected-to-normal vision and no reported history of linguistic, cognitive, or neurological impairments. Participants were unaware of the experiment's purpose until the end

of the testing session. The study was conducted after receiving approval from the SISSA Ethics Committee (Protocol Nr. 9001-III/13).

3.2.3 Stimuli

The stimuli employed in the FPVS tasks were the same employed in Ruzza and Crepaldi (2026)'s Experiment 2 and comprised two categories: words and pseudofont strings, each consisting of 16 five-element items, all selected following the criteria of De Rosa et al. (2022). Word stimuli and word foils were selected controlling for frequency, phon-grapheme overlapping ($\leq$2 characters) and semantic distance; pseudofont strings were constructed using the BACS-2 serif artificial script (Vidal et al., 2017) controlling for grapheme overlapping (overlapping $\leq$2 characters) and orthographic similarity. All stimuli are reported in the Supplementary Materials section (see Table S5 for words stimuli and Table S6 for pseudofont stimuli). Stimuli for the VSL study were directly taken from Siegelman et al. (2017b) and included 16 complex visual shapes (taken from Fiser and Aslin 2001); see Supplementary Materials section, Table S8.

3.2.4 Experimental Paradigm and Procedure

The testing session was organized as follows. Participants were firstly asked to complete the MoCA test and the VSL test. They were subsequently instructed and prepared for the EEG session and asked to complete the FPVS and related behavioral tasks. See the outline of experimental procedure in Fig.3.1.

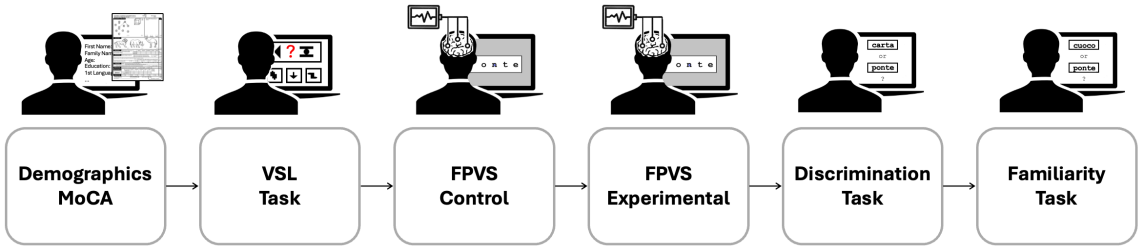


Figure 3.1: **Overview of the experimental design and procedure.** After completing the demographics and language background questionnaires and assessed their cognitive skills using the MoCA Test, participants were assigned to one of two groups that differed in the order of modality exposure: Visual–Auditory (VA) or Auditory–Visual (AV). The experiment consisted of two phases. In each phase, participants completed a familiarization (FAM) in one modality (visual in blue, auditory in red), followed by four test tasks: a two-alternative forced-choice task (2AFC) and a serial imitation and cued recall task (SICR) administered first in the same modality as familiarization and then in the other modality. In the second phase, the familiarization modality was reversed, and the same sequence of within- and cross-modal test tasks was repeated. This counterbalanced design controlled for order effects of modality exposure across participants. Adapted from (Ruzza and Crepaldi, 2026)

The Montreal Cognitive Assessment (MoCA; Nasreddine et al. 2005) is a screening tool designed to assess global cognitive function, including memory, attention, language, and executive skills (see Supplementary Materials section, Table S7). It was chosen for its specificity in identifying mild forms cognitive impairments (Pinto et al., 2019) and was used in this study to ensure participants were cognitively unimpaired. Participants scored an average of 27.71/30 (SD = 1.92), confirming their cognitive suitability for the study.

The VSL task designed by Siegelman et al. (2017b) was employed as a measure of participants' ability to learn and generalize visual patterns. In this task, 15 shapes were randomly organized for each participant into 8 triplets, with two types of transitional probabilities (TPs): four triplets with TPs of 0.33 and four with TPs of 1. During the familiarization phase, triplets were presented in a random order for 10 minutes; each of them appeared 24 times in the stimulus stream. Participants were instructed to watch the stream in preparation for the test phase, which consists of two tasks. The pattern recognition task required participants to select the triplet they were previously familiarized to. Items varied in terms of position violations and distractors number (two or four alternatives). Subsequently, participants completed a pattern completion task, where they were asked to select among 3 options the shape completing the triplet. The total score ranged from 0 to 42. See Fig.3.2 for an outline of the VSL task.

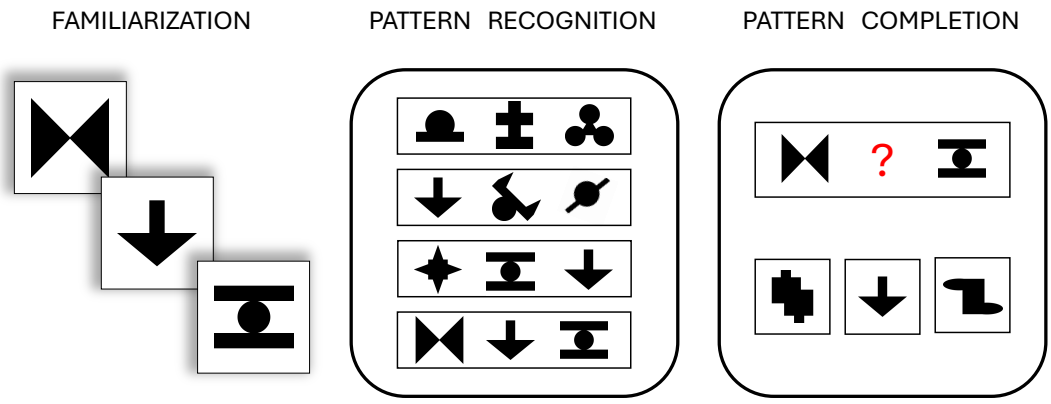


Figure 3.2: **Visual Statistical Learning (VSL) task design.** The VSL task consisted of three phases. During the familiarization phase (left panel), participants were exposed to a continuous stream of visual symbols organized into structured triplets. In the pattern recognition phase (middle panel), participants were asked to identify the previously learned triplet among alternative combinations of symbols. In the pattern completion phase (right panel), participants viewed incomplete triplets and selected the missing symbol from multiple alternatives. Adapted from (Siegelman et al., 2017b)

The FPVS-EEG study followed the same structure described in Ruzza and Crepaldi (2026), comprising both an experimental and a control condition, paired with two behavioural two-alternative forced choice (2AFC) tasks: the Discrimination and Familiarity tasks. See Fig. 3.3 for an overview of the FPVS paradigm and associated behavioural tasks.

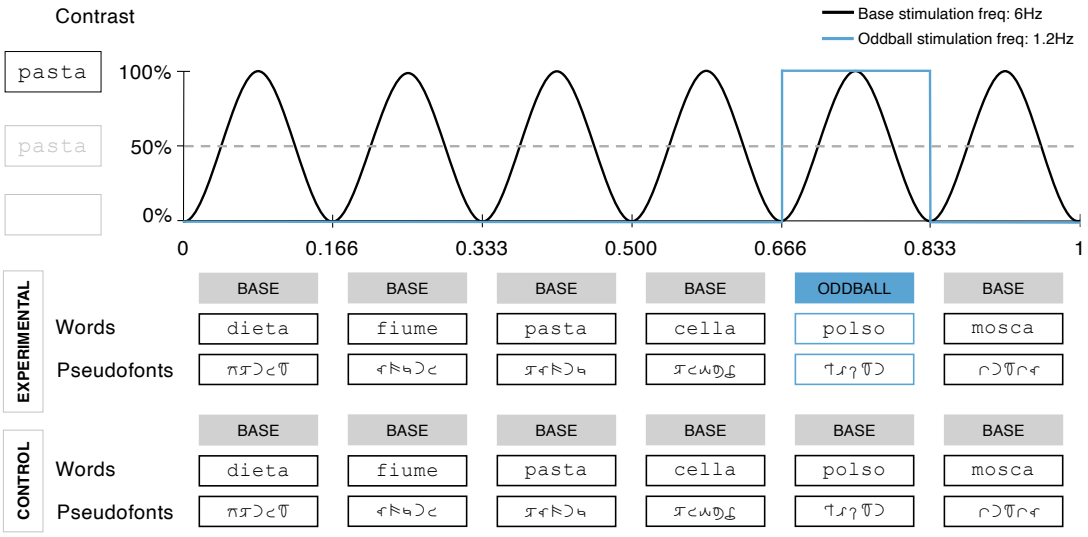


Figure 3.3: **FPVS experimental paradigm.** In each stimulation block (1 min), stimuli were presented at a frequency of 6Hz with sinusoidal contrast modulation, reaching maximum contrast 83 ms after onset and fading out completely after 166 ms, corresponding to one full cycle. In the experimental condition, streams of base stimuli were interspersed every fifth item with oddball stimuli (1.2 Hz). Unlike traditional oddball paradigms, both base and oddball stimuli belonged to the same category (words or pseudofonts) and differed only in presentation frequency, with oddballs appearing four times less frequently than bases. In the control condition, the statistical structure was disrupted and all stimuli were presented with equal probability in random order at 6 Hz. The figure illustrates both experimental and control conditions for word and pseudofont stimuli. Taken from Ruzza and Crepaldi (2026).

The experimental condition combined the FPVS technique with an oddball paradigm in which 16 stimuli were randomly assigned for each participant to either the base (N = 8) or oddball category (N = 8). Unlike traditional oddball paradigms, both stimulus types belonged to the same visual category (e.g., word oddballs embedded within streams of word bases; see De Rosa et al. 2022), differing exclusively in their presentation frequency. Base stimuli were presented at 6Hz, while oddballs appeared every fifth item, corresponding to a frequency of 1.2Hz. Each stimulation block consisted of 360 stimuli (6 items/s), including 72 oddball presentations (9 repetitions per oddball token) and 288 base presentations (36 repetitions per base token).

In the control condition, no base/oddball distinction was present and all stimuli were presented with equal probability in pseudorandom order (mean repetitions per stimulus = 22.5, SD < 2). This condition was designed to disrupt the statistical

structure of the stimulation stream and to verify that neural responses observed at the oddball frequency in the experimental condition were specifically driven by the frequency manipulation rather than by low-level visual properties or general periodic stimulation.

Participants completed 5 stimulation blocks for each stimulus type (words and pseudofonts) and condition (experimental and control), with a 30-second break between blocks. Throughout the FPVS tasks, participants were instructed to fixate on a central cross and press a button whenever its colour changed (red to blue or vice versa), which occurred six times per trial. This orthogonal task was included to ensure sustained attention during visual stimulation.

Participants' performance in the color-change detection task was accurate in both the control condition (Accuracy: M = 96.9%, SD = 0.05; Reaction Time: M = 545ms, SD = 0.15) and the in the experimental condition (Accuracy: M = 94.7%, SD = 0.05; Reaction Time: M = 628ms, SD = 0.21) and was comparable across the different stimuli types (Control task: Words - Accuracy = 96.54%, SD = 0.05, Reaction Time: M = 491ms, SD = 0.5; Pseudofonts - Accuracy = 97.2%, SD = 0.04, Reaction Time: M = 493ms, SD = 0.5; Experimental task: Words - Accuracy = 93.9%, SD = 0.06, Reaction Time: M = 484ms, SD = 0.5; Pseudofonts - Accuracy = 95.5%, SD = 0.05, Reaction Time: M = 488s, SD = 0.5).

After the FPVS sessions, participants completed two behavioural tasks. In these 2AFC tasks, participants were instructed to choose between two alternatives by pressing two predefined keys on a keyboard. A 2-second timeout was included to encourage instinctive responses.

The Discrimination task assessed participants' ability to differentiate base and oddball stimuli from foils and evaluated whether exposure differences between base and oddball stimuli influenced discrimination. The task included base vs. foil and oddball vs. foil trials. Eight foils per stimulus type (words and pseudofonts) were chosen, allowing each stimulus to pair with any foil. The task comprised three blocks of 32 trials each (N = 8 base vs. foil; N = 8 oddball vs. foil, for both stimulus type). Participants were asked to indicate which item they had seen during FPVS stimulation phase.

The Familiarity task assessed whether participants' had a familiarity preference according to frequency exposure to stimuli. Bases were randomly paired with oddballs across three blocks, ensuring no base was repeatedly paired with the same oddball. Each block included 16 trials (8 x stimulus type: words and pseudofonts). Participants were asked to identify the item they perceived as being presented more frequently or as being more familiar to them.

Participants were tested individually in a soundproof room, seated about 80 cm

away from a 27" BenQ XL2720Z monitor (screen resolution of 1920 x 1080 pixels and a refresh rate of 60Hz). Stimuli were displayed using PsychToolbox-3 (Brainard and Vision, 1997) on MATLAB R2021b for the FPSV tasks, PsychoPy (Peirce, 2007) for the 2AFC tasks and jsPsych (De Leeuw, 2015) for the VSL task.

3.2.5 EEG Data Acquisition and Pre-processing

During the FPVS tasks, EEG activity was recorded using an Active Two Biosemi system with 64 active electrodes arranged on an elastic cap following the radial-ABC system. Data were sampled at 256 Hz, with additional electrodes for reference (Common Mode Sense, CMS) and grounding (Driven Right Leg, DRL), maintaining electrode offset magnitudes below 20 mV. Pre-processing was conducted in MATLAB with custom scripts and EEGLAB toolbox (Delorme and Makeig, 2004). The pre-processing pipeline included band-pass filtering (0.1–100 Hz), visual inspection of channels for artifacts (time series and spectra), and Independent Component Analysis to remove eye and movement artifacts. Rejected channels were interpolated and data was segmented into 60-second epochs, resulting in five 60-second trials per participant for each stimulus type (word and pseudofont) and each condition (control and experimental).

3.2.6 Data Analysis

Data analysis comprised both electrophysiological (EEG) and behavioural measures, which were analyzed separately using complementary approaches.

EEG Data Analysis

The EEG data analysis followed the steps of previous similar studies (De Rosa et al., 2022; Ruzza and Crepaldi, 2026). Trials were averaged in the time domain for each condition and participant to minimize EEG activity not phase-locked to stimulation. A Fast Fourier Transform (FFT) was then applied to generate normalized amplitude spectra by dividing the absolute FFT values by the number of data points.

A local baseline correction was applied to account for noise variations across the frequency spectrum: we subtracted the mean of the surrounding 20 bins (10 on each side) from each frequency bin, as in previous studies (Dzhelyova and Rossion, 2014; De Rosa et al., 2022; Ruzza and Crepaldi, 2026). Base and oddball responses were defined as the sum of baseline-corrected amplitudes at the base (6 Hz) and oddball (1.2 Hz) frequencies, along with their first three harmonics (i.e., base: 12, 18, 24 Hz; oddball: 2.4, 3.6, 4.8 Hz).

The base response was analyzed in an occipito-parietal region comprising 5 electrodes centered around Oz (A29 in Biosemi's ABC system), where steady-state visual evoked potentials (SSVEPs) are maximal (Norcia et al., 2015). Oddball responses were analyzed in two predefined regions of interest (ROIs) based on prior studies using similar stimuli (De Rosa et al., 2022; Lochy et al., 2015). These ROIs consisted of six occipito-temporal electrodes clustered around PO7 (A25) on the left and PO8 (B30) on the right hemisphere.

To evaluate the significance of the neural responses, ROI-averaged amplitudes were subjected to one-sample t-tests against the noise level. Direct comparisons between younger and older adults were conducted using independent-samples t-tests. Effect sizes (Hedges' g) and 95% confidence intervals are reported throughout.

Behavioural Data Analysis

Performance on the behavioural tasks was assessed using accuracy as the primary metric. For the Discrimination task, accuracy was calculated as the proportion of correct responses in the base-vs.-foil and oddball-vs.-foil conditions. In the Familiarity task, we measured the proportion of trials on which participants selected oddball stimuli as more familiar. Performance in the VSL task was analysed following the original analysis pipeline proposed by Siegelman et al. (2017b). Accuracy was expressed as the percentage of correct responses, and overall VSL performance was computed by combining scores from the pattern recognition and pattern completion tasks.

Behavioural data were analysed in R (R Core Team, 2023). Trial-level accuracy data were analysed using generalized linear mixed-effects models (GLMMs) with a binomial error distribution and logit link function, implemented in the `lme4` package (Bates et al., 2015). Fixed effects varied according to the specific analysis and included factors such as Stimulus Type, Frequency Category, and Age Group. Random intercepts for participants were included to account for repeated observations. In addition, one-sample t-tests were used to assess performance against chance levels, and independent-samples t-tests were used to compare younger and older adults. Effect sizes and 95% confidence intervals are reported throughout.

Reliability and Correlational Analyses

To assess the reliability of our measures, we computed Cronbach's alpha coefficients. This approach was adopted to maintain consistency with previous work examining individual differences in statistical learning, including Ruzza and Crepaldi (2026) and studies using the classic VSL paradigm (e.g., Siegelman et al., 2017b). Cronbach's

alpha provides an estimate of the extent to which performance is consistent across trials within a task and is widely used in the statistical learning literature when evaluating the suitability of behavioural measures for individual-differences analyses.

Following the criterion adopted in our previous work, only measures showing reliability estimates above .50 were considered suitable for primary correlational analyses. Measures that did not meet this threshold were excluded from confirmatory analyses because low reliability is expected to attenuate observed correlations and complicate their interpretation. Given that several measures approached the reliability criterion, additional exploratory correlational analyses were conducted using a more liberal threshold of $\alpha \geq .40$. These exploratory analyses are reported separately and should be interpreted with caution.

Associations among reliable behavioural and neural measures, as well as demographic variables (e.g., age), were examined using Pearson correlation coefficients.

3.3 Results

We organized the results in three main stages. First, we report group-level findings for the neural and behavioural tasks, including comparisons between younger and older adults. For the FPVS paradigm, we begin with the control condition to validate basic visual entrainment and ensure that responses at the oddball frequency reflect sensitivity to frequency-based regularities rather than low-level stimulation properties, before turning to the experimental condition. We then present behavioural results from the discrimination, familiarity, and VSL tasks.

Second, we report reliability analyses for all neural and behavioural measures, including internal consistency and split-half reliability estimates. Following the criteria adopted in our previous work, only measures meeting the predefined reliability threshold were subsequently included in correlational analyses.

Finally, we report analyses examining associations between neural and behavioural measures, as well as relationships with age within the older adult group.

3.3.1 Group-Level Analyses

We first report group-level analyses separately for neural and behavioural measures in young and elderly participants. For each task, we first assessed whether performance or neural responses significantly differed from chance or noise level within each age group, and then directly compared young and elderly participants to examine potential age-related differences in statistical learning and frequency-based processing.

3.3.1.1 FPVS Control Task

The FPVS Control task was designed to validate basic visual entrainment and ensure that any response observed at the oddball frequency in the experimental condition could not be attributed to low-level stimulation properties alone. Specifically, stimuli were presented periodically without any frequency manipulation between categories, allowing us to assess neural synchronization to the visual stream in the absence of statistical regularities.

In younger adults, the FPVS Control task elicited a robust response at the base stimulation frequency of 6 Hz, localized in the occipito-temporal region and significantly above the level of noise for both words ($t(37) = 15.99$, $p < 0.0001$, $g = 2.44$ [95% CI: 1.835 - 3.034]) and pseudofonts ($t(37) = 16.7$, $p < 0.0001$, $g = 2.55$ [95% CI: 1.923 - 3.162]), indicating effective synchronization of the visual system with the periodic visual stimulation. No significant response was observed at the oddball stimulation frequency of 1.2 Hz for either words (Left ROI: $t(37) = 0.61$, $p = 0.571$, $g = 0.09$ [95% CI: -0.208 - 0.392]; Right ROI: $t(37) = 3.745$, $p = 0.057$, $g = 0.571$ [95% CI: 0.256 - 0.892]) or pseudofonts (Left ROI: $t(37) = 0.39$, $p = 0.7$, $g = 0.056$ [95% CI: -0.241 - 0.358]; Right ROI: $t(37) = -0.858$, $p = 0.4$, $g = -0.131$ [95% CI: -0.43 - 0.17]).

Similarly, in older adults, response at the base stimulation frequency of 6 Hz was localized in the occipito-temporal area and was significantly above the level of noise for both stimulus types (Words: $t(51) = 10.2$, $p < 0.001$, $g = 1.41$ [95% CI: 1.025–1.796]; Pseudofonts: $t(51) = 12.4$, $p < 0.001$, $g = 1.72$ [95% CI: 1.286–2.145]), which indicates effective synchronization of the visual system with the periodic visual stimulation. No significant response was observed at the oddball stimulation frequency for either words (Left ROI: $t(51) = 1.59$, $p = 0.12$, $g = 0.22$ [95% CI: -0.055–0.495]; Right ROI: $t(51) = -0.0124$, $p = 0.99$, $g = -0.002$ [95% CI: -0.274–0.270]) or pseudofonts (Left ROI: $t(51) = 0.215$, $p = 0.831$, $g = 0.03$ [95% CI: -0.242–0.301]; Right ROI: $t(51) = 1.1784$, $p = 0.224$, $g = 0.163$ [95% CI: -0.111–0.436]).

Direct comparisons between younger and older adults did not reveal significant group differences in the FPVS Control task for either the base stimulation frequency or the oddball stimulation frequency across stimulus types (all ps $> .05$).

3.3.1.2 FPVS Experimental Task

The FPVS Experimental task was designed to assess neural sensitivity to frequency-based statistical regularities. Specifically, base and oddball stimuli were presented at different occurrence frequencies within a continuous rapid visual stream, allowing us to examine whether participants showed selective neural responses to the less frequent category at the oddball stimulation frequency.

In younger adults, the FPVS Experimental task elicited a significant response at the base stimulation frequency of 6 Hz in the occipito-temporal area for both stimulus types (Words: $t(37) = 16.1$, $p < 0.0001$, $g = 2.33$ [95% CI: 1.87 - 2.91]; Pseudofonts: $t(37) = 15.8$, $p < 0.0001$, $g = 2.44$ [95% CI: 1.83 - 3.04]). A significant response to the oddball stimulation frequency of 1.2 Hz was also observed in the predefined ROIs for both words (Left ROI: $t(37) = 4.66$, $p < 0.0001$, $g = 0.72$ [95% CI: 0.38 - 1.06]; Right ROI: $t(37) = 3.39$, $p = 0.002$, $g = 0.52$ [95% CI: 0.20 - 0.84]) and pseudofonts (Left ROI: $t(37) = 3.52$, $p = 0.001$, $g = 0.54$ [95% CI: 0.22 - 0.87]; Right ROI: $t(37) = 3.90$, $p < 0.001$, $g = 0.60$ [95% CI: 0.27 - 0.93]). Scalp topographies were similar to the patterns observed in previous studies employing comparable stimuli and stimulation frequencies (Liu-Shuang et al., 2014; De Rosa et al., 2022).

The results of the older adults are illustrated in Fig. 3.4. A significant response at the base stimulation frequency of 6Hz was observed in the occipito-temporal area for both types of stimuli (Words: $t(51) = 10.66$, $p < 0.001$, $g = 1.48$ [95% CI: 1.08 - 1.87]; Pseudofonts: $t(51) = 11.154$, $p < 0.001$, $g = 1.55$ [95% CI: 1.139 - 1.948]). A significant response to the oddball stimulation frequency of 1.2Hz was visible in predefined ROIs for both words (Left ROI: $t(51) = 4.25$, $p < 0.0001$, $g = 0.589$ [95% CI: 0.292 - 0.882]; Right ROI: $t(51) = 4.7$, $p < 0.0001$, $g = 0.658$ [95% CI: 0.355 - 0.955]) and pseudofonts (Left ROI: $t(51) = 1.155$, $p < 0.0001$, $g = 0.715$ [95% CI: 0.407 - 1.017]; Right ROI: $t(51) = 4.29$, $p < 0.0001$, $g = 0.595$ [95% CI: 0.297 - 0.887]). Responses scalp topographies were similar to the patterns seen in previous studies that used the same types of stimuli at similar frequency rates (Liu-Shuang et al., 2014; De Rosa et al., 2022).

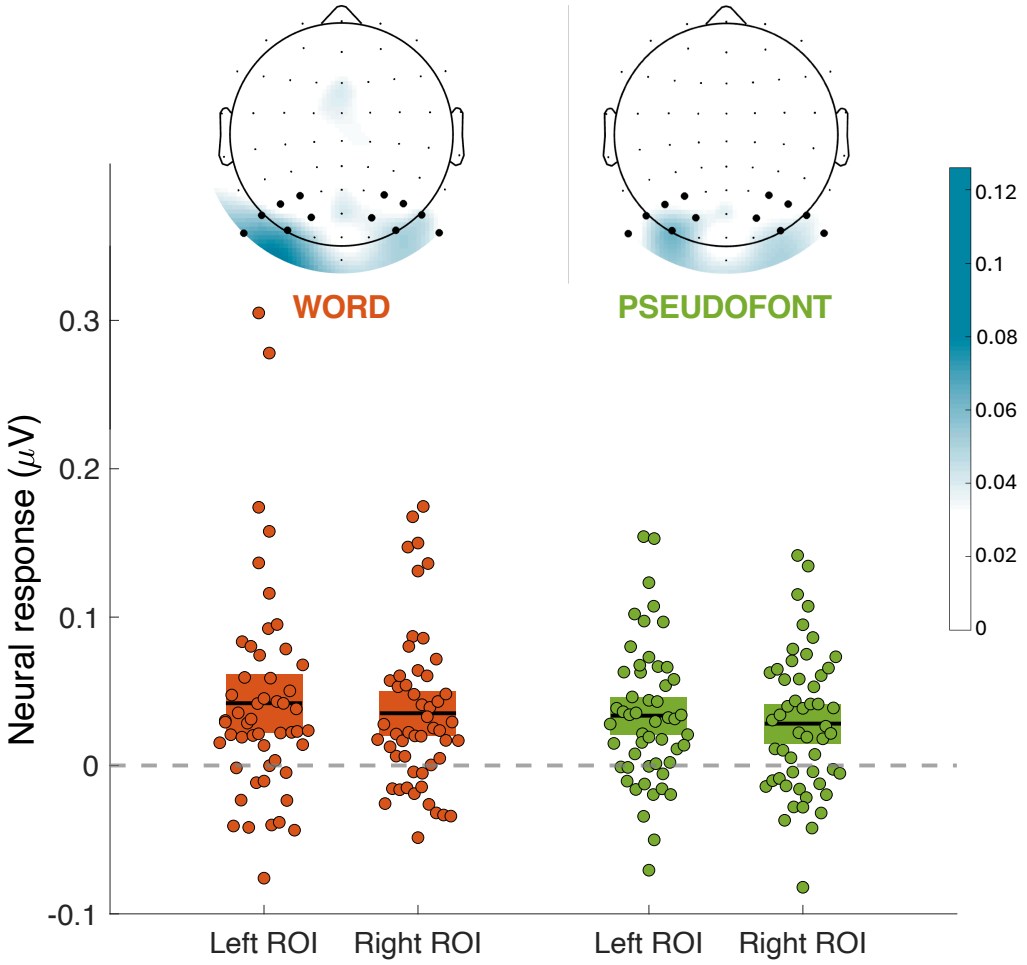


Figure 3.4: **FPVS grand-averaged token frequency response in the older adults.** Grand-averaged token frequency response estimated as the summed baseline corrected amplitudes for the oddball frequency (1.2Hz) and its three harmonics (2.4, 3.6, 4.8 Hz) for the two different conditions (Words in orange, Pseudofonts in green) in the left and right occipito-temporal ROIs, marked in the corresponding scalp topographies. Boxes illustrate the 95% t-interval around the mean. Dots stand for single participants. The dashed line represents the level of noise.

Direct comparisons between younger and older adults revealed reduced oddball response amplitudes in the elderly group for words in both ROIs (Left ROI: M_{young} = 0.12 μV, $M_{elderly}$ = 0.04 μV; t = 2.4, p = 0.02, 95% CI [0.012, 0.134]; Right ROI: M_{young} = 0.07 μV, $M_{elderly}$ = 0.04 μV; t = 2.1, p = 0.04, 95% CI [0.001, 0.068]) and for pseudofonts in the right ROI (Left ROI: M_{young} = 0.05 μV, $M_{elderly}$

= 0.03 μV; t = 1.24, p = 0.22, 95% CI [-0.012, 0.053]; Right ROI: M_{young} = 0.08 μV, $M_{elderly}$ = 0.03 μV; t = 2.7, p = 0.01, 95% CI [0.013, 0.094]; see Fig. 3.5). In contrast, no significant difference was observed between the two age groups for the base stimulation frequency of 6Hz in either words (M_{young} = 0.6 μV, $M_{elderly}$ = 0.55 μV; t = 0.69, p = 0.5, 95% CI [-0.084, 0.174]) or pseudofonts (M_{young} = 0.532 μV, $M_{elderly}$ = 0.536 μV; t = -0.06, p = 0.95, 95% CI [-0.12, 0.11]), suggesting that the reduced oddball response in the elderly is not due to a general decrease in brain activation, but rather to a selective reduction in sensitivity to statistical regularities.

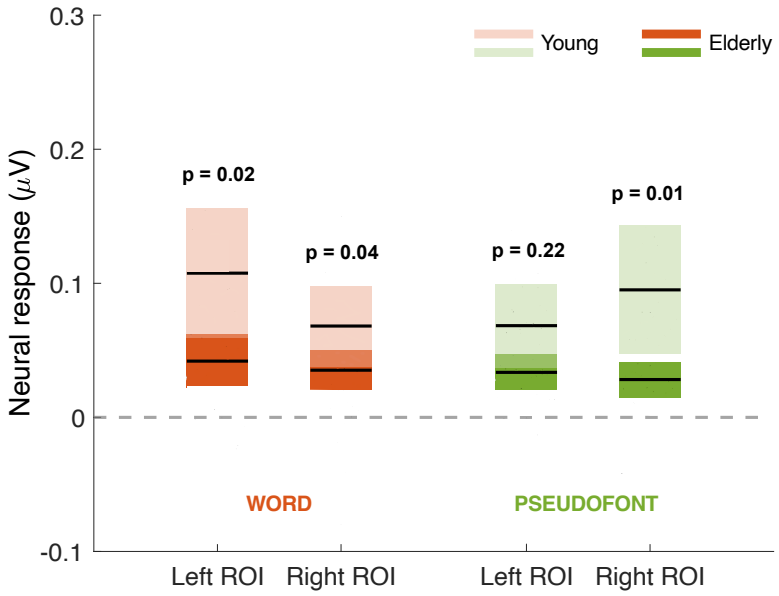


Figure 3.5: **Young vs. Elderly adults FPVS grand-averaged token frequency response** Grand-averaged token frequency response estimated as the summed baseline corrected amplitudes for the oddball frequency. Boxes depict the 95% t-interval around the mean for words (in orange) and pseudofonts (in green) in the left and right occipito-temporal ROIs, with light boxes representing brain activation of young subjects, and dark boxes the one of elderly adults. The dashed line represents the level of noise.

3.3.1.3 Discrimination Task

The Discrimination task was designed to assess participants' explicit memory for stimuli presented during the FPVS paradigm. Specifically, participants were asked to discriminate previously presented items from novel foils separately for base and oddball stimuli, allowing us to evaluate whether frequency-based exposure differences translated into explicit behavioural sensitivity.

In younger adults, performance in the discrimination task was close to ceiling for words in both the base and oddball conditions (Bases: M = 83.9%, SD = 13.0, t(37) = 18.9, p < 0.0001, d = 2.93 [95% CI: 79.8% - 88.0%]; Oddballs: M = 82.1%, SD = 14.4, t(37) = 16.7, p < 0.0001, d = 2.61 [95% CI: 77.5% - 86.7%]). Performance with pseudofonts was substantially lower and did not reliably differ from chance (Bases: M = 53.7%, SD = 14.0, t(37) = 1.75, p = 0.08, d = 0.27 [95% CI: 49.2% - 58.1%]; Oddballs: M = 53.3%, SD = 13.8, t(37) = 1.59, p = 0.12, d = 0.25 [95% CI: 48.9% - 57.7%]). A positive association was observed between performance in the word base vs. foil and oddball vs. foil conditions (r = 0.67, t(43) = 5.79, p < 0.0001), whereas no reliable association was present for pseudofonts (r = 0.079, t(43) = 0.51, p = 0.61). Following Ruzza and Crepaldi (2026), we further examined whether participants showed differential performance for base and oddball items. Because base stimuli were presented more frequently during the FPVS task, one possibility was that participants would later discriminate base items more accurately due to repeated exposure. Alternatively, oddball items might benefit from increased saliency or surprise-related processing. However, no reliable difference emerged between base and oddball performance for either words ($\beta = 0.005$, SE = 0.020, $t = 0.26$, $p = .80$) or pseudofonts ($\beta = -0.004$, SE = 0.020, $t = -0.21$, $p = .84$), indicating that frequency differences did not translate into explicit behavioural sensitivity.

A similar pattern emerged in older adults (see Fig.3.6). Participants performed well with words (Bases: M = 84.98%, SD = 13.0, t(51) = 18.64, p < 0.0001, d = 2.69 [95% CI: 81.21% – 88.76%]; Oddballs: M = 85.91%, SD = 14.31, t(51) = 17.53, p < 0.0001, d = 2.53 [95% CI: 82.04% - 90.35%]), while performance for pseudofonts remained at chance level (Bases: M = 48.44%, SD = 13.25, t(51) = -0.82, p = 0.42, d = -0.12 [95% CI: 44.59% - 56.75%]; Oddballs: M = 51.91%, SD = 11.69, t(51) = 1.13, p = 0.26, d = 0.16 [95% CI: 48.56% - 55.3%]). Relation analyses again revealed a positive correlation between performance in the word base vs. foil and oddball vs. foil conditions (r = 0.51, t = 4.07, p = 0.0001), whereas no reliable association was observed for pseudofonts (r = 0.13, t = 0.88, p = 0.38). Importantly, the same pattern observed in younger adults was replicated in the elderly group, with no reliable asymmetry between base and oddball performance for either words ($\beta = 0.006$, SE = 0.025, $t = 0.29$, $p = .75$) or pseudofonts ($\beta = -0.003$, SE =

0.025, $t = -0.26$, $p = .89$), suggesting that differences in exposure frequency did not reliably influence explicit discrimination performance across age groups.

Direct comparisons between younger and older adults revealed a significant group difference only in the pseudofont base vs. foil condition, with younger adults outperforming older adults (t(73) = 2.55, p = 0.01, 95% CI [0.38, 3.04]). No significant group differences were observed in the remaining conditions (Pseudofont – oddball vs. foil: t(73) = 0.38, p = 0.71, 95% CI [-1.13, 1.67]; Word – base vs. foil: t(73) = -0.64, p = 0.53, 95% CI [-1.97, 1.01]; Word – oddball vs. foil: t(73) = -1.08, p = 0.28, 95% CI [-2.43, 0.72]).

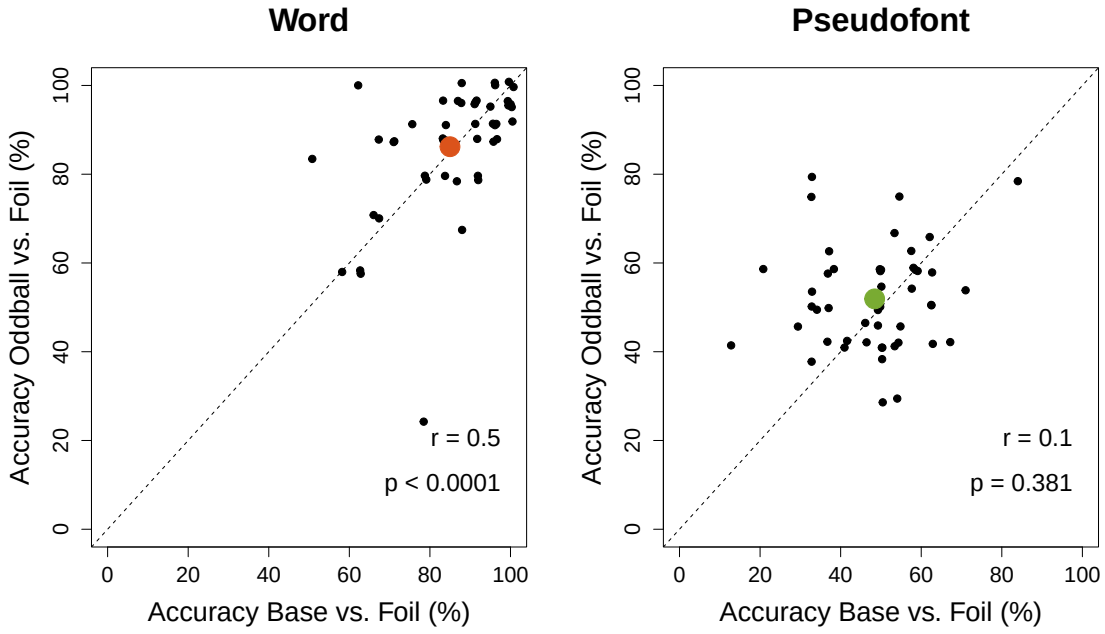


Figure 3.6: **Discrimination task performance** The left panel illustrates accuracy in discriminating bases from foils and oddballs from foils for words. The right panel displays the same comparison for pseudofonts. Each black dot represents the performance of an individual participant, with the larger colored dot representing the group mean. The diagonal dashed line indicates equal performance across both tasks.

3.3.1.4 Familiarity Task

The Familiarity task was designed to assess whether participants internalized the frequency manipulation introduced during the FPVS task. Specifically, base and oddball items were directly contrasted, and participants were asked to indicate which of the two categories they perceived as more familiar.

In younger adults, participants showed no reliable preference for either frequency category at the group level (rate of preference for oddballs; Words: $M = 47.16\%$, $SD = 12.77$, $t(43) = -0.79$, $p = .44$, $d = -0.12$ [95% CI: 44.6%–52.37%]; Pseudofonts: $M = 49.71\%$, $SD = 13.44$, $t(43) = -1.40$, $p = .17$, $d = -0.21$ [95% CI: 43.07%–51.25%]). No significant association was observed between preference for oddballs in the word and pseudofont conditions ($r = 0.24$, $t(43) = 1.60$, $p = .12$, 95% CI $[-0.06, 0.51]$).

A similar pattern emerged in older adults (see Fig.3.7, Left Panel). At the group level, participants again showed no reliable preference for frequency category at the group level (rate of preference for oddballs; Words: M = 51.68%, SD = 13.56, t(51) = 0.89, p = 0.38, d = 0.12 [95% CI: 47.91% - 55.46%]; Pseudofonts: M = 48.32%, SD = 12.76, t(51) = -0.95, p = 0.35, d = -0.13 [95% CI: 44.76% - 51.87%]). As in younger adults, no reliable association was observed between preference for oddballs in the word and pseudofont conditions ($r = 0.16$, $t(50) = 1.12$, $p = .27$, 95% CI $[-0.12, 0.41]$).

Direct comparisons between younger and older adults did not reveal significant group differences in familiarity judgments for either words (t(92.90) = -1.19, p = 0.24, 95% CI [-2.05, 0.52]) or pseudofonts (t(89.64) = -0.43, p = 0.67, 95% CI [-1.56, 1.00]). Taken together, these findings indicate that either younger nor older adults showed a reliable explicit preference for the more or less frequent stimulus category, despite having been exposed to the frequency manipulation during the FPVS task.

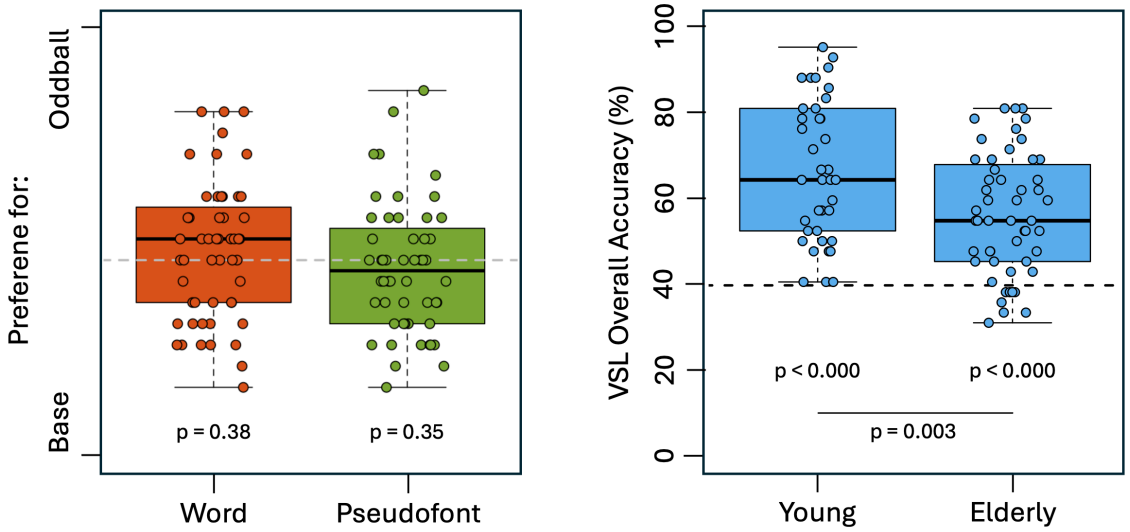


Figure 3.7: **Familiarity Task and VSL Task performances. Left panel:** Familiarity task performance in older adults. Rate of preference for oddball items in the familiarity task for words and pseudofonts. Values above the dashed line indicate preference for oddballs, whereas values below the line indicate preference for base items. . Each dot represents an individual participant, while boxplots illustrate the group distribution and 95% confidence intervals. The dashed line indicates chance level. **Right panel:** Overall VSL task performance in young and elderly adults. Overall accuracy in the VSL task, computed by combining performance across the pattern recognition and pattern completion conditions. The dashed line indicates chance level. Each dot represents an individual participant, while boxplots illustrate the group distribution and 95% confidence intervals.

3.3.1.5 VSL Task

Overall, performance in the VSL task was consistent with previous studies (see Fig.3.7, right Panel; Siegelman et al., 2017b; Solaja et al., 2026).

In younger adults, overall VSL performance was significantly above chance ($M = 66.48\%$, $SD = 16.41$, $t(37) = 10.07$, $p < 0.0001$, $d = 1.63$ [95% CI: 61.09% - 71.87%]). Performance on the pattern recognition and pattern completion components was positively correlated ($r = 0.58$, $t(36) = 4.28$, $p = 0.0001$, 95% CI [0.32, 0.76]), supporting their combination into a single composite measure of visual statistical learning.

Similarly, older adults performed significantly above chance on the overall VSL measure ($M = 56.30\%$, $SD = 14.11$, $t(51) = 8.41$, $p < 0.0001$, $d = 1.18$ [95% CI: 52.33% - 60.27%]). As in younger adults, performance on the pattern recognition and pattern completion components was positively correlated ($r = 0.53$, $t(49) = 4.34$,

$p < 0.0001$, 95% CI [0.29, 0.70]).

Direct comparisons between younger and older adults revealed significantly higher overall VSL performance in younger adults than in older adults ($t(72.68) = 3.07$, $p = 0.003$, 95% CI [3.57, 16.78]), indicating an age-related decline in visual statistical learning.

3.3.2 Reliability Analyses

Reliability analyses are particularly important in studies focusing on individual differences, because tasks that robustly produce group-level effects do not necessarily provide reliable measurements at the individual level (Hedge et al., 2018). In the context of statistical learning research, limited reliability can substantially attenuate correlations between measures, reduce statistical power, and contribute to inconsistent findings across studies (Bogaerts et al., 2026). Recent consensus recommendations have therefore emphasized the importance of systematically reporting reliability indices when investigating individual variability in SL (Bogaerts et al., 2026). Following these recommendations, we assessed the internal consistency of all neural and behavioural measures included in the present study before conducting correlation analyses.

3.3.2.1 FPVS Tasks

Reliability analyses were conducted separately in the younger and older adult samples using Cronbach's α as an index of internal consistency across stimulation blocks. In the younger sample, reliability of the FPVS neural response at the oddball stimulation frequency was low for both words ($\alpha = .21$) and pseudofonts ($\alpha = .4$), indicating that the FPVS oddball response did not provide a stable estimate of individual differences under the present experimental conditions.

A somewhat different pattern emerged in the older adult sample. Reliability remained low for the word condition ($\alpha = .32$), but reached acceptable levels for pseudofonts ($\alpha = .55$).

One possible explanation for the modest reliability of the FPVS metrics is the relatively small number of stimulation blocks included in each condition (five blocks per stimulus type), which may limit the stability of individual estimates. Longer recording sessions or a greater number of stimulation blocks may therefore improve the reliability of FPVS-based measures of individual differences in future studies. Importantly, limited reliability is expected to attenuate correlations and may therefore lead to wrong estimates of the true association between neural responses, behavioural performance, and age.

3.3.2.2 Behavioural Tasks

Reliability analyses were conducted separately in the younger and older adult samples using Cronbach's α as an index of internal consistency.

In younger adults, reliability was highest for the discrimination task with words ($\alpha = .77$), indicating that this measure captured relatively stable individual differences in performance. Reliability was substantially lower for pseudofonts ($\alpha = .46$), suggesting that performance in this condition provided a less stable estimate of individual differences. A similar pattern emerged in older adults. Reliability was again high for the discrimination task with words ($\alpha = .76$) and acceptable for pseudofonts ($\alpha = .52$), indicating greater consistency of performance in this condition than was observed in the younger sample.

The familiarity task showed mixed reliability in the younger sample. Cronbach's α was quite low for the word condition ($\alpha = .45$), but reached acceptable levels for pseudofonts ($\alpha = .51$). In older adults, reliability was acceptable for both the word ($\alpha = .60$) and pseudofont ($\alpha = .52$) conditions, suggesting that familiarity judgments provided a fairly stable estimate of individual differences in this age group.

Reliability of the VSL task was assessed separately in the younger and older adult samples, considering the overall task results, as in Siegelman et al. (2017b). Reliability of the overall VSL composite score was acceptable in both younger adults ($\alpha = .71$) and older adults ($\alpha = .73$), supporting its use as a measure of individual differences in statistical learning performance across the lifespan.

Taken together, these findings have important implications for the individual-differences analyses reported below. While limited reliability does not affect the group-level analyses presented above, it can substantially attenuate correlations between measures and thereby hinder the detection of meaningful associations. Reliability varied substantially across behavioural measures and age groups, indicating that some tasks provide more stable estimates of individual differences than others. These psychometric properties were taken into account when selecting measures for the correlation analyses reported in the following section.

3.3.3 Correlation Analyses

We next examined the relationships among neural and behavioural measures of statistical learning, as well as the extent to which performance generalized across tasks. Because correlational analyses are strongly influenced by measurement reliability, all correlations were interpreted in light of the psychometric properties reported above.

Following the criterion established in the Reliability Analyses section, primary correlational analyses were restricted to measures showing reliability estimates of at

least $r = .50$. This threshold was applied separately within each age group, resulting in different sets of measures being retained for analysis. In the elderly sample, several behavioural measures and the FPVS pseudofont response exceeded the reliability criterion and were therefore included in the primary analyses. In contrast, fewer measures met this criterion in the younger sample, limiting the number of correlations that could be meaningfully interpreted.

Measures that failed to reach the primary reliability threshold were excluded from confirmatory correlational analyses, as low reliability is expected to attenuate observed associations and complicate their interpretation. However, because several measures approached acceptable reliability, we additionally conducted exploratory analyses using a more liberal threshold of $r \geq .40$. These analyses were intended to assess whether similar patterns emerged when considering measures with somewhat lower, but still potentially informative, reliability estimates.

Accordingly, the sections below distinguish between primary correlational analyses based on the most reliable neural and behavioural indices of statistical learning and exploratory analyses involving measures that did not meet the primary reliability criterion. All exploratory findings should therefore be interpreted with appropriate caution.

3.3.3.1 Age-Related Effects

To investigate whether statistical learning abilities vary as a function of age within the elderly sample, we examined the relationship between chronological age and performance across neural and behavioural measures.

We first examined the relationship between age and the FPVS neural response in the pseudofont condition, which was the only neural measure exceeding the reliability threshold adopted for the primary analyses. No significant correlation emerged between age and the FPVS signal. In the left ROI, the correlation was positive but non-significant ($r = 0.17$, $t(49) = 1.20$, $p = .24$, 95%. CI $[-0.11, 0.43]$); in the right ROI, the correlation was negative and non-significant ($r = -0.12$, $t(49) = -0.82$, $p = .42$, 95%. CI $[-0.37, 0.17]$). No association emerged when combining responses across ROIs ($r = 0.03$, $t(49) = 0.21$, $p = .84$, 95%. CI $[-0.25, 0.30]$). These findings suggest that neural sensitivity to the statistical regularities embedded in the FPVS streams remains relatively stable across the age range sampled here.

We next examined the relationship between age and performance in the discrimination task (see Fig.3.8, Left Panel). A significant negative correlation was observed for word discrimination in the base vs. foil condition ($r = -0.37$, $t(49) = -2.76$, $p = .008$, 95%. CI $[-0.58, -0.10]$), indicating reduced performance with increasing

age. No significant correlation emerged in the oddball vs. foil condition, although the association was in the same direction ($r = -0.21$, $t(49) = -1.47$, $p = .15$, 95%. CI [−0.46, 0.07]). For discrimination in the pseudofont condition, no significant association with age was observed in the base vs. foil condition ($r = -0.12$, $t(49) = -0.86$, $p = .39$, 95%. CI [−0.38, 0.16]), while a marginal negative association emerged in the oddball vs. foil condition ($r = -0.26$, $t(49) = -1.86$, $p = .069$, 95%. CI [−0.50, 0.02]), suggesting a possible tendency for discrimination performance to decline with age.

We then examined the familiarity task. No significant relationship between age and familiarity judgments was observed in the word condition ($r = 0.11$, $t(49) = 0.77$, $p = .44$, 95%. CI [−0.17, 0.37]); in the pseudofont condition, a marginal negative association emerged ($r = -0.24$, $t(49) = -1.73$, $p = .089$, 95%. CI [−0.48, 0.04]), suggesting a tendency for preference for oddball items to decrease with increasing age.

Finally, we examined the relationship between age and VSL performance (see Fig.3.8, Right Panel). A significant negative relationship was observed between age and overall VSL accuracy ($r = -0.49$, $t(49) = -3.99$, $p = .0002$, 95% CI [−0.68, −0.25]), indicating that statistical learning performance decreased with increasing age.

Overall, these findings suggest that neural entrainment to statistical regularities remains relatively stable across the age range sampled here, whereas behavioural measures reveal age-related declines. The strongest age-related effect was observed for the VSL task, which showed a robust negative association between age and learning performance. Similar, although weaker, tendencies were observed for discrimination and familiarity measures, particularly in the pseudofont conditions.

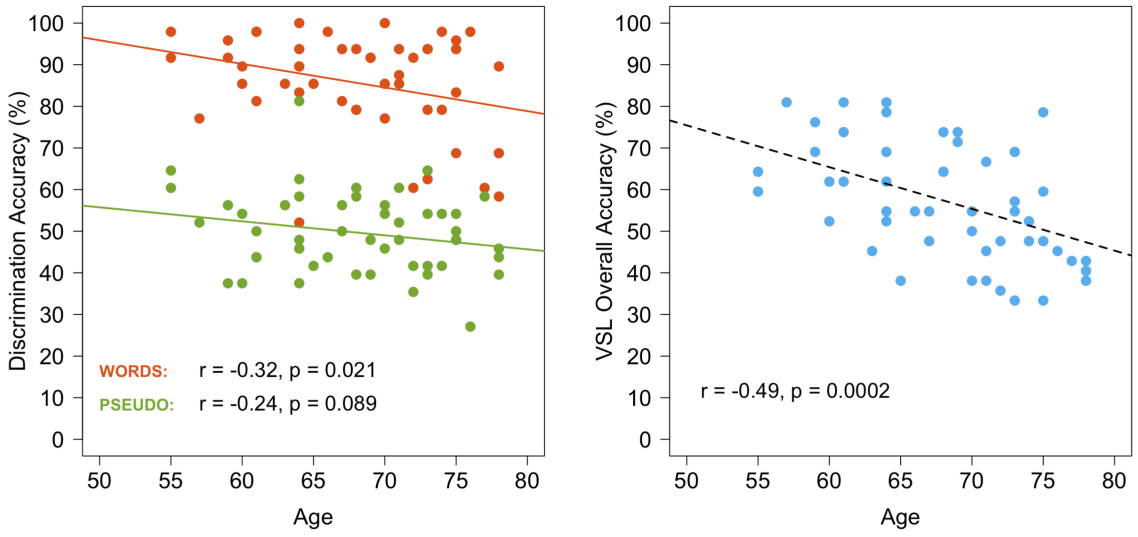


Figure 3.8: **Age-related effects in behavioural measures of statistical learning. Left panel:** Relationship between age and overall VSL task accuracy within the elderly sample. Accuracy reflects combined performance across the pattern recognition and pattern completion conditions. Each dot represents one participant, and the dashed regression line illustrates the significant negative association between age and VSL performance. **Right panel:** Relationship between age and overall discrimination task performance for words (orange) and pseudofonts (green). Accuracy reflects combined performance across base vs. foil and oddball vs. foil conditions. Each dot represents one participant, and solid regression lines illustrate the association between age and performance for each stimulus type.

3.3.3.2 Neural–Behavioural Relations

Given that the FPVS pseudofont response in the elderly subjects group was the only neural measure exceeding the reliability threshold of $\alpha > .50$, primary neural–behavioural analyses focused on this measure in the elderly sample. We examined its relationship with behavioural indices of statistical learning, including discrimination performance, familiarity judgments, and VSL performance.

No significant association emerged between FPVS pseudofont responses and pseudofont discrimination performance. Correlations were non-significant in both the left ROI ($r = .18$, $p = .20$ for base trials; $r = -.12$, $p = .38$ for oddball trials) and the right ROI ($r = .22$, $p = .12$ for base trials; $r = .00$, $p = .99$ for oddball trials). Likewise, when neural responses were combined across ROIs and compared with overall pseudofont discrimination performance, no significant relationship was observed ($r = .13$, $t(49) = 0.89$, $p = .38$, 95%. CI $[-0.15, 0.39]$).

Similarly, no significant relationship was observed between FPVS pseudofont responses and familiarity judgments. Correlations were non-significant in both the left ROI ($r = -.16$, $p = .26$) and the right ROI ($r = .04$, $p = .77$). The combined ROI measure was likewise unrelated to familiarity performance ($r = -.07$, $t(49) = -0.50$, $p = .62$, 95%. CI [−0.34, 0.21]).

Finally, no association emerged between FPVS pseudofont responses and VSL performance. Correlations were non-significant in the left ROI ($r = .08$, $p = .57$), the right ROI ($r = .14$, $p = .34$), and for the combined ROI measure ($r = .13$, $t(49) = 0.93$, $p = .36$, 95%. CI [−0.15, 0.39]).

We additionally conducted exploratory analyses using a more liberal reliability criterion ($\alpha \geq .40$), which allowed the inclusion of the FPVS pseudofont measure in the younger sample. As in the elderly group, no significant associations were observed between FPVS pseudofont responses and behavioural measures of statistical learning. Correlations with pseudofont discrimination performance were non-significant (combined ROI: $r = .13$, $t(36) = 0.80$, $p = .43$, 95%. CI [−0.20, 0.43]), as were correlations with familiarity judgments (combined ROI: $r = -.08$, $t(50) = -0.57$, $p = .57$, 95%. CI [−0.35, 0.20]) and VSL performance (combined ROI: $r = .14$, $t(36) = 0.82$, $p = .42$, 95%. CI [−0.19, 0.44]). Correlations in the individual ROIs were likewise non-significant in all cases (all ps $> .13$).

Overall, despite evidence of statistical learning at both the neural and behavioural levels, neural FPVS responses did not reliably predict behavioural measures of statistical learning in either age group.

3.3.3.3 Behavioural–Task Relations

We next examined the relationships among the behavioural measures of statistical learning to assess whether tasks hypothesized to tap into related learning mechanisms were associated at the individual level.

We first examined the relationship between discrimination performance and familiarity judgments. In the elderly sample, no reliable associations emerged between the two measures. Specifically, discrimination accuracy for words was not associated with familiarity-based preferences for words ($r = .08$, $t(49) = 0.59$, $p = .561$, 95%. CI [−0.20, 0.35]). Similarly, no significant relationship was observed between discrimination and familiarity performance for pseudofonts ($r = .19$, $t(49) = 1.33$, $p = .189$, 95%. CI [−0.09, 0.44]). These findings suggest that discrimination-based and familiarity-based measures captured largely independent sources of individual variability in the elderly group.

We next examined relations with performance in the classic VSL task. In younger

adults, no reliable associations were observed. Word discrimination performance showed a positive but non-significant relationship with VSL accuracy ($r = .28$, $t(36) = 1.73$, $p = .092$, 95%. CI [−0.05, 0.55]). Likewise, familiarity performance for pseudofonts was not significantly related to VSL performance ($r = .15$, $t(36) = 0.91$, $p = .371$, 95%. CI [−0.18, 0.45]).

In contrast, several significant associations emerged in the elderly sample. VSL performance was positively associated with discrimination accuracy for words ($r = .30$, $t(49) = 2.19$, $p = .034$, 95%. CI [0.02, 0.53]) and pseudofonts ($r = .29$, $t(49) = 2.15$, $p = .036$, 95%. CI [0.02, 0.53]). Thus, individuals showing stronger statistical learning in the VSL task also tended to perform better on discrimination-based measures of statistical learning. No relationship was observed between VSL performance and familiarity judgments for words ($r = -.03$, $t(49) = -0.24$, $p = .814$, 95%. CI [−0.31, 0.24]). However, a significant positive association emerged between VSL performance and familiarity judgments for pseudofonts ($r = .28$, $t(49) = 2.01$, $p = .050$, 95%. CI [0.001, 0.51]), indicating that individuals with higher VSL performance showed a greater tendency to prefer pseudofont oddballs.

Finally, we conducted exploratory analyses using a more liberal reliability threshold ($r \geq .40$) for the younger sample. Under this criterion, we examined relations involving measures that did not meet the primary reliability threshold but approached acceptable reliability. No significant association was observed between word discrimination and word familiarity performance ($r = -.01$, $t(36) = -0.05$, $p = .957$, 95%. CI [−0.33, 0.31]). Similarly, pseudofont discrimination was not associated with pseudofont familiarity ($r = .09$, $t(36) = 0.55$, $p = .587$, 95%. CI [−0.24, 0.40]). No reliable relationship emerged between pseudofont discrimination and VSL performance ($r = .20$, $t(36) = 1.21$, $p = .235$, 95%. CI [−0.13, 0.49]), nor between word familiarity and VSL performance ($r = -.04$, $t(36) = -0.26$, $p = .793$, 95%. CI [−0.36, 0.28]).

Taken together, these findings indicate that behavioural measures of statistical learning showed limited convergence overall. The most consistent associations emerged in the elderly sample, where better performance on the VSL task was associated with higher discrimination accuracy for both words and pseudofonts, as well as stronger familiarity-based preferences for pseudofont oddballs. By contrast, younger adults showed no reliable associations between tasks, even under the more liberal exploratory reliability criterion.

3.4 General Discussion

The present study investigated Statistical Learning (SL) in healthy aging by combining neural and behavioural measures within a unified experimental framework. Specifically, we used a Fast Periodic Visual Stimulation (FPVS) oddball paradigm to capture implicit neural sensitivity to frequency-based regularities, alongside multiple behavioural tasks—discrimination, familiarity and a classic Visual Statistical Learning (VSL) task. This multimethod approach allowed us to explore whether neural sensitivity to statistical regularities is preserved in aging, whether behavioural measures reveal comparable patterns of preservation or decline, and whether neural and behavioural indices of SL are related at the individual level. By comparing younger and older adults and examining individual variability, the study aimed to address the mixed evidence reported in the literature and to provide preliminary insights toward a more integrated account of SL across the lifespan.

Preserved but Reduced Neural Sensitivity in Aging

We firstly aimed at investigating whether neural sensitivity to statistical regularities, as indexed by FPVS oddball responses, is preserved in healthy aging. Our results provide evidence that this sensitivity is largely preserved. Older adults showed significant responses at the oddball frequency for both words and pseudofonts, indicating that the brain remains capable of tracking frequency-based regularities even in later life. Importantly, the effect was observed for both familiar and unfamiliar visual stimuli, suggesting that neural sensitivity to these regularities does not depend on prior lexical knowledge. However, oddball responses were significantly reduced in amplitude relative to younger adults, while base-rate responses at 6 Hz were comparable across groups. This dissociation indicates that aging spares general visual entrainment, but selectively affects neural discrimination of statistical structure.

This pattern aligns with and extends previous findings. Prior FPVS and neural entrainment studies show that older adults retain category-selective responses when stimulation parameters are adapted to age-related changes (Marchive et al., 2025; Sáringer et al., 2024; Peykarjou et al., 2024), alongside qualitative changes such as reduced amplitude, decreased selectivity, and increased neural noise (Mao et al., 2025; Huizeling et al., 2021). Our results replicate this coexistence of preservation and alteration, and further demonstrate that it applies to statistical regularities in continuous input.

Importantly, preserved base-rate responses rule out a general decline in neural responsiveness, pointing instead to a specific reduction in the efficiency of the

differentiation between frequent and infrequent stimuli. This may reflect reduced representational precision or lower signal-to-noise ratio, consistent with accounts of increased neural noise in aging (Li et al., 2001; Bäckman et al., 2006; Huizeling et al., 2021). Neural sensitivity was observed for both words and pseudofonts, indicating that FPVS indexes SL mechanisms operating independently of prior knowledge. The parallel reduction in amplitude across stimulus types further supports a more general effect of aging on neural sensitivity to statistical regularities.

Finally, within the older group, no reliable relationship emerged between chronological age and the amplitude of the neural response, suggesting that sensitivity to statistical regularities may not decline linearly across older adulthood. Instead, this points to heterogeneity in aging, potentially driven by factors such as cognitive reserve and neural efficiency (Bogaerts et al., 2026). At the same time, the relatively modest reliability of the FPVS measure across blocks suggests that individual-differences findings should be interpreted cautiously, as reduced measurement precision likely hinders our ability to observe age-related effects.

Heterogeneous Behavioural Expressions of Statistical Learning

Our second aim was to investigate whether behavioural measures related to statistical learning show similar patterns of preservation or decline in aging. The results revealed a heterogeneous, task-dependent pattern.

The clearest picture is probably in the VSL task, where both groups performed above chance, but older adults showed lower overall accuracy and a robust negative relationship with age. In the discrimination task, participants performed above chance for words but not for pseudofonts, suggesting better explicit memory performance for familiar verbal material than for unfamiliar visual stimuli. Performance also declined with age, particularly for words. In contrast, the familiarity task did not reveal reliable preferences for either frequent or infrequent items at the group level, despite substantial inter-individual variability, and no robust relationship with age. Taken together, these findings indicate that behavioural expressions of sensitivity to regularities vary substantially across tasks and stimulus types. While participants' brains showed clear sensitivity to frequency regularities in the FPVS task, this sensitivity was not consistently expressed behaviourally across paradigms and appeared more vulnerable to aging, at least under some task conditions.

This pattern aligns with previous literature showing relatively preserved SL in older adults in some contexts (Bulgarelli et al., 2021; Johannsen et al., 2022; Schevenels et al., 2023), alongside age-related reductions in more demanding tasks (Ge et al., 2025; Palmer et al., 2018; Lega et al., 2023). Our results fit this broader framework:

performance remained above chance in several conditions, yet age-related reductions emerged particularly in the VSL task. The strong negative relationship between age and VSL performance is especially informative. Unlike FPVS, which captures rapid and implicit neural sensitivity to statistical regularities, the VSL task requires sustained exposure, integration of sequential information, and explicit behavioural decisions, placing greater demands on domain-general processes such as attention and working memory (Conway et al., 2010; Misyak and Christiansen, 2012). The association with age therefore suggests that these broader cognitive constraints may critically shape behavioural outcomes, even when implicit neural sensitivity to regularities remains relatively preserved.

Importantly, the behavioural tasks differed substantially in their psychometric properties. The discrimination task with words and the overall VSL measure showed the highest reliability, whereas familiarity judgments and pseudofont-based measures were considerably more variable. Accordingly, the weaker and less consistent behavioural effects observed in some conditions should be interpreted cautiously, as reduced measurement precision may attenuate observable effects and increase inter-individual variability.

Overall, the behavioural findings support a nuanced account: sensitivity to statistical regularities in aging is neither fully preserved nor uniformly impaired, but depends on task demands, stimulus properties, and the cognitive processes required to behaviourally express learned information.

Neural and Behavioural Markers of Statistical Learning: Limited Convergence Across Levels of Processing

A third question we aimed at investigating was whether neural and behavioural indices of SL converge at the individual level. Given the low reliability of the FPVS measure, particularly in the word condition, neural–behavioural correlations were interpreted cautiously and were not considered sufficiently robust to support strong individual-differences conclusions. Nevertheless, an informative dissociation emerged at the group level. While the FPVS task revealed clear neural sensitivity to statistical regularities in both young and older adults, behavioural tasks showed weaker and more variable evidence of learning. In contrast, behavioural measures showed selective coherence with one another: VSL performance was positively associated with discrimination accuracy for both words and pseudofonts, and more weakly with familiarity judgments for pseudofonts. Overall, these findings suggest limited convergence between neural and behavioural indices of SL, with FPVS capturing robust implicit sensitivity to regularities, and behavioural measures reflecting more variable

and task-dependent expressions of learning.

This pattern aligns with previous findings showing that neural entrainment paradigms can reveal learning-related sensitivity even when behavioural evidence is weak or absent (Frost et al., 2015; Stothart and Smith, 2020; Rawal and Tseng, 2021). We take this evidence to suggest that neural measures index early and implicit stages of SL, which are not always accessible to conscious report or explicit decision-making. Behavioural performance is also likely to reflect additional cognitive demands beyond learning itself. Tasks such as discrimination, familiarity judgments, and VSL require retrieval, comparison, maintenance of information, and explicit decisions, which rely on domain-general resources such as attention and working memory (Sklenar and Mienaltowski, 2019; Donohue et al., 2019). These processes are known to vary substantially across individuals and to decline with aging, potentially contributing to the observed variability in behavioural performance.

Importantly, the lack of strong convergence between neural and behavioural measures remained evident even when considering the VSL task, which showed comparatively higher reliability than the other behavioural measures. This suggests that the dissociation cannot be explained solely by measurement noise, but may instead reflect differences in the processes captured by neural and behavioural paradigms. More generally, these findings are consistent with multidimensional accounts of SL (Schmalz et al., 2019; Siegelman, 2020a), according to which different paradigms engage at least partially distinct learning mechanisms and cognitive operations. In aging, this dissociation may become particularly pronounced if implicit sensitivity to regularities remains relatively preserved, while explicit behavioural expression increasingly depends on age-sensitive cognitive resources. Such a framework may help reconcile the heterogeneous findings reported in the aging and SL literature.

Overall, these findings highlight the importance of a multimethod approach: neural and behavioural measures should be considered complementary rather than interchangeable, as they likely capture different levels and expressions of statistical learning.

Individual Differences in Statistical Learning: Variability and Dissociation Across Measures

A final question concerned the extent to which individual variability shapes SL in aging, and whether such variability is consistently reflected across neural and behavioural measures. We observed substantial inter-individual variability across all measures. At the neural level, oddball responses varied widely, ranging from robust to near-noise levels. Behaviourally, participants ranged from chance to high per-

formance in the discrimination task, showed variable preferences in the familiarity task, and differed markedly in VSL performance. Importantly, behavioural measures showed only partial coherence, with VSL performance correlating with discrimination accuracy, but not so consistently with familiarity judgments.

This pattern supports the view that individual variability in SL is a fundamental property rather than a mere reflection of measurement noise (Bogaerts et al., 2026). Prior work has shown large individual differences in learning ability (Kaufman et al., 2010; Siegelman and Frost, 2015), and our findings extend this observation to aging, where variability may be amplified by interactions between preserved learning mechanisms and age-related changes in broader cognitive resources. Importantly, considerable heterogeneity was observed even within the elderly group, despite relatively restricted age ranges and a generally preserved neural sensitivity to regularities.

This interpretation aligns with evidence that neural measures can reveal learning in the absence of behavioural effects (Stothart and Smith, 2020; Rawal and Tseng, 2021), and that similar behaviour can arise from different neural mechanisms (Orpella et al., 2022; Verty et al., 2024). At the same time, the association between VSL and discrimination performance indicates some coherence within behavioural measures, consistent with evidence that SL abilities can generalize across tasks with partially overlapping demands (Arciuli and Simpson, 2012; Kidd and Arciuli, 2016). However, the overall pattern of partial convergence supports multidimensional views of SL (Schmalz et al., 2019; Siegelman, 2020a), according to which different paradigms capture distinct but related aspects of learning.

In aging, these findings imply that chronological age alone is not a sufficient predictor of SL ability. Although some behavioural measures declined with age, substantial individual variability remained observable across tasks. This suggests that factors such as cognitive reserve, attentional resources, neural efficiency, and compensatory mechanisms may play an important role in shaping learning outcomes in older adulthood (Strzelczyk et al., 2025; Zając-Lamparska et al., 2025).

Overall, these results highlight the importance of moving beyond group averages to consider individual variability. A multimethod approach combining neural and behavioural measures remains essential for capturing the complexity of SL and disentangling preserved mechanisms from age-related changes in performance.

Psychometric Considerations in Statistical Learning Research

An important aspect emerging from the present study concerns the psychometric properties of SL measures. Reliability analyses revealed substantial variability across tasks and conditions. Behavioural measures such as discrimination with word stimuli

and the overall VSL task showed acceptable reliability and appeared to provide relatively stable estimates of individual differences. In contrast, familiarity judgments, pseudofont-based measures, and especially FPVS oddball responses showed lower and more variable reliability. This pattern is broadly consistent with growing concerns regarding the psychometric robustness of SL paradigms and their suitability for individual-differences research (Siegelman et al., 2017b; West and Vadillo, 2018; Bogaerts et al., 2026).

Importantly, reduced reliability has direct implications for interpreting correlations between measures. Low internal consistency attenuates observable associations and may obscure relationships that are genuinely present at the latent level. In the present study, this issue was particularly relevant for the FPVS task, where neural responses showed clear and robust group-level effects despite limited reliability across blocks. One likely explanation is that the FPVS design was originally optimized to maximize sensitivity to group-level entrainment effects rather than stable individual-differences estimation. In particular, the relatively small number of stimulation blocks per condition may have limited reliability, as reliability in neural measures generally increases with longer recording durations and larger trial counts.

At the same time, the present findings indicate that not all SL measures are equally affected by these psychometric constraints. The comparatively stronger reliability of the VSL task and of discrimination with word stimuli suggests that some paradigms may provide more stable behavioural estimates than others, particularly when a larger number of trials is available or when tasks rely on more explicit behavioural decisions. These differences likely contribute to the heterogeneous pattern of convergence observed across measures.

More broadly, these findings highlight the importance of considering psychometric properties as a central methodological issue in SL research, particularly in aging populations where increased variability and reduced signal-to-noise ratio may further complicate measurement. Future work would benefit from developing paradigms specifically optimized for individual-differences analyses, increasing trial counts and recording durations, and combining neural and behavioural measures within longitudinal and multimethod frameworks. Such approaches may help disentangle genuine dissociations between measures from limitations imposed by measurement reliability itself.

3.5 Conclusion

The present study provides a multimethod account of SL in healthy aging by integrating neural and behavioural measures within a unified framework. The findings demonstrate that neural sensitivity to statistical regularities, as indexed by FPVS oddball responses, is largely preserved in older adulthood, although reduced in amplitude relative to younger adults. In contrast, behavioural performance showed a more heterogeneous pattern, with some measures revealing preserved performance, others showing age-related decline, and familiarity judgments failing to provide robust evidence of learning at the group level. Together, these findings suggest that aging affects the behavioural expression of statistical learning more strongly than the underlying neural sensitivity to regularities.

Importantly, the limited convergence observed across neural and behavioural measures further suggests that different paradigms capture partially distinct aspects of SL. At the same time, the present findings highlight the importance of considering psychometric properties when interpreting individual differences in SL, as variability in reliability substantially shapes the strength and interpretability of observed associations across tasks.

More broadly, these results support a view of SL as a multifaceted and heterogeneous process shaped by implicit sensitivity to regularities, task demands, domain-general cognitive resources, and individual variability. Understanding how these components interact across the lifespan will be essential for developing more precise and psychometrically robust models of learning and cognition in aging.

Chapter 4

Cross-Modal Transfer of Statistical Learning: Evidence from auditory and visual sequence learning

Claudia Ruzza[1]
Erin S. Isbilen[2]
Richard N. Aslin[2]

[1]Department of Cognitive Neuroscience
Scuola Internazionale Superiore di Studi Avanzati (SISSA)
Trieste, Italy

[2]Yale Child Study Center
Yale University
New Haven, CT, USA

This chapter is adapted from a manuscript currently under revision for publication in *Journal of Experimental Psychology: General*. The pre-print can be found on PsyArXiv at https://osf.io/preprints/psyarxiv/fevds_v1.

This research was in part supported by a National Institutes of Health K99 Pathway to Independence Award (1K99HD113838-01A1) awarded to Erin S. Isbilen.

Abstract

Statistical learning (SL) enables the extraction of regularities across sensory modalities, yet it remains unclear whether learned representations are modality-specific or modality-general. Here, we investigated cross-modal generalization by exposing participants to spoken and written regularities using a sequential unimodal design with linguistic stimuli, manipulating the transfer direction (auditory-to-visual vs. visual-to-auditory), exposure order (visual-first vs. auditory-first), and task demands. Adults were tested using both a two-alternative forced-choice (2AFC) task and a Statistically Induced Chunking Recall (SICR) task. Results showed robust cross-modal transfer, with no differences between modality-match and modality-mismatch trials and no directional asymmetries, indicating symmetric generalization across modalities. Exposure order affected 2AFC performance but not SICR, suggesting effects on explicit decision processes rather than implicit learning. Critically, recall revealed greater modality sensitivity than recognition, indicating that although learning transfers across modalities, modality-specific constraints emerge under more demanding retrieval conditions. These findings support a hybrid account in which SL operates over shared representations, while modality-specific constraints influence retrieval and behavioral expression under varying task demands.

I would like to thank Melissa Vergara and Mitchell Kim for their assistance with the SICR coding.

4.1 Introduction

Statistical learning (SL) refers to the ability to extract structural regularities, either temporal or spatial, from exposure to a corpus of stimulus input even in the absence of a task or instructions on what aspects of the stimuli should be the focus of attention (see Aslin and Saffran, 2025 for a review) This capacity was first observed by Saffran et al. (1996a,b) in the context of continuous sequences of speech syllables, showing that adults and infants can segment speech based on syllable statistics, such as the transitional probability that each syllable is followed by a restricted set of subsequent syllables.

Subsequent research has shown that SL operates across multiple sensory modalities, including vision (Fiser and Aslin, 2005; Turk-Browne et al., 2008), audition (Batterink et al., 2015), and touch (Conway and Christiansen, 2005; Lengyel et al., 2019) Further evidence suggests that learning in these modalities is not fully independent but can interact across sensory systems (Conway and Christiansen, 2006; Mitchel and Weiss, 2011a; Glicksohn and Cohen, 2013; Murray and Shams, 2023).

Modality-Specific vs. Modality-General Representations in SL

A central question in SL research is whether extracted regularities are encoded in modality-specific representations or at a more abstract, modality-general level that enables cross-modal transfer.

Although SL has been observed across a range of sensory modalities and structural regularities, including segmentation of auditory, visual and tactile input, (Saffran et al., 1996a; Conway and Christiansen, 2005) as well as linguistic and non-linguistic stimuli (Nissen and Bullemer, 1987; Howard et al., 2004; Teinonen et al., 2009; Bulf et al., 2011), the widespread evidence of SL does not bear directly on the question of domain-generality and cross-modal transfer. Furthermore, direct evidence of cross-modal transfer remains mixed. Some studies suggest that SL can generalize across stimulus formats, consistent with modality-general learning mechanisms (Saffran et al., 1996a; Kirkham et al., 2002; Thiessen et al., 2013). In contrast, other studies report limited transfer across modalities, suggesting that learning may be constrained by modality-specific processing systems (Conway and Christiansen, 2005; Johansson, 2009; Misyak and Christiansen, 2012).

To reconcile these findings, (Frost et al., 2015) proposed that SL reflects domain-general computational principles implemented within modality-specific processing systems. That is, each sensory modality has its own biases and constraints (Siegelman and Frost, 2015), such as which stimulus features are encoded and how those

features interact with working memory and other processing capacities within that sensory modality. However, at an abstract level beyond the sensory modality of the input there can be structural regularities (e.g., whether speech syllables or letters in the alphabet are organized into temporally ordered triplets) regardless of the modality (e.g., visual or auditory) or domain (e.g., linguistic or non-linguistic). Thus, the same structural regularities may not result in consistent SL performance across modalities or domains either because of domain-specific biases and constraints or because of the absence of a domain-general SL mechanism. Consistent with this account, recent work by (Lukics and Lukács, 2022) shows that learning depends on input characteristics, such as modality and domain, as well as the amount of exposure during the learning phase and the way the outcome of learning is assessed.

To illustrate how cross-modal transfer of SL might occur, consider two general models depicted in Fig.4.1 for the simple case of speech syllables as in (Saffran et al., 1996a). Consistent with (Frost et al., 2015), Model 1 proposes that SL operates at the uni-modal level consistent with the sensory properties of the input, with cross-modal transfer operating at a higher level of abstraction – the domain-general representation. This model is particularly relevant for stimuli that have no a priori cross-modal correspondence. That is, this model assumes that the presentation of speech syllables does not automatically elicit an orthographic representation of those syllables. This is what we would expect for infants and young children who have not yet learned the sound-spelling correspondences of the written version of their native language. According to Model 1, then, cross-modal transfer could be 100% if there is a domain-general representation of the structural regularities, or less than 100% if uni-modal biases and constraints during the learning phase differ across the two modalities.

In contrast, Model 2 proposes that for some types of stimuli the sensory input, even if unimodal, automatically triggers a domain-general representation because of prior exposure to multi-modal input. For example, in contrast to infants who have no knowledge of orthography, literate children have learned the correspondence between spoken words and their text-based representation. As a result, when hearing the syllable /ba/ there is a domain-general representation of that syllable in both the auditory and visual modalities (and vice versa when seeing the syllable "ba"). Thus, Model 2 proposes that cross-modal transfer occurs automatically even before the structural regularity in the corpus of input has been learned. In turn, this model predicts 100% transfer unless there are uni-modal biases in the testing phase that assesses SL after the learning phase.

In the present set of experiments, we lay the groundwork for assessing cross-modal transfer of SL in pre- and early-reading children by testing adults using stimulus

materials similar to those depicted in Fig.4.1. Because adults are exposed to and tested with speech and text that is highly familiar to them, we expect robust cross-modal transfer. However, we examine whether there are biases in the direction of transfer (auditory-to-visual vs. visual-to-auditory), as well as two different post-learning assessments of SL that place different demands on the uni-modal encoding and retrieval of the test items.

SL, Reading and challenges in investigating Cross-Modal Transfer

The prior literature on cross-modal transfer of SL relevant to reading is unclear for a variety of reasons. First, while it has been proposed that SL plays a crucial role in reading acquisition, SL studies rarely use orthographic stimuli and often relate SL performance to very broad measures of reading proficiency (e.g., general comprehension), rather than measures that test sensitivity to orthographic structure, rather than measures that rely on sensitivity to orthographic structure (see Bogaerts et al., 2022a; Isbilen et al., 2022, 2025 for further discussion). Learning to read requires sensitivity to the statistical regularities of written language, including probabilistic correspondences between orthographic and phonological units and the distributional patterns of letters within words (Arciuli and Simpson, 2012; Frost et al., 2013a; Sawi and Rueckl, 2019). Consistent with this view, individual differences in both visual and auditory SL have been associated with variation in word reading and reading fluency (Arciuli and Simpson, 2012; von Koss Torkildsen et al., 2019b; Qi et al., 2019). Importantly, recent evidence further suggests that these learning mechanisms may operate across languages: sensitivity to statistical regularities in one language can predict reading performance in another, indicating bidirectional cross-linguistic transfer between linguistic SL and literacy outcomes (Ren and Wang, 2024).

However, the relationship between spoken and written SL has been mostly indirect. Many studies rely on non-linguistic stimuli and examine transfer either within a single modality across low-level features (e.g., shapes, colors, or fractals in vision; Turk-Browne et al., 2008; Lengyel et al., 2019; Uno et al., 2020 or across modalities using simple auditory and visual materials (e.g., tones or environmental sounds; Conway and Christiansen, 2006; Mitchel and Weiss, 2011a; Walk and Conway, 2011; Duato et al., 2023; Tang et al., 2026. Studies employing linguistic input are comparatively scarce and typically adopt artificial grammar learning paradigms, which require considerably different task demands than reading. In these studies, uni-modal auditory and visual conditions are typically used as baselines, and cross-modal transfer is inferred from performance differences observed in multi-modal contexts(Conway and Christiansen, 2006; Walk and Conway, 2011; Mitchel and Weiss, 2011a; Li et al.,

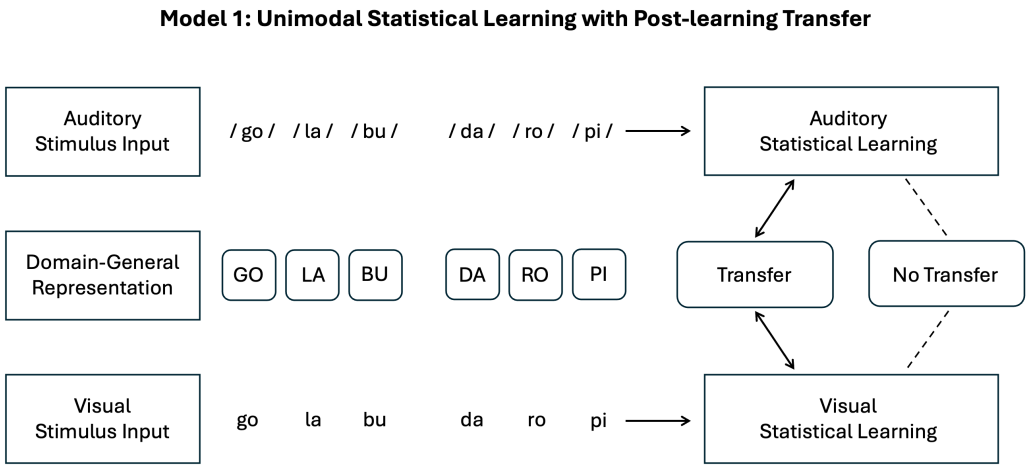


Figure 4.1: **Two models of cross-modal transfer in SL for auditory and visual stimulus materials in the domain of speech syllables and their corresponding text.** In Model 1 transfer occurs via a domain-general representation after the uni-modal structural regularities have been learned. In Model 2 the speech-text correspondence has already been learned prior to the structural regularities, thereby enabling automatic transfer of those regularities once they have been learned in the uni-modal domain.

2018). Such designs make it difficult to isolate transfer mechanisms, as learners receive concurrent information from multiple modalities and learning may proceed in parallel across channels. As a result, direct evidence for cross-modal transfer following sequential uni-modal learning remains scarce, and critical factors such as transfer directionality and the of exposure order (e.g., which modality was learned first) have rarely been examined.

Modality directionality and exposure order effects in cross-modal SL

Previous research suggests that cross-modal transfer of sequence learning is possible but limited. While some studies demonstrate transfer across modalities in adults (Altmann et al., 1995; Hupp et al., 2009), other work indicates that sequence learning is often more efficient in the auditory than in the visual modality (Conway and Christiansen, 2005; Robinson and Sloutsky, 2007). Developmental studies further show that cross-modal transfer emerges later than within-modality transfer (Hupp and Sloutsky, 2011) and that such modality asymmetries may reflect the phonological foundations of literacy. In children, sensitivity to auditory statistical regularities predicts non-word reading accuracy through phonological processing, whereas visual SL shows weaker associations (Qi et al., 2019). Similarly, auditory SL has been linked to a broader range of literacy- related skills (Spencer et al., 2015) and to reading performance in both typical and struggling readers (Gabay and Holt, 2015), in line with a long legacy of studies demonstrating a strong association between phonological awareness and reading (e.g., Frost et al., 2009b; Clayton et al., 2020b; Powell and Atkinson, 2021).

More broadly, SL abilities have been associated with language and reading development across modalities (Evans et al., 2009; Arciuli and Simpson, 2012; Frost et al., 2013a), and meta-analytic evidence confirms a reliable relationship between SL and reading outcomes (Ren and Wang, 2024). These findings align with the central role of phonological decoding in early reading (Wagner and Torgesen, 1987; Hulme et al., 2012; Melby-Lervåg et al., 2012). From this perspective, early cross-modal asymmetries may reflect a dominance of auditory–phonological representations, with learning in the auditory domain exerting stronger influence on visual processing than the reverse.

Reading proficiency itself has been proposed to depend on the acquisition of statistical regularities linking orthography, phonology, and meaning (Frost et al., 2019). In adults, however, the evidence is less consistent and provides no clear consensus regarding directional asymmetries. Some studies report bidirectional transfer, with learning acquired in one modality generalizing equally well to another (Yildirim and

Jacobs, 2015; Tang et al., 2026). At the same time, other findings suggest potential imbalances in cross-modal influence, with auditory information exerting a stronger effect on visual processing than the reverse (Robinson and Sloutsky, 2007), consistent with evidence that the auditory system may be more efficient at extracting temporal structure (Conway and Christiansen, 2005).

Developmental research further indicates that modality dominance changes over time, reflecting increasing integration between sensory and linguistic representations rather than a stable hierarchy (Robinson and Sloutsky, 2004; Nava et al., 2020; Russo et al., 2024). If literacy experience strengthens bidirectional links between orthographic and phonological representations, directional asymmetries observed earlier in development may attenuate with reading expertise, leading to comparable transfer across modality directions in skilled adult readers.

Another underrated factor in SL studies is the impact of exposure order (visual-first vs. auditory-first) on the ability to learn subsequent structural regularities. Previous studies have shown that exposure order may influence SL outcomes, as prior experience with one regularity can shape the processing of subsequent input: learning a first structure can create expectations or memory traces that either facilitate or interfere with the acquisition of a second, different regularity (Gebhart et al., 2009; Gómez and Gerken, 2000; Conway and Christiansen, 2005; Benitez et al., 2020; Quentin et al., 2021; Varela et al., 2024). Consequently, performance observed after a second exposure may partly reflect carryover effects from the initial learning phase rather than transfer per se. An example of positive carryover comes from Marcus et al. (2007) who showed that 9-month-old infants, who fail on a visual SL task, will succeed on that task if it was preceded by an auditory SL task that contained the same underlying structural regularity.

Taken together, the adult literature provides mixed and often indirect evidence on the trajectory of linguistic cross-modal transfer, leaving open the question of whether cross-modal transfer is symmetric or shaped by modality-specific processing biases and task characteristics. Moreover, studies rarely examine transfer in both modality directions within the same participants, as testing one direction first may introduce exposure-order effects that influence performance in subsequent phases. To address this issue, the present study counterbalanced modality order across participants and explicitly tested for order effects, allowing cross-modal transfer to be assessed while controlling for potential biases related to exposure sequence.

Research Goals and Questions

In this study, we directly examine cross-modal transfer between uni-modal learning contexts while explicitly manipulating both modality direction (auditory- to-visual vs. visual-to-auditory) and order of exposure (visual-first vs. auditory-first). Furthermore, we assess learning using both a discrimination task and a more demanding recall task, allowing us to capture complementary aspects of learning (i.e., explicit reflection vs. implicit knowledge; Christiansen, 2019) and potential individual differences therein.

Specifically, we ask: (1) whether performance differs when the test modality matches or mismatches the training modality; (2) whether transfer is symmetric across sensory modalities; (3) whether prior exposure to one set of regularities influences learning of a subsequent set and (4) whether these factors interact in shaping the expression of statistical learning. By addressing these questions, the present study aims to clarify the conditions under which statistical learning generalizes across modalities and the extent to which such generalization reflects abstract versus modality-bound learning processes.

4.2 Materials and Methods

Given that SL has been proposed as a core mechanism in reading acquisition, the present study examines cross-modal transfer between auditory and visual language representations (i.e., spoken and written words). We deploy a sequential uni-modal design that allows a more direct test of transfer mechanisms than previous multi-modal paradigms.

In classic SL paradigms, participants first undergo a familiarization phase in which they are exposed to a continuous stream of stimuli containing recurring regularities without explicit instruction about the underlying structure. Learning in adults is typically assessed using a two-alternative forced-choice (2AFC) task, where participants select the more familiar sequence between a previously encountered item and a foil; above-chance performance is interpreted as evidence of successful learning. Despite its widespread use, the 2AFC method has important limitations. Because SL is largely implicit, explicit familiarity judgments may not accurately reflect the underlying learning process (see Christiansen, 2019 for a review). Performance can be influenced by response biases, meta-cognitive strategies, and individual differences in decision-making, while the binary scoring provides only coarse information and may yield unreliable estimates at the individual level (Frost et al., 2015; Siegelman and Frost, 2015; Isbilen et al., 2020). To mitigate these limitations, the present study complemented the traditional 2AFC test with a Statistically Induced Chunking Recall (SICR) task (Isbilen et al., 2020), a processing-based measure that indexes the formation of chunked representations and has been shown to provide more sensitive and reliable evidence of learning in adults and children (Isbilen et al., 2020; Kidd et al., 2020).

All data analysis code and experimental stimuli are openly available on the Open Science Framework at `https://osf.io/xfmnp`.

4.2.1 Participants

Forty adults participated in the study (17 males; $M_{\text{age}} = 23.68$ years, $SD = 3.85$, range: 18–30) who were recruited using the Prolific participant platform (`www.prolific.com`). Twenty-one participants were assigned to the Auditory–Visual (AV) condition, where participants learned an auditory language followed by a visual one, and nineteen to the Visual–Auditory (VA) condition, where participants learned a visual language followed by an auditory one. All participants provided informed consent prior to participation. Most participants were right-handed (37 right-handed, 3 left-handed). All were native English speakers who reported normal or corrected-

to-normal vision and normal hearing, and none reported a history of neurological, cognitive, or language-related disorders. Participants were unaware of the purpose of the experiment until completion of the study. Participants were compensated at the Prolific recommended rate of $8.00/hour.

4.2.2 Stimuli

The experiment used artificial linguistic stimuli composed of consonant–vowel (CV) syllables, adapted from classic statistical learning paradigms (Saffran et al., 1996a,b; Kidd et al., 2020). Syllables were combined to form trisyllabic words and were presented either visually (as written syllables) or auditorily (as spoken syllables), depending on the experimental phase. Two distinct artificial languages were used. Language 1 consisted of four trisyllabic words delivered in the auditory modality (*takapo, modipa, kibudu, and lomari*); Language 2 consisted of a different set of four trisyllabic words delivered in the visual modality (*kalomo, bumapo, dikita, and duripa*). Words from both languages were built using the same set of syllables (*bu, di, du, ka, ki, lo, ma, mo, pa, po, ri, ta*). The structure of the two languages were non-overlapping to prevent interference effects across phases (i.e., AV and VA). Across both languages, syllables were distributed such that no syllable consistently occurred in the same serial position within a word, minimizing the availability of cues other than the transitional probabilities between syllables.

4.2.3 Experimental Paradigm and Procedure

The overall structure of the experimental design is illustrated in Fig.4.2. After collecting demographic information (age, gender, native language, handedness), participants were asked to complete a language check. This was implemented to ensure that participants were all English native speakers and to exclude participants whose language background might impact task performance. The language checks consisted of listening to an English sentence and typing it correctly and then reading aloud *The Rainbow Passage* Fairbanks (1940), a standardized reading passage widely used in speech, language, and hearing research.

Subsequently, participants were randomly assigned to one of two groups that differed in the order of modality exposure. In the Visual–Auditory (VA) group, participants completed a visual familiarization phase followed by an auditory familiarization phase, whereas in the Auditory–Visual (AV) group this order was reversed. Each familiarization phase was followed by four test tasks: a two-alternative forced-choice (2AFC) task and a Statistically Induced Chunking Recall (SICR) task Isbilen

et al. (2020) administered first in the same modality as familiarization and then in the other modality. The 2AFC task was always administered before the SICR task, as findings from Isbilen et al. (2020) suggested that completing SICR can influence subsequent 2AFC performance, whereas completing 2AFC does not comparably affect SICR performance. Thus, each phase included both modality match and modality mismatch testing. In the second phase, the familiarization modality was reversed, and the same sequence of tasks was repeated. This counterbalanced design ensured that any effects of learning and transfer could not be attributed to a fixed order of modality exposure.

During familiarization, participants were exposed to three randomized blocks of a continuous stream created by concatenating tri-syllabic words without pauses or explicit boundary cues. Transitional probabilities were structured such that the probability between syllables within a word was 1, whereas the probability between syllables spanning word boundaries was .33. Each tri-syllabic word occurred 72 times in each stream. In the visual modality, syllables were presented sequentially at the center of the screen, whereas in the auditory modality they were presented as a continuous speech stream. In both modalities, syllables were presented at a rate of 3 Hz Batterink and Paller (2017). Participants were exposed to three blocks of the stream separated by 30-second pauses before proceeding to the testing phase.

Test materials consisted of stimuli for 2AFC and SICR tasks. For the 2AFC task, test trials paired one familiar word from the familiarization stream with one novel word constructed from the same syllable inventory but not presented during familiarization. Four novel word foils were created to avoid reproducing the transitional probabilities of the trained words and paired with the four familiar word items with all possible combinations, counterbalancing presentation order, resulting in a total of 16 trials. For the SICR task, test stimuli consisted of six-syllable sequences: eight sequences were experimental items formed by concatenating two familiar tri-syllabic words from the familiarization language; eight were random items, created by pseudo-randomly combining six syllables while avoiding familiar word-level structure (all transitional probabilities in the random items were zero). Participants were not informed of this manipulation. SICR stimuli were presented visually or auditorily, depending on task modality, and responses were collected via typed input (visual modality) or spoken repetition recorded via microphone (auditory modality).

The same principles of creating and organizing the stimulus materials were applied across modalities and phases, ensuring that any observed effects reflected learning and transfer of sequential structure rather than surface differences between stimulus sets.

The experiment was conducted online using the Gorilla platform Anwyl-Irvine

et al. (2020), a web-based tool for designing and running behavioral experiments, which enabled controlled stimulus presentation and response collection via web browsers.

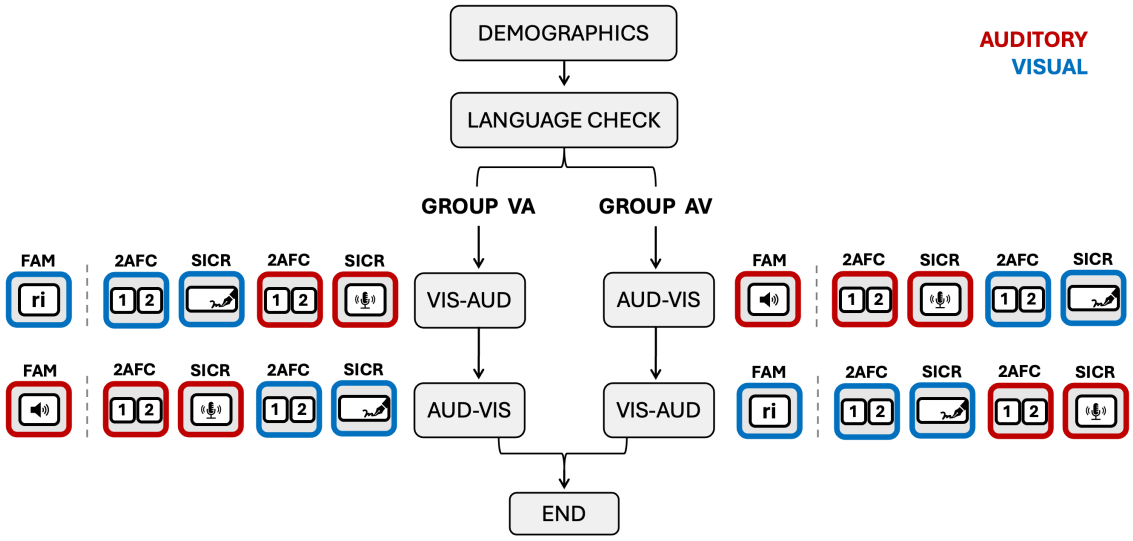


Figure 4.2: **Overview of the experimental design and procedure.** After completing the demographics and language background questionnaires, participants were assigned to one of two groups that differed in the order of modality exposure: Visual–Auditory (VA) or Auditory–Visual (AV). The experiment consisted of two phases. In each phase, participants completed a familiarization (FAM) in one modality (visual in blue, auditory in red), followed by four test tasks: a two-alternative forced-choice task (2AFC) and a Statistically Induced Chunking Recall (SICR) administered first in the same modality as familiarization and then in the other modality. In the second phase, the familiarization modality was reversed, and the same sequence of within- and cross-modal test tasks was repeated. This counterbalanced design controlled for order effects of modality exposure across participants.

4.2.4 Behavioral Data Analysis

Performance on the behavioral tasks was assessed using accuracy-based measures. All statistical analyses were conducted in R (version 4.3.2 (2023-10-31) using generalized linear mixed-effects models implemented using the lme4 package, within the RStudio integrated development environment (Posit team, 2023). Mixed-effects modeling was chosen to account for the repeated-measures structure of the data and for individual differences in overall task performance. In the 2AFC task, accuracy was defined as the proportion of correct responses on each test trial. Trial-level accuracy (correct vs. incorrect) was analyzed using logistic mixed-effects models

with random intercepts for participants. Fixed effects included transfer type (same-modality vs. cross-modality), order of exposure (visual-first vs. auditory-first), , and training modality (auditory vs. visual), as well as their interactions. Counterbalancing group (AV vs. VA) was not included in the final models, as it did not correspond to a theoretically meaningful factor and differed only in task order. This modeling approach allowed us to directly test whether cross-modal generalization occurred, whether performance differed as a function of learning phase, and whether transfer depended on the modality in which learning took place.

Coding of participants' responses in SICR task was performed following the methods described by Isbilen et al. (2020). Performance was analyzed using trigram recall accuracy as the primary dependent measure, since it measures sensitivity to the tri-syllabic words present in the input languages and provides a more direct comparison to 2AFC, which also presents tri-syllabic words (Isbilen et al., 2022). Supplementary analyses were also conducted using total syllable recall as a secondary measure. Under this scoring method, a syllable was counted as correct only when reproduced in the appropriate serial position within the six-syllable response sequence. These analyses yielded a pattern comparable to the trigram-based results and are therefore not reported in detail. For each trial, the number of correctly recalled trisyllabic strings (out of the two possibilities in each 6-syllable test item) was modeled as a binomial outcome. Analyses were conducted at the trial level using logistic mixed-effects models with random intercepts for participants. Fixed effects included item type (familiar vs. random sequences), transfer type (same-modality vs. cross-modal), order of exposure (visual-first vs. auditory-first), and training modality, along with their theoretically motivated interactions. This approach allowed us to assess whether structured sequences were recalled more accurately than random sequences, and whether learning generalized across modalities depending on learning phase and modality of familiarization. In addition, to quantify learning relative to baseline performance in SICR task, we computed differential scores by subtracting performance on random (unstructured) sequences from performance on familiar (structured) sequences for each participant and condition, thereby isolating sensitivity to the statistical regularities of the input.

For both tasks, model structures were determined using a theory-driven approach, focusing on interactions relevant to the research questions rather than on exhaustive inclusion of all higher-order terms. Model comparisons based on likelihood-ratio tests were used to evaluate the contribution of complex interactions.

4.3 Results

To address the research questions outlined above, we first tested whether statistical learning generalized across modalities by comparing performance in modality-match and modality-mismatch trials. We then examined whether transfer differed as a function of modality direction (auditory-to-visual vs. visual-to-auditory), whether exposure order (visual-first vs. auditory-first) influenced learning outcomes across phases, and whether these factors interacted.

As a preliminary check, we assessed whether performance differed between participants assigned to the Auditory–Visual (AV) and Visual–Auditory (VA) groups, to ensure that any observed effects could be attributed to exposure order rather than pre-existing group differences. We therefore directly compared performance between groups across both tasks.

In the 2AFC task, mean accuracy was comparable between groups (AV: $M = 8.79$, $SD = 2.46$; VA: $M = 9.25$, $SD = 2.50$), and a linear model confirmed that group did not significantly predict performance, $t(158) = 1.18$, $p = .239$. In the SICR task, there was likewise no reliable difference between groups for either experimental sequences, $t(37.92) = -1.79$, $p = .081$, or random sequences, $t(38.87) = -1.32$, $p = .195$. To further assess potential group differences while controlling for baseline performance, we computed a differential trigram recall score for each participant by subtracting performance on random sequences from performance on experimental sequences. This analysis confirmed that the difference between AV and VA groups was not significant, $t(34.81) = -0.88$, $p = .384$, indicating comparable learning-related gains across groups. Taken together, these analyses indicate that performance was largely comparable between the AV and VA groups across both tasks.

4.3.1 Transfer Effects

Cross-modal transfer of statistical learning was defined as above-chance learning when the testing modality differed from the training modality (modality mismatch), relative to trials in which training and testing occurred in the same modality (modality match). Accordingly, the data were collapsed into modality-match (auditory–auditory and visual–visual) and modality-mismatch (auditory–visual and visual–auditory) categories, irrespective of experimental condition (AV vs. VA), allowing direct comparison of unimodal versus cross-modal expressions of learning (see Fig.4.3).

Before collapsing across experimental conditions, we compared performance between the AV and VA groups separately for modality-match and modality-mismatch trials. No significant difference emerged between groups in 2AFC modality-match

performance, $t(78) = 0.27$, $p = .79$, indicating comparable learning. For modality-mismatch trials, however, a marginally significant difference was observed, $t(73.33) = -2.02$, $p = .047$, with the VA group showing slightly higher accuracy than the AV group. To assess whether performance differed between groups prior to collapsing, we also compared overall accuracy across AV and VA participants in the SICR task. Trigram recall did not differ significantly between groups for either experimental sequences, $t(37.92) = -1.79$, $p = .081$, or random sequences, $t(38.87) = -1.32$, $p = .195$. These results indicate broadly comparable performance across the AV and VA groups. Given that group differences were marginal and inconsistent across conditions and tasks, data were collapsed across groups in subsequent analyses to focus on overall cross-modal transfer effects and increase statistical power.

In the 2AFC task, performance was significantly above chance in both modality-match and modality-mismatch trials. Mean accuracy in modality-match trials was 56.09% ($SD = 11.37$), reliably higher than chance level (50%), $t(39) = 4.01$, $p = .0008$, corresponding to a medium effect size ($d = 0.54$). Likewise, performance in modality-mismatch trials averaged 56.48% ($SD = 11.15$), also exceeding chance, $t(39) = 4.26$, $p = .0004$, with a comparable medium effect size ($d = 0.58$). Critically, a paired comparison revealed no difference between modality-match and modality-mismatch performance, $t(39) = -0.18$, $p = .86$, with a negligible effect size ($d = -0.03$), indicating that learning generalized across modalities at a level indistinguishable from within-modality performance.

SICR performance likewise revealed reliable learning in both modality-match and modality-mismatch conditions (see Appendix, Fig. A4.1). In modality-match trials, participants recalled more trigrams for experimental ($M = 3.23$, $SD = 1.88$) than random ($M = 2.11$, $SD = 1.32$) sequences, $t(70.07) = 3.06$, $p = .003$. A similar pattern was observed in modality-mismatch trials ($M_{\text{exp}} = 3.35$, $SD = 1.96$; $M_{\text{ran}} = 2.52$, $SD = 1.81$), $t(79.45) = 1.99$, $p = .050$. However, recall did not differ between modality-match and modality-mismatch conditions for either experimental sequences, $t(78.97) = -0.30$, $p = .76$, or random sequences, $t(73.34) = -1.17$, $p = .24$, indicating comparable performance across conditions. This pattern was confirmed by analyses on difference scores (see Fig.4.3). Learning was significant in both modality-match ($M = 1.11$, $t(39) = 4.28$, $p < .001$) and modality-mismatch trials ($M = 0.83$, $t(40) = 3.43$, $p = .001$). Critically, the magnitude of learning did not differ between conditions, $t(39) = 0.93$, $p = .36$, indicating that cross-modal transfer was as robust as unimodal learning.

Together, these findings demonstrate robust cross-modal transfer of statistical learning, with participants generalizing learned structure across modalities to the same extent as within-modality learning.

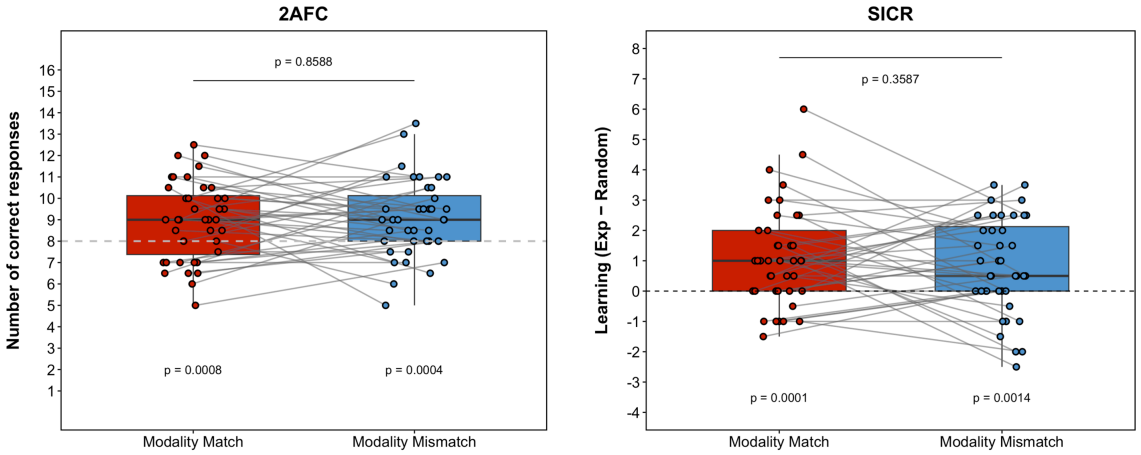


Figure 4.3: **Performance as a function of transfer. Left panel (2AFC task):** Mean number of correct responses collapsed across trials in which the testing modality matched the training modality (Modality Match) or differed from it (Modality Mismatch). The dashed horizontal line indicates chance performance. **Right panel (SICR task):** Difference scores (experimental – random) are shown for modality-match and modality-mismatch conditions, indexing learning while controlling for baseline recall performance. Values above zero indicate successful statistical learning. Each point represents an individual participant, boxplots summarize group-level distributions, and grey lines connect paired observations. The dashed horizontal line indicates zero (no learning). Reported p-values refer to one-sample tests against chance (reported below each box) and paired comparisons between trials (reported above the connecting lines).

4.3.2 Modality Directionality Effects

To examine whether the direction of cross-modal transfer modulated statistical learning (SL), performance was grouped according to the modality in which regularities were acquired and the modality in which they were tested. Specifically, we contrasted conditions in which learning occurred in the auditory modality and was tested visually (auditory-to-visual) with conditions in which learning occurred visually and was tested auditorily (visual-to-auditory), collapsing across order and group (AV vs. VA; see Fig.4.4).

Prior to collapsing across AV and VA groups, we compared performance separately for each modality direction. For 2AFC accuracy, a significant difference was found for conditions involving auditory-based learning (auditory-to-auditory and auditory-to-visual), $t(74.03) = -2.65$, $p = .010$, with higher accuracy in the VA group. In contrast, no significant difference was observed for conditions involving visual-based learning, $t(77.11) = 1.04$, $p = .30$. SICR results revealed a mixed pattern. In auditory-to-visual conditions, no differences were observed between AV

and VA groups in either experimental trials, $t(36.12) = 0.25$, $p = .805$, or random trials, $t(77.84) = -1.89$, $p = .062$. In contrast, in the visual-to-auditory condition, the VA group showed significantly higher performance than the AV group in both experimental trials, $t(77.69) = -2.39$, $p = .019$, and random trials, $t(77.72) = -2.29$, $p = .024$. Because these differences were not consistent across modality directions and, critically, were present in both experimental and random conditions, they likely reflect general performance differences rather than differences in learning. Therefore, data were collapsed across AV and VA groups in subsequent analyses.

In the 2AFC task, performance was significantly above chance in both modality directions. Accuracy in the auditory-to-visual condition was above chance, $t(39) = 3.87$, $p < .001$, $d = 0.61$, as was accuracy in the visual-to-auditory condition, $t(39) = 2.51$, $p = .008$, $d = 0.40$. A direct within-subject comparison revealed no significant difference between modality directions, $t(39) = 1.11$, $p = .27$, with a small effect size ($d = 0.18$), indicating comparable behavioral performance regardless of transfer direction.

A similar pattern emerged in the SICR task (see Appendix, Fig. A4.2). In the auditory-to-visual condition, participants recalled more trigrams for familiar than random sequences ($M_{\text{exp}} = 3.28$, $SD = 2.32$; $M_{\text{ran}} = 2.38$, $SD = 1.71$), $t(73.56) = 2.00$, $p = .049$. In the visual-to-auditory condition, this advantage was also significant ($M_{\text{exp}} = 3.30$, $SD = 1.67$; $M_{\text{ran}} = 2.29$, $SD = 1.64$), $t(77.96) = 2.73$, $p = .008$. However, performance did not differ between modality directions for either experimental sequences, $t(72.77) = -0.04$, $p = .97$, or random sequences, $t(78.97) = 0.24$, $p = .81$. Analyses of difference scores (see Fig.4.4) confirmed reliable learning in both directions, with significant effects for auditory-to-visual transfer ($M = 0.90$, $t(40) = 3.18$, $p = .003$) and visual-to-auditory transfer ($M = 1.01$, $t(39) = 3.55$, $p = .001$). Critically, the magnitude of learning did not differ between directions, $t(39) = -0.27$, $p = .79$, indicating that cross-modal generalization was symmetric across sensory modalities.

Overall, these results indicate that cross-modal statistical learning is bidirectional and symmetric, with comparable generalization from audition to vision and from vision to audition.

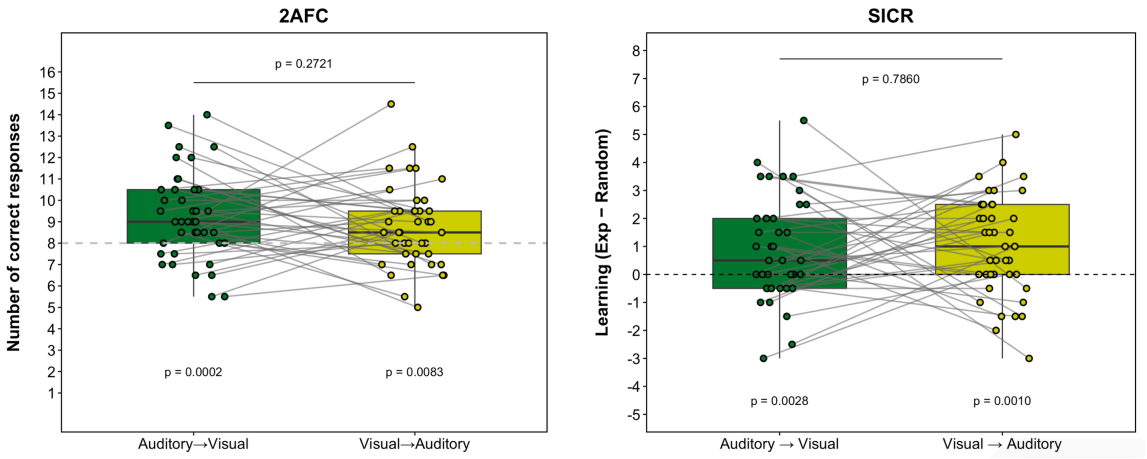


Figure 4.4: **Performance as a function of modality direction. Left panel (2AFC task):** Mean number of correct responses for conditions in which participants were trained in the auditory modality and tested in the visual modality (Auditory-to-Visual) or trained in the visual modality and tested in the auditory modality (Visual-to-Auditory). The dashed horizontal line indicates chance performance. **Right panel (SICR task):** Number of correctly recalled trigrams for experimental stimuli (concatenations of two familiarized words) and random stimuli, shown separately for Auditory-to-Visual and Visual-to-Auditory transfer directions. Each dot represents an individual participant; box-plots summarize group-level performance, and grey lines connect paired observations within participants. Reported p-values refer to one-sample tests against chance (reported below each box) and paired comparisons between conditions (reported above the connecting lines).

4.3.3 Exposure Order Effects

To assess whether prior exposure to one set of regularities influenced performance on a subsequent set of different regularities, we compared accuracy in tests administered after the first familiarization phase (Phase 1) with accuracy in tests administered after the second familiarization phase (Phase 2), collapsing across training modality, test modality, and group (see Fig.4.5).

To determine whether performance differed between AV and VA groups as a function of exposure order (first vs. second familiarization phase), we conducted preliminary comparisons separately for Phase 1 and Phase 2. For 2AFC accuracy, no significant differences were found between groups in either Phase 1, $t(78) = 0.15$, $p = .88$, or Phase 2, $t(77.11) = 1.04$, $p = .30$. In the SICR task, no differences were observed between AV and VA groups in either experimental trials, $t(35.06) = -0.22$, $p = .831$, or random trials, $t(77.66) = -1.96$, $p = .053$, during the first phase. In the second phase, however, the VA group showed higher performance than the AV group in both experimental trials, $t(77.69) = -2.39$, $p = .019$, and random

trials, $t(77.72) = -2.29$, $p = .024$. Because this difference in the second phase was present in both experimental and random trials, it likely reflects general performance differences rather than differences in learning. Therefore, data were collapsed across AV and VA groups in subsequent analyses to focus on the overall effects of exposure order.

In the 2AFC task, performance after the first familiarization phase was modestly but significantly above chance ($M = 53.20\%$, $SD = 9.94$), $t(39) = 2.04$, $p = .024$, $d = 0.32$. Performance after the second familiarization phase was substantially higher ($M = 59.38\%$, $SD = 13.94$), $t(39) = 4.25$, $p < .001$, $d = 0.67$. A within-subject comparison confirmed a significant improvement from the first to the second language, $t(39) = -2.38$, $p = .022$, with a medium effect size ($d = 0.38$), indicating enhanced behavioral performance following a second round of exposure to a different set of regularities.

In the SICR task (see Appendix, Fig. A4.3), participants recalled significantly more trigrams for familiar than random sequences during the first familiarization phase ($M_{\text{exp}} = 3.31$, $SD = 2.22$; $M_{\text{ran}} = 2.04$, $SD = 1.67$), $t(72.36) = 2.90$, $p = .005$. In the second familiarization phase, the advantage for experimental sequences was numerically smaller and did not reach significance ($M_{\text{exp}} = 3.27$, $SD = 1.82$; $M_{\text{ran}} = 2.62$, $SD = 1.63$), $t(79.05) = 1.69$, $p = .094$. Importantly, trigram recall did not differ between the first and second familiarization phases for either experimental sequences, $t(75.29) = 0.10$, $p = .92$, or random sequences, $t(78.82) = -1.59$, $p = .12$. Difference score analyses indicated significant learning in both the first ($M = 1.28$, $t(39) = 4.33$, $p < .001$) and second phase ($M = 0.65$, $t(40) = 2.43$, $p = .019$). The magnitude of learning did not differ significantly between phases, $t(39) = 1.60$, $p = .12$, suggesting comparable learning across familiarization orders. To determine whether the reduced raw contrast in Phase 2 reflected weaker learning or improved recall, participant-level trigram difference scores (experimental − random) were computed. Difference scores (see Fig.4.5) were significantly above zero in both Phase 1, $t(39) = 3.16$, $p = .003$, and Phase 2, $t(39) = 3.24$, $p = .002$, indicating reliable learning in both phases. Critically, learning magnitude did not differ between phases, $t(39) = 0.19$, $p = .85$. This suggests that the smaller raw experimental-versus-random contrast in Phase 2 was driven by improved recall of random items rather than reduced statistical learning.

Overall, these results indicate that prior exposure enhances general task performance, but does not modulate the strength of statistical learning, which remains stable across successive learning phases.

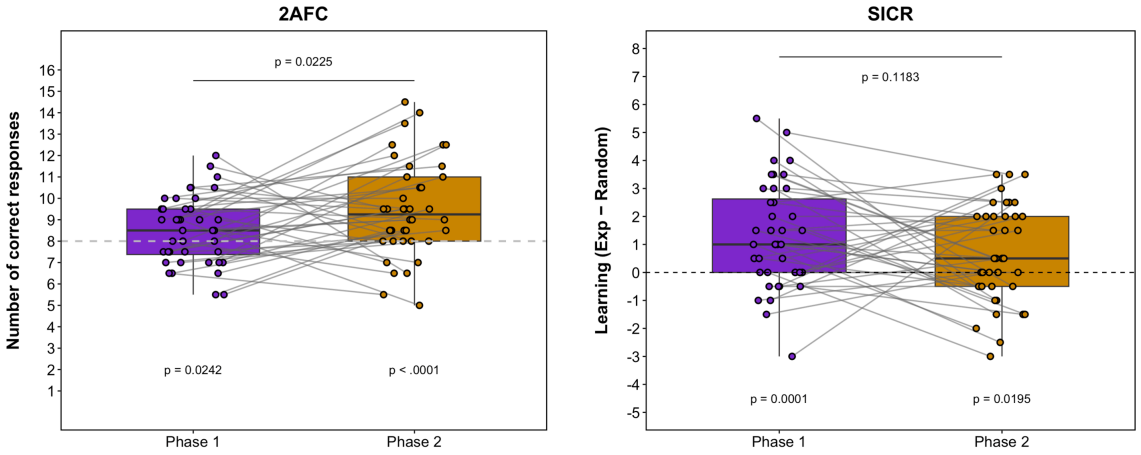


Figure 4.5: **Performance as a function of language order. Left panel (2AFC task):** Mean number of correct responses for Phase1 and Phase2. The dashed horizontal line indicates chance performance. **Right panel (SICR task):** Number of correctly recalled trigrams for experimental stimuli (concatenations of two familiarized words) and random stimuli, shown separately for Phase1 and Phase2. Each dot represents an individual participant; box-plots summarize group-level performance, and grey lines connect paired observations within participants. Reported *p*-values refer to one-sample tests against chance (reported below each box) and paired comparisons between languages (reported above the connecting lines).

4.3.4 Interaction Effects

To examine whether transfer effects depended jointly on modality direction and language order, we analyzed trial-level performance using mixed-effects logistic regression models. These models included transfer (modality match vs. mismatch), training modality (auditory vs. visual), order (first vs. second familiarization phase), and their interactions as fixed effects, with random intercepts for participants to account for repeated measurements. This approach allowed us to test theoretically relevant higher-order interactions while appropriately modeling the structure of the data.

In the 2AFC task (see Table1), there was no overall effect of transfer, indicating comparable performance for modality-matched and modality-mismatched trials ($\beta = 0.01$, $SE = 0.16$, $z = 0.08$, $p = .938$). However, accuracy was higher following the second familiarization phase, reflecting an order effect ($\beta = 0.41$, $SE = 0.18$, $z = 2.28$, $p = .023$). Training modality alone did not reliably affect performance ($\beta = 0.15$, $SE = 0.18$, $z = 0.83$, $p = .407$).

Critically, exposure order interacted with training modality ($\beta = -0.56$, $SE = 0.28$, $z = -2.02$, $p = .044$), indicating that the improvement from the first to the sec-

ond phase depended on whether learning occurred in the auditory or visual modality. This interaction was further qualified by a significant three-way interaction between transfer, order, and training modality ($\beta = 0.64$, $SE = 0.32$, $z = 1.97$, $p = .049$), suggesting that modality-specific effects varied as a function of both learning phase and training modality.

No interaction between transfer and order ($\beta = -0.05$, $SE = 0.23$, $z = -0.24$, $p = .811$) or between transfer and training modality ($\beta = -0.33$, $SE = 0.23$, $z = -1.48$, $p = .139$) was observed.

Additional analyses focusing specifically on cross-modal trials revealed no asymmetry in transfer direction: performance did not differ between auditory-to-visual and visual-to-auditory transfer ($\beta = -0.12$, $SE = 0.11$, $z = -1.08$, $p = .280$), nor did transfer direction interact with group/order ($\beta = -0.27$, $SE = 0.23$, $z = -1.16$, $p = .246$). Thus, cross-modal transfer was symmetric in the 2AFC task.

Table 1: **Mixed-effects logistic regression results for the 2AFC task**

Predictor	β	SE	z	p
Intercept	0.133	0.123	1.08	.279
Transfer (Match)	0.012	0.156	0.08	.938
Order (Second)	0.414	0.182	2.28	.023*
Training Modality (Visual)	0.149	0.180	0.83	.407
Transfer × Order	-0.055	0.230	-0.24	.811
Transfer × Training Modality	-0.335	0.227	-1.48	.139
Order × Training Modality	-0.563	0.279	-2.02	.044*
Transfer × Order × Training Modality	0.636	0.323	1.97	.049*

To determine whether similar interaction patterns would emerge in a more demanding recall measure, we analyzed trigram recall accuracy in the SICR task using mixed-effects logistic regression (see Table2).

Participants recalled significantly more trigrams from experimental than random sequences ($\beta = 0.37$, $SE = 0.11$, $z = 3.48$, $p < .001$), confirming successful statistical learning.

In contrast to the 2AFC results, recall performance showed a strong effect of transfer: accuracy was lower when training and test modalities matched ($\beta = -0.79$, $SE = 0.16$, $z = -4.84$, $p < .001$).

To clarify this pattern, we conducted additional analyses using participant-level difference scores (familiar − random), which provide a measure of learning independent of baseline recall performance. These analyses revealed a strong effect of

modality direction: learning was substantially greater in the auditory-to-visual condition than in the visual-to-auditory condition, $t(39) = 19.12$, $p < .001$. This indicates a pronounced asymmetry in cross-modal transfer in the SICR task. Critically, this result suggests that the apparent disadvantage for modality-matched trials in the raw accuracy data was not driven by transfer per se, but rather by differences in learning strength across modality directions. In other words, learning acquired in the auditory modality transferred more robustly than learning acquired in the visual modality, consistent with modality-specific constraints on retrieval under more demanding recall conditions.

Importantly, the transfer effect was modulated by both language order and training modality. A significant interaction between transfer and order emerged ($\beta = 0.64$, $SE = 0.15$, $z = 4.17$, $p < .001$), indicating that the recall disadvantage associated with modality-matched testing was smaller for the second language. In addition, transfer interacted with training modality ($\beta = 0.56$, $SE = 0.15$, $z = 3.64$, $p < .001$), such that modality-matched costs were larger following auditory training than visual training. No interaction between item type and transfer was observed ($\beta = 0.17$, $SE = 0.15$, $z = 1.09$, $p = .275$), indicating that these effects reflected general recall dynamics rather than differences specific to learned regularities.

Table 2: **Mixed-effects logistic regression results for the SICR task**

Predictor	β	SE	z	p
Intercept	-1.592	0.149	-10.67	$< .001$***
Item Type (Experimental)	0.369	0.106	3.48	$< .001$***
Transfer (Match)	-0.795	0.164	-4.84	$< .001$***
Order (Second)	-0.228	0.107	-2.13	.033*
Training Modality (Visual)	-0.235	0.107	-2.20	.028*
Item Type × Transfer	0.167	0.153	1.09	.275
Transfer × Order	0.643	0.154	4.17	$< .001$***
Transfer × Training Modality	0.560	0.154	3.64	$< .001$***

Taken together, these results suggest a potential order effect in the 2AFC task, with higher accuracy following the second familiarization phase. However, this pattern was not mirrored in the SICR task, where learning was evident for the first language but did not significantly increase for the second. Importantly, internal consistency differed markedly across tasks. The 2AFC measure showed relatively low reliability (Cronbach's $\alpha = .52$), with very weak inter-item correlations, consistent with previous evidence that forced-choice tasks provide noisy estimates of statistical learning (Frost et al., 2015; Siegelman and Frost, 2015). In contrast, the

SICR task demonstrated substantially higher reliability across conditions. For experimental sequences, internal consistency was high in both modalities (auditory: $\alpha = .78$; visual: $\alpha = .76$), indicating strong coherence across trials. Reliability for random sequences was lower but remained moderate to high (auditory: $\alpha = .57$; visual: $\alpha = .76$).

Because the order effect emerged only in the less reliable 2AFC measure and not in the more reliable recall-based SICR task, its robustness should be interpreted with caution. This discrepancy suggests that exposure order may influence the behavioral expression of learning in recognition-based judgments more strongly than in free recall, although further evidence is needed to confirm the reliability of this effect.

4.4 General Discussion

A central question motivating this study was whether statistical learning (SL) gives rise to representations that are tied to the sensory modality in which regularities are experienced, or whether learning operates over more abstract representations that can generalize across modalities. Put differently, we asked whether sequential regularities learned through audition or vision remain modality-bound, or whether they transfer flexibly between spoken and written forms. By combining a sequential unimodal learning design with bidirectional transfer tests and complementary outcome measures, the present study provides a direct test of the extent and limits of cross-modal transfer in linguistic SL.

The present study investigated cross-modal statistical learning using a sequential unimodal design that allowed transfer to be assessed within participants across both auditory and visual modalities. Participants learned two artificial languages presented in different modalities and were tested after each familiarization phase using both a traditional 2AFC task and a processing-based SICR recall measure. To interpret these findings, we framed cross-modal transfer in terms of two alternative models introduced in the Introduction. Model 1 assumes that SL operates within modality-specific systems and that transfer is a secondary, post-learning mapping process. Model 2, in contrast, proposes that for familiar linguistic stimuli, unimodal input automatically activates shared orthographic–phonological representations, such that transfer is an intrinsic consequence of learning. Overall, the present findings are more consistent with Model 2, while also revealing task-dependent constraints that suggest a more nuanced, hybrid architecture that is affected by the requirements of the post-exposure test phase.

Evidence for Cross-Modal Transfer in SL

A central question of the present study was whether statistical regularities acquired in one modality would generalize to another. Across both the 2AFC and SICR tasks, participants demonstrated reliable learning not only when training and testing occurred within the same modality but also when the modalities differed. Crucially, in the 2AFC task, performance did not differ between modality-match and modality-mismatch conditions. This pattern is difficult to reconcile with a strong version of Model 1, which predicts that cross-modal transfer should depend on an additional mapping process and may therefore be weaker or less reliable than within-modality performance. Instead, the observed equivalence between conditions suggests that transfer was not a secondary or effortful process, but rather an intrinsic outcome of

learning, consistent with Model 2.

These findings extend previous demonstrations of cross-modal statistical learning, which have largely relied on non-linguistic stimuli or multi-modal presentation paradigms (e.g., tones, shapes, or environmental sounds; see Conway and Christiansen, 2006; Mitchel and Weiss, 2011a; Walk and Conway, 2011; Duato et al., 2023; Tang et al., 2026). In many of these studies, transfer has been interpreted as evidence that statistical learning can operate over representations that are not tied to modality-specific sensory features. However, most prior work relied either on simple non-linguistic stimuli or on paradigms in which information from multiple modalities is presented simultaneously, making it difficult to isolate the mechanisms underlying transfer. Moreover, these reports of similar SL in different modalities rarely used stimuli that had pre-established cross-modal equivalencies, a design that formed the basis of the present study.

The present findings extend this literature by demonstrating cross-modal transfer using linguistic materials presented sequentially across modalities. Prior cross-modal studies employing linguistic input have been comparatively scarce and have largely relied on artificial grammar paradigms or developmental samples Conway and Christiansen (2006); Walk and Conway (2011); Shafto et al. (2012); Emberson and Rubinstein (2016). By contrast, the current design used sequential uni-modal exposure, allowing transfer to be assessed more directly. The finding that statistical regularities learned in one modality readily transferred to performance in another modality therefore suggests that learners were able to extract structure at a level that is not strictly bound to the sensory characteristics of the input. At the same time, the SICR results qualify this interpretation. Although cross-modal transfer was evident, recall performance revealed stronger sensitivity to modality-related factors. This pattern suggests that, while learning may operate over shared representations, the expression of those representations depends on task demands, with modality-specific representations becoming more relevant under more demanding retrieval conditions.

An important feature of the present stimuli is that they were not only cross-modal (auditory vs. visual), but also remained within the linguistic domain, as spoken and written language are linked through well-established phonological-to-orthographic correspondences. For literate adults and children, seeing printed syllables can automatically activate phonological representations, just as hearing speech can evoke associated orthographic forms (Cortese and Balota, 2012; Frost, 2012). This bidirectional mapping is central to reading, particularly during early acquisition when decoding depends heavily on translating visual symbols into speech-based representations. Consistent with this view, phonological skills are among the strongest predictors of reading success in both children and adults (Wagner and Torgesen,

1987; Hulme et al., 2012; Melby-Lervåg et al., 2012).The robust transfer observed here may therefore reflect not only modality-general statistical learning, but also the operation of highly practiced links between orthographic and phonological systems that allow regularities learned in one format to support performance in the other.

Taken together, these findings align with proposals that statistical learning reflects domain-general computational mechanisms operating within modality-specific processing systems Frost et al. (2015). Within this framework, regularities can be encoded at a level that supports cross-modal access, while the perceptual characteristics of the input continue to shape how these representations are expressed in behavior. Consistent with recent evidence that learning performance depends on properties of the input such as modality and presentation format Lukics and Lukács (2022), the present results suggest that cross-modal transfer is possible but not entirely independent of the conditions under which learning occurs.

In the context of language, such cross-modal access is particularly relevant, as spoken and written forms rely on shared phonological and lexical representations despite being conveyed through different sensory channels. The present findings therefore provide evidence that statistical learning mechanisms can operate over linguistic regularities in a way that supports communication across modalities, offering a closer approximation to the multi-modal nature of real-world language use than studies relying solely on non-linguistic stimuli.

No Evidence for Directional Asymmetry in Cross-Modal Transfer

A second question addressed by the present study concerned whether cross-modal transfer was asymmetric across sensory directions. Across both 2AFC and SICR tasks, learning transferred successfully from audition to vision and from vision to audition, with no reliable differences between directions. Thus, the behavioral expression of SL did not depend on whether regularities were acquired in the auditory modality and tested visually or vice versa, providing evidence for bidirectional cross-modal transfer. This symmetry provides further support for Model 2. If transfer required a sequential mapping process (as in Model 1), one might expect asymmetries reflecting modality-specific constraints or differences in processing efficiency. Instead, the absence of directional effects suggests that both modalities accessed a common representational format, consistent with automatic co-activation of orthographic and phonological representations in literate adults.

This pattern of transfer aligns with adult studies reporting symmetric generalization across modalities Yildirim and Jacobs (2015); Tang et al. (2026), and contrasts with developmental accounts proposing a relative advantage for auditory-

driven transfer during early literacy acquisition. In children, stronger associations between auditory statistical learning and reading-related skills have been interpreted as reflecting the functional dominance of phonological representations when learners rely heavily on mapping print to sound Qi et al. (2019); Spencer et al. (2015); Gabay and Holt (2015). From this perspective, learning in the auditory domain may exert a stronger influence on visual processing during early reading development, consistent with the central role of phonological decoding in literacy acquisition Wagner and Torgesen (1987); Hulme et al. (2012); Melby-Lervåg et al. (2012).

The absence of directional asymmetries in the present data suggests that such biases may not characterize mature learners. Instead, the findings are consistent with the idea that literacy experience strengthens bidirectional links between orthographic and phonological representations, allowing information learned in one modality to be accessed equivalently from another Frost et al. (2019); Ren and Wang (2024). As these mappings become more integrated with reading experience, modality-specific biases may diminish, supporting more balanced cross-modal interactions in skilled readers. Thus, the degree of cross-modal transfer in a linguistic SL paradigm could form a useful diagnostic tool for the assessment of individual differences in the strength of phonological-orthographic mappings in early readers, thereby enabling interventions to strengthen these mappings.

More broadly, the present findings help clarify mixed evidence in the adult literature. While some studies have reported auditory advantages Conway and Christiansen (2005); Robinson and Sloutsky (2007), such asymmetries may be more likely with non-linguistic or unfamiliar stimuli. In contrast, linguistic materials appear to engage shared representations that support balanced cross-modal access. In this sense, modality-specific advantages may reflect properties of the stimuli and task demands rather than inherent constraints of the statistical learning system.

Finally, the present design provides a stronger test of directional asymmetry than many previous studies. Prior work has often examined only one transfer direction or tested directions sequentially (Conway and Christiansen, 2006; Mitchel and Weiss, 2011a; Walk and Conway, 2011), making it difficult to rule out exposure-order effects. By counterbalancing modality order and assessing both transfer directions within the same experimental design, the current study reduces this potential confound. Taken together, the results suggest that, in adult learners and with familiar linguistic materials, cross-modal statistical learning operates in a bidirectional manner. Directional constraints observed in developmental contexts may therefore reflect transient learning demands associated with early literacy acquisition rather than stable properties of the statistical learning system.

Task-Specific Effects of Exposure Order

Because participants were exposed to two sets of regularities sequentially, we also examined whether prior experience with the first language influenced performance during the second learning phase. In the 2AFC task, accuracy was higher for the second language, indicating an order effect. In contrast, performance in the SICR task showed clear evidence of learning for the first language that did not increase for the second. This dissociation suggests that exposure order influenced recognition-based judgments (2AFC task) more than representation-based learning (SICR task).

The improvement observed in the 2AFC task is consistent with previous evidence showing that prior experience with one statistical structure can shape the processing of subsequent input by establishing expectations, strengthening sensitivity to regularities, or promoting more efficient learning strategies Gómez and Gerken (2000); Conway and Christiansen (2005). In sequential learning contexts, exposure to an initial regularity can create memory traces or processing biases that influence how subsequent input is interpreted Benitez et al. (2020); Quentin et al. (2021); Varela et al. (2024). From this perspective, the higher accuracy observed for the second language may reflect facilitation arising from familiarity with the statistical structure of the task or with the demands of the forced-choice test (i.e., practice effects), hence partly reflecting task-specific noise or decision processes rather than robust changes in underlying representations. This reinforces the importance of considering measurement reliability when interpreting apparent learning effects in statistical learning paradigms.

Importantly, this pattern does not necessarily indicate stronger learning of the second statistical regularity itself. Instead, it likely reflects a change in how learned information was expressed behaviorally. Participants may have become more attuned to the presence of structure, more confident in making familiarity-based decisions, or more efficient in applying decision criteria during the second test phase.

The absence of a comparable increase in SICR performance across the two familiarization phases suggests that exposure order primarily affected decision-based judgments rather than the underlying strength of learned representations. Recall-based performance already showed clear evidence of learning following the first familiarization phase and remained stable thereafter, indicating that participants were able to form chunk-like representations early in exposure. This dissociation aligns with previous work suggesting that forced-choice measures may be particularly sensitive to response strategies, decision thresholds, or meta-cognitive factors (Frost et al., 2015; Siegelman and Frost, 2015; Isbilen et al., 2020), whereas recall-based tasks provide a more direct index of representation-based learning (Isbilen et al., 2020, 2022).

Within the framework of the two models, these findings further argue against a

simple sequential-transfer account (Model 1). If transfer depended strongly on prior learning episodes, one might expect consistent order effects across measures. Instead, the selective emergence of order effects in the 2AFC task suggests that exposure history primarily modulates how learning is expressed, not how it is acquired.

Taken together, these findings highlight the importance of controlling for exposure order in sequential statistical learning designs. When learners encounter multiple regularities in succession, performance during later phases may partly reflect carryover effects from prior exposure rather than transfer between specific structures. More broadly, the results underscore the value of combining multiple behavioral measures of statistical learning, as order-related influences may differentially affect recognition-based and recall-based indices of learning.

Task-Dependent Expression of Cross-Modal Transfer

Beyond the main effects, we examined whether transfer patterns depended jointly on modality, learning order, and task demands. Across analyses, a clear dissociation emerged between the recognition-based 2AFC task and the recall-based SICR measure.

In the 2AFC task, transfer was generally robust and modality-independent, as performance did not differ between modality-match and modality-mismatch conditions. However, interaction analyses revealed that the behavioral expression of transfer depended on learning order and training modality. In particular, 2AFC accuracy was higher for the second familiarization phase, and a three-way interaction between transfer, order, and training modality indicated that modality-related advantages varied depending on both whether a language was learned first or second and the modality in which it was acquired. Specifically, modality-matched advantages were strongest when the second language was learned auditorily, whereas visual training in the second phase attenuated modality-specific effects. These findings suggest that recognition-based judgments are sensitive to contextual factors such as learning history and modality context. Consistent with previous concerns about forced-choice measures, familiarity judgments may reflect not only the strength of learned representations but also decision processes, response strategies, or meta-cognitive factors that shape how statistical knowledge is expressed behaviorally Frost et al. (2015); Siegelman and Frost (2015).

In contrast, the SICR task revealed a different pattern. Although participants showed robust learning overall, recall performance exhibited systematic modality-related constraints. Specifically, recall was lower when training and test modalities matched, and this transfer effect was modulated by both language order and train-

ing modality. Modality-matched costs were reduced for the second language and were larger following auditory than visual training. Importantly, these effects were independent of item type, indicating that they reflected general retrieval dynamics rather than differences in the acquisition of experimental versus random sequences. Crucially, this apparent disadvantage for modality-matched trials should not be interpreted as a genuine cost of transfer. Rather, it reflects differences in the strength of learning across modality directions, with auditory-acquired regularities supporting stronger and more accessible representations than visually acquired ones. One possible explanation for this asymmetry lies in the nature of the auditory signal. Spoken input is inherently transient and often acoustically variable, which may place stronger demands on rapid encoding processes and lead to the formation of more robust, chunked representations when successfully learned. When these representations are subsequently tested in the visual modality—where input is stable and unambiguous—performance may be facilitated. This account is consistent with the "Now-or-Never" bottleneck hypothesis (Christiansen and Chater, 2016), which proposes that linguistic input must be rapidly recoded into compact representations due to the fleeting nature of speech. This pattern suggests that under more demanding retrieval conditions characteristic of the SICR task, access to learned representations remains sensitive to the sensory context in which they were encoded.

This dissociation highlights the importance of measurement in the study of cross-modal statistical learning. Forced-choice tasks may capture flexible familiarity-based access to learned structure, whereas recall-based measures provide a more direct index of the accessibility of representations of statistical chunks Frost et al. (2015); Siegelman and Frost (2015); Isbilen et al. (2020). More specifically, this pattern suggests that different tasks access different levels of representation. Recognition-based judgments may rely on relatively abstract representations that are readily accessible across modalities, whereas recall-based performance may depend on reconstructing sequences, making it more sensitive to modality-specific encoding.

Within the two-model framework, this dissociation can be interpreted more precisely. At the level of representation, the data are consistent with Model 2: learning appears to operate over shared, modality-general structures that support robust and symmetric transfer. At the level of retrieval and expression, however, aspects of Model 1 remain relevant, as performance continues to be shaped by modality-specific constraints and task demands. Thus, rather than selecting between the two models, the present findings suggest that they capture complementary stages of processing, with shared representations supporting cross-modal learning and modality-specific factors influencing how that learning is expressed.

More broadly, the interaction results indicate that cross-modal statistical learning

is not strictly fixed but dynamically shaped by learning history, sensory experience, and task demands. Rather than reflecting either a purely modality-general or strictly modality-specific mechanism, the present findings support a hybrid architecture in which shared, modality-independent representations are activated automatically during learning, while modality-specific constraints emerge at the level of retrieval and task demands.

Rethinking Modality Constraints in SL

The present findings can be interpreted in light of the two alternative models of cross-modal transfer outlined in the Introduction. Model 1 assumes that statistical learning occurs independently within each modality and that transfer reflects a secondary process operating over abstracted representations. In contrast, Model 2 proposes that, for familiar stimuli such as speech and text, uni-modal input automatically activates shared representations due to prior experience with orthography–phonology mappings, enabling immediate cross-modal access.

The absence of any difference between modality-match and modality-mismatch conditions, together with the lack of directional asymmetries, is more consistent with Model 2. Specifically, transfer appeared to be an intrinsic consequence of learning rather than a secondary process, suggesting that auditory and visual inputs accessed shared representations from the outset. At the same time, the modulation of performance by task demands in SICR indicates that modality-specific factors still influence how these representations are retrieved, supporting a hybrid account in which shared representations coexist with modality-dependent constraints at the level of expression.

Taken together, the present findings contribute to the ongoing debate regarding whether statistical learning relies on modality-general or modality- specific representations. Evidence from both tasks showed reliable cross-modal transfer and no directional asymmetry, indicating that learners can access representations that generalize across sensory modalities. At the same time, the interaction analyses revealed that the accessibility of these representations depended on learning history, sensory context, and task demands. In particular, modality-related retrieval costs emerged under the more demanding recall conditions of the SICR task and were modulated by exposure order and training modality. In contrast, recognition-based performance in the 2AFC task showed relatively flexible transfer across modalities.

One interpretation of this dissociation is that different tasks access different levels of the learned representation. Recognition-based judgments may rely on relatively abstract representations of sequential structure that can be accessed independently of

the modality in which they were learned. In contrast, recall-based performance may depend more strongly on modality-linked features of the encoded sequences, making retrieval more sensitive to the sensory context of learning. In this sense, cross-modal recognition may reflect access to higher-level regularities, whereas recall may require reconstructing sequences using representations that retain modality-specific traces.

Within the two-model framework, this pattern can be understood as reflecting different stages of processing. At the level of representation, the data support Model 2, as learning appears to operate over shared, modality-general structures. At the level of retrieval and expression, however, aspects of Model 1 remain relevant, as performance continues to be shaped by modality-specific constraints and task demands. Thus, rather than selecting between the two models, the present findings suggest that they capture complementary components of the same learning system.

This complementary perspective aligns with frameworks proposing that SL reflects domain-general computations implemented within modality-specific systems Frost et al. (2015). Rather than producing purely modality-general representations, statistical learning may generate hierarchical representations that include both modality-independent structural information (supporting transfer) and modality-linked encoding details (influencing retrieval).

Under this hybrid account, cross-modal generalization becomes possible because learners can access abstract representations of sequential structure, but task demands determine the extent to which modality-specific information influences performance. More broadly, the present findings help reconcile previously mixed results by suggesting that conclusions about modality generality depend not only on learning itself but also on the level of representation accessed and the cognitive demands of the task used to measure it.

4.5 Conclusion

The present study addressed four key questions about cross-modal SL. First, performance did not differ between modality-match and modality-mismatch conditions, indicating that learning generalized across modalities as effectively as within a single modality. Second, prior exposure influenced subsequent performance, with evidence of order-related changes, particularly in the decision-based task, suggesting that learning history shapes the expression of knowledge. Third, transfer was symmetric across sensory modalities, with comparable learning from auditory to visual and from visual to auditory contexts. Fourth, interaction analyses showed that the behavioral expression of transfer depended jointly on exposure order, training modality, and task demands, revealing greater modality sensitivity under effortful recall than in forced-choice judgments.

Taken together, these findings are most consistent with a model in which learning operates over shared, modality-general representations that enable cross-modal transfer as an intrinsic consequence of learning. At the same time, the task-dependent modulation of performance supports a model in which modality-specific constraints remain relevant at the level of retrieval and expression. This second model captures the representational basis of learning, whereas the first model explains how learning is accessed and expressed under different task demands.

More broadly, cross-modal statistical learning appears to reflect a flexible system in which abstract representations emerge from modality-bound experiences, while their behavioral expression remains sensitive to modality-specific processing factors, task demands, and prior learning history. Finally, the current findings from adults, who have a long history of over-learned associations between speech and text that facilitates robust cross-modal transfer in a linguistic context, suggests that for children who are still learning these speech-text associations the assessment of cross-modal transfer in a SL paradigm could serve as a sensitive indicator of emerging literacy.

Appendix

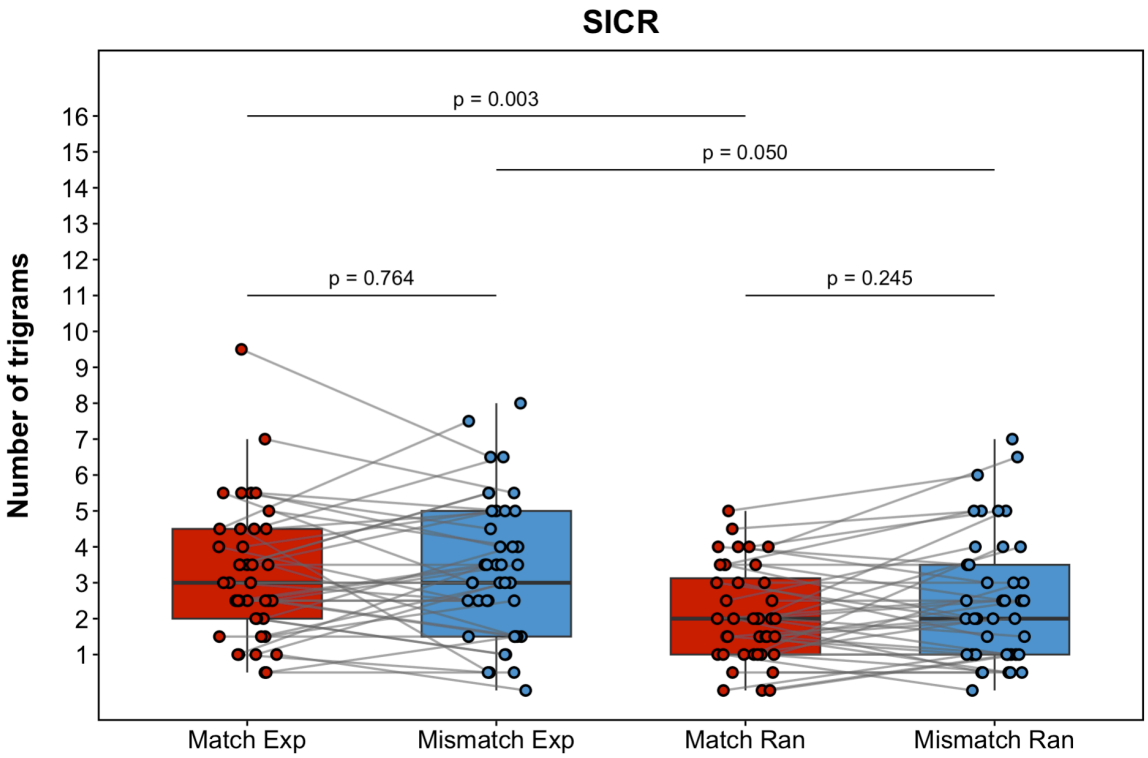


Figure A4.1: **Performance as a function of transfer in SICR task.** Number of correctly recalled trigrams for experimental stimuli (concatenations of two familiarized words) and random stimuli. Each dot represents an individual participant; boxplots summarize group-level performance, and grey lines connect paired observations within participants. Reported p-values refer to one-sample tests against chance (reported below each box) and paired comparisons between trials (reported above the connecting lines).

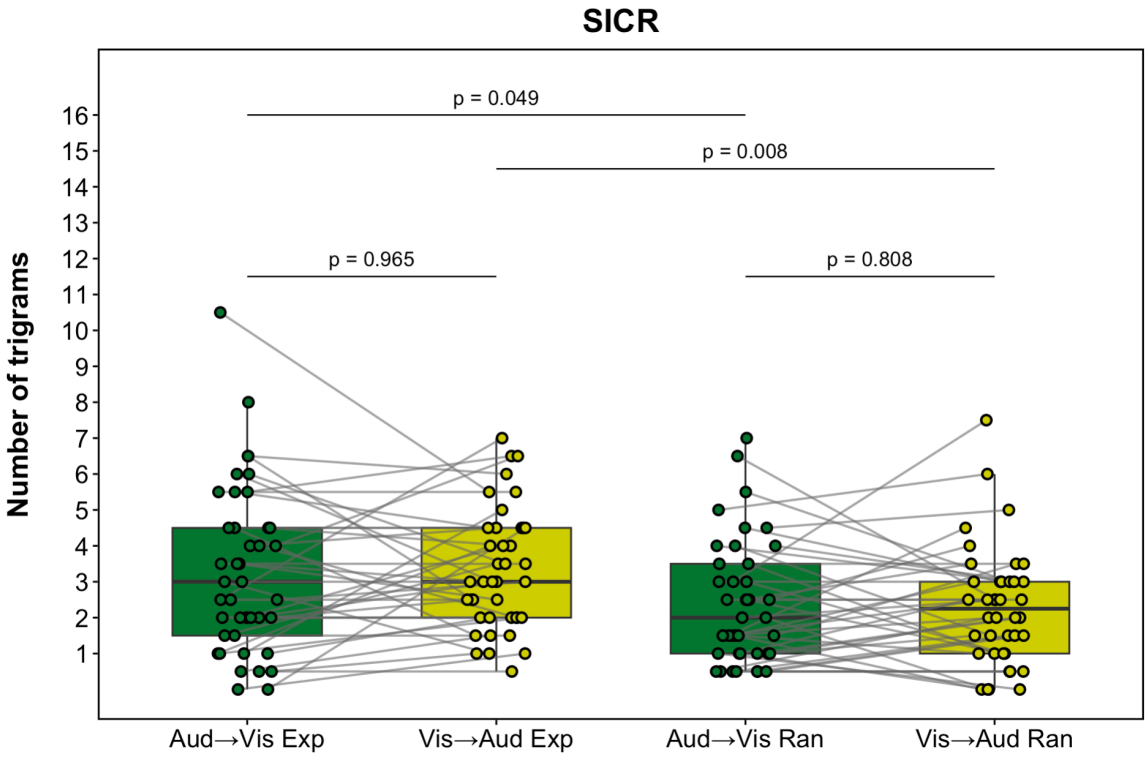


Figure A4.2: **Performance as a function of modality in SICR task.** Number of correctly recalled trigrams for experimental stimuli (concatenations of two familiarized words) and random stimuli, shown separately for Auditory-to-Visual and Visual-to-Auditory transfer directions. Each dot represents an individual participant; boxplots summarize group-level performance, and grey lines connect paired observations within participants. Reported p-values refer to one-sample tests against chance (reported below each box) and paired comparisons between conditions (reported above the connecting lines).

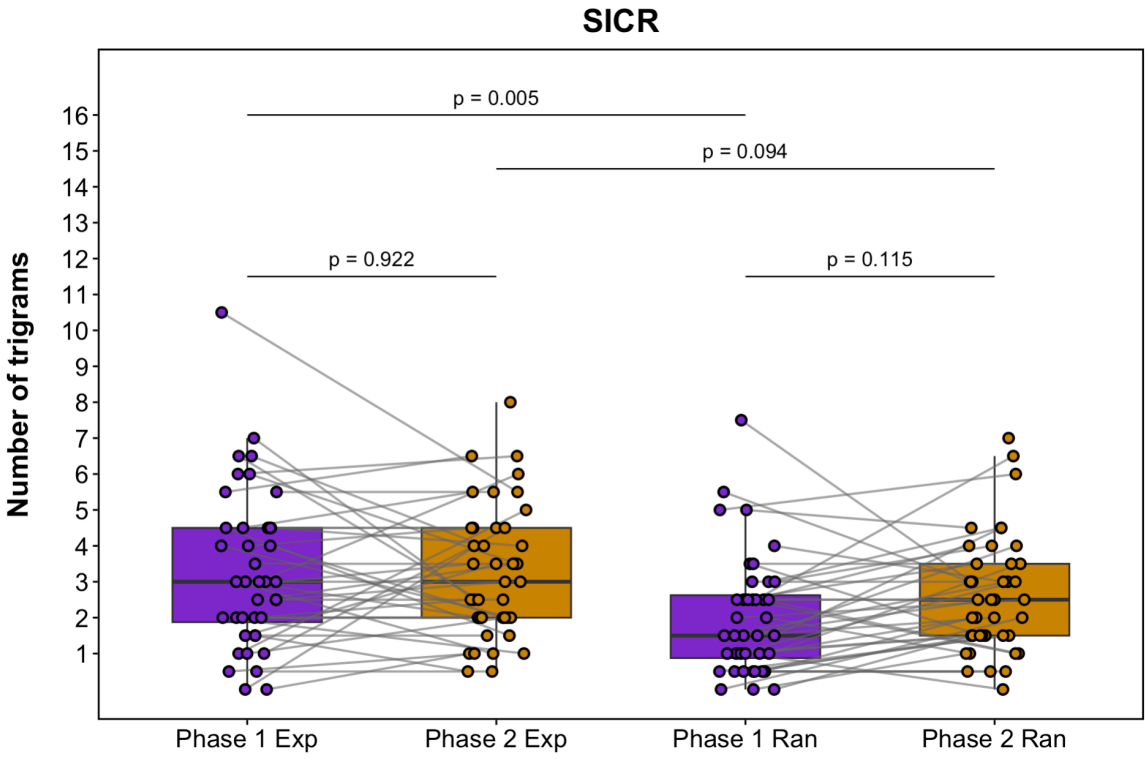


Figure A4.3: **Performance as a function of language order in SICR task.** Number of correctly recalled trigrams for experimental stimuli (concatenations of two familiarized words) and random stimuli, shown separately for Phase 1 and Phase 2. Each dot represents an individual participant; boxplots summarize group-level performance, and grey lines connect paired observations within participants. Reported p-values refer to one-sample tests against chance (reported below each box) and paired comparisons between languages (reported above the connecting lines).

Chapter 5

Listening to Read: the relationship between Auditory and Visual Statistical Learning and Literacy

Claudia Ruzza[1]
Erin S. Isbilen[2]
Richard N. Aslin[2]

[1]Department of Cognitive Neuroscience
Scuola Internazionale Superiore di Studi Avanzati (SISSA)
Trieste, Italy

[2]Yale Child Study Center
Yale University
New Haven, CT, USA

This research was in part supported by a National Institutes of Health K99 Pathway to Independence Award (1K99HD113838-01A1) awarded to Erin S. Isbilen.

Abstract

Statistical learning (SL) has been proposed as a key mechanism supporting language and reading development. The present study investigated auditory and visual SL, cross-modal transfer, and their relation to literacy skills in children aged 5–9 years. Following exposure to an artificial language, children completed auditory and visual judgment tasks together with standardized reading assessments (TOWRE-2). Results showed reliable auditory SL and weaker but directionally similar visual transfer effects, suggesting partial cross-modal generalization of learned regularities. Importantly, individual differences in auditory learning and visual transfer performance were positively associated with standardized reading ability, particularly age-normed word reading performance. In contrast, age showed relatively limited associations with learning performance within the sampled developmental range. Together, these findings suggest that SL and cross-modal transfer may contribute meaningfully to early literacy development while also reflecting modality-specific constraints on learning and generalization.

I would like to thank Melissa Vergara and Mary Shamon Ayala for their assistance with the SICR coding.

5.1 Introduction

Statistical learning (SL) refers to the ability to extract regularities from structured input through exposure, enabling learners to detect patterns such as transitional probabilities, co-occurrence statistics, and distributional properties of stimuli (Saffran et al., 1996a; Erickson and Thiessen, 2015). This process is typically considered implicit, as it occurs without explicit instruction or conscious awareness, and allows individuals to build internal representations of structured environments.

Statistical Learning in Childhood: Foundations and Mechanisms

Evidence for SL has been documented from early infancy across auditory and visual domains. Seminal work by Saffran et al. (1996a) demonstrated that 8-month-old infants can segment continuous speech streams into word-like units based on transitional probabilities between syllables. Subsequent studies extended these findings to visual sequences (Kirkham et al., 2002; Fiser and Aslin, 2002) and non-linguistic auditory stimuli such as tones (Saffran et al., 1999), suggesting that learners are sensitive to statistical regularities across multiple forms of sensory input.

Auditory statistical learning (ASL) has been proposed as a key mechanism for language acquisition because spoken language unfolds sequentially and contains rich probabilistic structure (Abreu et al., 2023). By tracking regularities in speech, infants and children can segment continuous input into recurring units that support vocabulary and lexical development (Conway and Pisoni, 2008; Evans et al., 2009). Children continue to demonstrate robust learning beyond infancy, including sensitivity to sequential and non-adjacent dependencies (Gómez, 2002; Thiessen et al., 2013).

Visual statistical learning (VSL) likewise supports the extraction of regularities from structured visual input and has been implicated in the acquisition of orthographic knowledge relevant for reading development. Children reliably demonstrate above-chance learning across age groups, with evidence for developmental improvements in performance (Arciuli and Simpson, 2011; Ren and Wang, 2024). Learning also appears to occur online during exposure, as children show sensitivity to statistical structure during familiarization itself (Sáringer et al., 2024).

An important question concerns how different forms of statistical learning develop across childhood and whether developmental trajectories differ across modalities and representational domains. Reading acquisition, for example, requires sensitivity to auditory regularities, visual orthographic regularities, and the mappings between phonological and orthographic representations. Statistical learning may therefore

contribute to literacy development through multiple interacting mechanisms operating across sensory modalities, each with their own distinct perceptual affordances, and possibly, distinct developmental trajectories. Some accounts propose that statistical learning reflects an early-emerging capacity that remains relatively stable throughout development, particularly in the auditory domain because of its proposed role in spoken language acquisition (Saffran et al., 1997). Consistent with this view, Saffran et al. (1997) found comparable auditory statistical learning performance in 6-year-old children and adults. However, other findings suggest that developmental improvements may emerge depending on the modality, representational domain, and task demands.

While auditory-linguistic SL appears relatively stable across the lifespan, evidence from the visual domain has been mixed. Some studies report little evidence for developmental change across childhood. For example, Bertels et al. (2015) found that children aged 9–12 years performed similarly to adults on a visual SL task, while Schlichting et al. (2017) reported that adults outperformed both children and adolescents, but observed no significant differences between the younger age groups themselves. These findings have been interpreted as suggesting relative age-invariance in visual SL during childhood. However, these studies relied on relatively small samples and examined limited age ranges, potentially reducing sensitivity to gradual developmental changes (Arciuli and Simpson, 2012; Bertels et al., 2015). In contrast, a larger-scale developmental study by Arciuli and Simpson (2011) examined visual SL in 183 children aged 5–12 years and found clear age-related improvements, with older children demonstrating significantly stronger learning abilities.

Even less work has examined developmental trajectories in auditory statistical learning despite its theorized importance for language acquisition. Although some evidence suggests stable auditory statistical learning across childhood (Saffran et al., 1997), other auditory sequence-learning paradigms, such as artificial grammar learning, have shown age-related improvements (Saffran, 2001). Importantly, artificial grammar learning tasks often involve more complex sequential structures and place greater demands on working memory, attention, and explicit processing than simpler syllable-triplet learning paradigms. Developmental improvements observed in these tasks may therefore reflect not only changes in statistical learning itself, but also broader maturation of cognitive systems supporting the processing of more complex sequential information.

Addressing these inconsistencies more directly, Raviv and Arnon (2018) compared auditory and visual statistical learning across children aged 5–12 years using closely matched paradigms. Their findings revealed modality- and domain-specific developmental trajectories: auditory statistical learning for linguistic stimuli re-

mained comparatively stable across the tested age range, whereas visual statistical learning for non-linguistic stimuli improved significantly with age. Together, these findings suggest that developmental changes in statistical learning may depend on modality-specific perceptual, representational, and cognitive constraints rather than reflecting a fully unified learning mechanism.

Evidence comparing learning across modalities further suggests that SL may involve both modality-general and modality-specific mechanisms. Importantly, relatively few developmental studies have directly compared the learning of phonological and orthographic regularities within the same paradigm, as auditory and visual statistical learning tasks often rely on stimuli from different representational domains (e.g., linguistic vs. non-linguistic materials). This distinction complicates interpretation of modality effects because observed differences may partly reflect domain-specific cognitive and representational demands in addition to sensory modality itself. Studies examining auditory, visual, and tactile learning in children report limited correlations across modalities despite above-chance learning within individual domains (Pavlidou and Bogaerts, 2019). Similarly, distributional learning studies indicate stronger associations within than across sensory modalities, with auditory tasks favoring temporal regularities and visual tasks favoring spatial regularities (Zhou et al., 2024). Importantly, statistical learning paradigms vary substantially in complexity, reliability, and response demands, which can strongly influence observed developmental effects and individual-differences findings (Isbilen and Christiansen, 2022; Siegelman and Frost, 2017). Together, these findings highlight the importance of considering modality, task structure, and measurement characteristics when interpreting developmental differences in SL.

Statistical Learning and Literacy Development: Theoretical Links and Empirical Evidence

Reading is a complex cognitive skill that requires mapping visual symbols onto phonological and semantic representations. Successful reading acquisition depends on multiple interacting abilities, including phonological awareness, decoding, orthographic processing, and reading fluency (Pikulski and Chard, 2005). From this perspective, reading can be understood as a form of skill learning that relies heavily on detecting and internalizing regularities present in written and spoken language (Chater and Christiansen, 2018).

SL has therefore been proposed as a key mechanism supporting literacy development. Written language contains rich statistical structure, including regularities in letter co-occurrences, grapheme–phoneme correspondences, and morphological pat-

terns that are often acquired implicitly through exposure (Pacton et al., 2001; Frost et al., 2009a). Connectionist models of reading similarly propose that successful word recognition and decoding emerge through sensitivity to the probabilistic structure of linguistic input (Seidenberg and McClelland, 1989; Harm and Seidenberg, 2004). According to this view, children gradually acquire orthographic and phonological knowledge by extracting recurring statistical regularities from written and spoken language experience.

Empirical findings provide substantial support for a relationship between SL and reading-related abilities. For example, Arciuli and Simpson (2012) reported that SL predicts reading performance in both children and adults, while Spencer et al. (2015) found that SL explains unique variance in early literacy skills, including vocabulary and phonological processing. More recently, Ren and Wang (2024) demonstrated that multiple forms of SL contribute to reading performance across languages in bilingual children learning English as a second language. Together, these findings suggest that SL may support literacy development both within and across linguistic systems.

Evidence further suggests that ASL and VSL learning may contribute differently to reading acquisition. ASL appears particularly relevant for phonological processing and spoken language representations (Qi et al., 2019; Clayton et al., 2020a), whereas VSL may more directly support orthographic pattern learning (Pavlidou and Bogaerts, 2019). Supporting this distinction, Pavlidou and Bogaerts (2019) found that VSL performance was associated with phonological awareness in children, despite weak correlations across modalities. Similarly, Ktori et al. (2026) showed that children exposed to recurring visual chunks within printed strings were more likely to later endorse novel strings containing those chunks, suggesting implicit learning of morpho-orthographic regularities through visual exposure.

Research on developmental disorders further supports a connection between SL and literacy. Children and adults with dyslexia have been reported to show impairments in SL, particularly for verbal and orthographic material (Mateus-Moreno et al., 2025). However, evidence linking SL and reading remains mixed, with some studies reporting weak or inconsistent associations (Schmalz et al., 2019; West et al., 2018). TThese inconsistencies may reflect differences in task demands, modality, developmental stage, or the specific reading skills being assessed (for further discussion, see Bogaerts et al., 2022b; Isbilen and Christiansen, 2022). With reading requiring the assimilation of both auditory and visual regularities, these findings suggest that different forms of SL may contribute differently to literacy development.

Together, these findings suggest that SL may provide an important foundation for literacy development by supporting the acquisition of phonological and orthographic

regularities. However, the extent to which auditory SL and auditory-to-visual transfer contribute to reading development remains incompletely understood. In particular, relatively little work has directly examined whether statistical regularities learned through auditory experience can generalize to orthographic representations during childhood, despite the likely importance of such cross-modal transfer processes for linking phonological and orthographic information during early literacy acquisition, where the acquisition of phonological wordforms precedes the acquisition of their orthographic counterparts.

Cross-Modal Transfer: Evidence and Mechanisms

Statistical learning has been demonstrated across different domains, raising important questions regarding whether learning acquired in one modality transfers to another or instead depends on modality-specific mechanisms (Marcus et al., 1999; Conway and Christiansen, 2005; Frost et al., 2015). Understanding cross-modal transfer may be particularly important for literacy development because learning to read requires children to map auditory phonological representations onto visual orthographic symbols. Successful reading acquisition therefore depends not only on extracting regularities within auditory and visual modalities independently, but also on integrating information across modalities (Dehaene et al., 2015; Garcia et al., 2019). Some researchers have argued that the predictive power of phonological awareness and rapid automatized naming for reading may partly derive from children's ability to implicitly learn phonological regularities and map them onto written forms (Qi et al., 2019). Investigating auditory-to-visual transfer may therefore provide important insight into how children connect spoken and written language representations during the early stages of literacy acquisition.

In the auditory domain, children reliably extract regularities from speech and nonspeech input, supporting the development of phonological and lexical representations (Saffran et al., 1996a; Evans et al., 2009). Visual statistical learning likewise enables sensitivity to structured visual sequences and has been linked to orthographic processing and written language acquisition (Kirkham et al., 2002; Arciuli and Simpson, 2012). By tracking recurring visual patterns, children may gradually acquire sensitivity to orthographic regularities such as letter combinations, grapheme patterns, and sublexical structure that support fluent word recognition and reading development. Some evidence further suggests that statistical learning in both auditory and visual modalities contributes to literacy skills. For example, Qi et al. (2019) reported that auditory statistical learning predicted reading proficiency above and beyond visual statistical learning in both children and adults, potentially because written text

automatically activates phonological representations in skilled readers. Other work suggests that visual statistical learning may show particularly strong associations with orthographic and phonological awareness skills (Pavlidou and Bogaerts, 2019), possibly because sensitivity to recurring visual regularities supports the acquisition of written language structure.

A growing body of behavioral evidence further suggests that children can generalize statistical regularities across modalities. Studies examining orthographic learning have shown that children form orthographic expectations following exposure to novel spoken words (Beyersmann et al., 2023), consistent with the idea that phonological representations can support emerging orthographic learning. Similarly, visual–phonological associative learning paradigms indicate that successful reading development depends on efficient integration between auditory and visual language representations (Garcia et al., 2019). More recently, Ren and Wang (2024) argued that cross-modal transfer may play an important role in literacy acquisition but noted that relatively few studies have directly examined how statistical regularities learned in one modality generalize to another during childhood.

Neural evidence also supports the possibility of cross-modal transfer. Neuroimaging studies have shown that written text can evoke automatic phonological representations in skilled readers and that speech and orthographic processing recruit partially overlapping neural systems (Dehaene et al., 2015). Recent EEG frequency-tagging paradigms further demonstrate that neural entrainment can track the online extraction of statistically defined word structure in both auditory and visual streams (Choi et al., 2020; Moreau et al., 2022). These paradigms provide implicit measures of learning that may be more sensitive than traditional behavioral tasks, particularly in children (Batterink et al., 2019). Importantly, emerging work proposes that neural entrainment may also capture cross-modal transfer itself, indexing the extent to which previously learned phonological or orthographic regularities are automatically reactivated when encountered in a different sensory modality.

At the same time, evidence for robust cross-modal transfer remains mixed. Several studies report weak or inconsistent transfer effects across modalities. For example, Pavlidou and Bogaerts (2019) found no significant correlations between auditory, visual, and tactile SL performance in children despite above-chance learning within individual modalities. Similarly, Mitchel and Weiss (2011b) reported that transfer depended heavily on task structure and was not consistently observed across conditions. Evidence from temporal adaptation paradigms additionally suggests that transfer may be asymmetrical across modalities, with stronger visual-to-auditory than auditory-to-visual transfer effects (Batikh et al., 2025). Together, these findings suggest that transfer may not emerge automatically even when SL occurs successfully

within individual modalities.

One possible explanation for these mixed findings is that SL may rely on both modality-general and modality-specific mechanisms. Early demonstrations of SL across auditory, visual, and motor domains supported proposals of a unified domain-general learning system (Saffran et al., 1996a; Conway and Christiansen, 2005; Kirkham et al., 2002). However, subsequent work showing weak cross-modal correlations and limited transfer challenged this interpretation (Misyak and Christiansen, 2012; Siegelman and Frost, 2015). Supporting a more nuanced account, Frost et al. (2015) proposed that SL reflects shared computational principles implemented within modality-specific perceptual systems, such that learning outcomes remain shaped by sensory and task-related constraints.

Additional evidence for modality-specificity comes from neurocognitive and clinical studies. Neuroimaging findings suggest that ASL and VSL rely on partially distinct neural systems (Prem et al., 2025), while adults with dyslexia have been reported to show impaired ASL despite relatively preserved VSL (Bulut and Şenol, 2026). At the same time, evidence for modality-general processing also exists, with studies reporting overlapping neural networks involved in SL and abstraction across modalities (Sáringer et al., 2024; Zhou et al., 2024). Together, these findings suggest that SL likely reflects an interaction between shared learning computations and modality-specific representational systems.

Methodological factors may also contribute substantially to inconsistencies in the literature. Cross-modal transfer studies often vary considerably in task structure, stimulus type, exposure duration, and response demands, making findings difficult to compare directly. In developmental populations, these issues may be compounded by age-related variability in attention, working memory, and meta-cognitive abilities. Concerns have additionally been raised regarding the psychometric reliability of commonly used SL tasks, particularly forced-choice paradigms that may underestimate implicit learning and show poor sensitivity to individual differences in children (Arnon, 2020; Isbilen and Christiansen, 2022; Batterink et al., 2019). Such methodological limitations may partially explain why some studies observe strong transfer effects whereas others report weak or absent associations. Recent developmental work has therefore increasingly employed child-friendly rating paradigms that provide more graded measures of learning and may be better suited for capturing individual variability in young children (Schuler et al., 2026).

Overall, the current literature suggests that cross-modal transfer likely plays an important role in literacy development, but the mechanisms underlying this process remain incompletely understood. In particular, relatively few developmental studies have directly examined whether children can generalize statistical regularities learned

through auditory experience to orthographic representations, despite the potential importance of such transfer processes for early reading acquisition.

Research Goals and Questions

Despite substantial progress in understanding SL during development, several important questions remain unresolved. First, relatively few studies have examined SL during the critical period of early reading acquisition, particularly between 5 and 7 years of age, limiting understanding of developmental trajectories during the emergence of literacy skills (Raviv and Arnon, 2018; Arciuli, 2017). Second, auditory-to-visual transfer has received comparatively little empirical attention in children, despite its potential importance for linking phonological and orthographic representations during reading acquisition. Third, although SL has been associated with reading-related abilities, relatively few studies have simultaneously examined auditory learning, cross-modal transfer, and standardized literacy measures within the same developmental sample (Arciuli and Simpson, 2012; Schmalz et al., 2019). In addition, concerns regarding the reliability and ecological validity of SL paradigms highlight the importance of using developmentally appropriate tasks and complementary literacy measures when investigating individual differences in learning abilities (Siegelman et al., 2017a; Isbilen and Christiansen, 2022).

The present study investigated the relationship between auditory statistical learning, auditory-to-visual transfer, and literacy skills in children aged 5 to 7 years, a developmental period during which foundational reading abilities rapidly emerge and formal reading instruction typically begins in the United States. Children were exposed to an artificial language through auditory familiarization and subsequently completed auditory and visual rating tasks assessing their ability to discriminate between familiar and novel patterns. Performance in the visual modality was designed to assess auditory-to-visual transfer of previously learned regularities to orthographic representations. In addition, independent literacy measures, including letter identification and standardized assessments of word and nonword reading (TOWRE-2), were collected.

This design addressed three main research questions: (1) Do children exhibit reliable auditory SL? (2) Is there evidence of auditory-to-visual transfer following auditory familiarization? and (3) How are auditory learning and cross-modal transfer related to literacy skills?

We hypothesized that children with stronger literacy skills would show better performance on statistical learning tasks, particularly on tasks involving auditory regularities relevant for phonological processing. We further predicted that children

showing stronger auditory learning and cross-modal transfer would demonstrate better literacy outcomes, consistent with the idea that reading development relies in part on the ability to extract and integrate regularities across auditory and visual domains. More broadly, if auditory statistical learning and cross-modal transfer are associated with literacy performance during early childhood, this would support theoretical accounts proposing that reading acquisition depends in part on domain-general learning mechanisms that enable children to detect and generalize structured regularities across modalities. Such findings would also provide further insight into how phonological and orthographic representations become integrated during the emergence of reading.

5.2 Materials and Methods

The present study employed an artificial language learning paradigm to investigate auditory SL, auditory-to-visual transfer, and their relation to literacy skills during early childhood. Children were first familiarized with an artificial auditory language and subsequently completed auditory and visual rating tasks assessing their ability to discriminate trained stimuli from novel distractors. Performance in the visual modality was designed to assess cross-modal transfer of previously learned auditory regularities to orthographic representations. In addition, children completed measures of letter identification and standardized reading ability (TOWRE-2).

5.2.1 Participants

Fourty-four children participated in the study. Ten participants were excluded prior to analysis for the following reasons: five due to corrupted or incomplete experimental video recordings that prevented reliable scoring of the TOWRE-2 task; one due to poor performance on the letter identification task (4/12 correct letters); one because all responses in both auditory and visual judgment tasks were uniformly rated as 5, preventing meaningful assessment of discrimination performance; and three because they failed the practice task, indicating that they did not appropriately use the response scale (e.g., assigning low ratings to animal stimuli or high ratings to object stimuli when asked whether the image depicted an animal).

The final sample therefore consisted of 34 children (20 female, 14 male) between 5.6 and 9.9 years of age (M = 7.91 years, SD = 1.23). All children were native

English speakers with normal or corrected-to-normal vision and no reported history of neurological, developmental, or language disorders according to parental report.

Parents provided informed consent prior to participation, and children gave verbal assent before beginning the experiment. The study was approved by the local institutional ethics committee and conducted in accordance with the Declaration of Helsinki.

5.2.2 Stimuli

The experiment employed artificial linguistic stimuli composed of consonant–vowel (CV) syllables, adapted from classic SL paradigms (Saffran et al., 1996a; Kidd et al., 2020): *bu, di, du, ka, ki, lo, ma, mo, pa, po, ri, ta.* These syllables were concatenated to form trisyllabic words and were presented auditorily as continuous speech streams.

The artificial language consisted of four trisyllabic words: *takapo, kibudu, modipa, lomari.* Syllables were arranged such that no syllable systematically occupied the same serial position across words. Transitional probabilities between adjacent syllables were 1.0 within words and 0.33 between words, thereby minimizing segmentation cues other than transitional probabilities. Foils for the rating tasks were constructed using the same syllable inventory while respecting the phonotactic constraints of the artificial language. Specifically, two foils contained repeated vowels and two did not (*dilota, kaburi, makidu, popamo*). These items served as distractors in the test phases. Visual stimuli in the rating task consisted of lowercase orthographic strings corresponding to trained words and distractors. Stimuli were presented centrally on a light grey background using a large sans-serif font in black lowercase letters to maximize readability and minimize visual complexity for young children. All strings were displayed in the same font, size, and formatting across conditions.

In the practice phase, visual stimuli consisted of images depicting animals, objects, and ambiguous animal–object hybrids. The hybrid images were generated using artificial intelligence tools to create perceptually uncertain stimuli. Practice stimuli can be found in the Supplementary Materials, Table S9.

In the letter identification, only letters used in the experimental stimuli were included (i.e., *a, b, d, i, k, l, m, o, p, r, t, u*). Stimuli used in the TOWRE-2 word and nonword task can be found in the Supplementary Materials section, respectively Table S2 and Table S3; TOWRE-2 task stimuli are taken from Torgeson et al. (1999) and can be found in Table S4 (word stimuli) and Table S5 (nonword stimuli).

5.2.3 Experimental Paradigm and Procedure

After collecting demographic information (age, gender, native language, and handedness), caregivers were instructed to remain present during the session without interfering with task performance. This procedure was implemented based on pilot testing, which indicated that some children tended to disengage from the TOWRE-2 task despite having adequate reading abilities. Caregiver presence was therefore intended to support sustained engagement while minimizing external influence on responses.

Children completed a multi-phase experimental protocol designed to assess auditory SL (ASL), auditory-to-visual transfer, and their relation to reading abilities. In the familiarization phase, children were exposed to an artificial "penguin language" composed of four trisyllabic words. The familiarization consisted of three 2-minute blocks, during which each word was presented 74 times in total. To enhance engagement, the auditory speech stream was overlaid onto short clips from the claymation cartoon *Pingu*. *Pingu* was selected because it is highly engaging for young children while containing no intelligible spoken language, thereby minimizing interference with the artificial language stimuli. Children were told they were listening to the secret language of penguins, and static images of Pingu, the titular character, accompanied task instructions to further support engagement throughout the experiment.

Following familiarization, children completed a brief practice phase of the rating task (6 trials) to ensure they understood how to use a 5-point Likert scale adapted for children. The rating procedure was adapted from child-friendly judgment paradigms previously shown to be successful with children aged 6–8 years (Schuler et al., 2026). Specifically, the numerical scale was replaced with a colored gradient ranging from red (accompanied by a thumbs-down icon) through yellow to green (accompanied by a thumbs-up icon), making the response options more intuitive for young children. Rating-based measures have additionally been argued to provide sensitive and reliable indices of SL relative to some commonly used forced-choice paradigms (Isbilen and Christiansen, 2022). During practice, children were asked to judge whether an image depicted an animal. The stimuli included two animal images, expected to elicit high ratings, two object images, expected to elicit low ratings, and two ambiguous hybrid images designed to encourage use of intermediate scale points.

Children then completed an auditory rating task in which they listened to either trained words or foils (random combinations of syllables) and rated whether each item belonged to the penguin language using the same response scale.

Next, children completed a visual rating task with written letter strings corre-

sponding to trained words and foils, again using the same response scale. Importantly, children did not receive visual familiarization prior to this phase. Performance in the visual modality was therefore interpreted as reflecting auditory-to-visual transfer of previously learned regularities to orthographic representations rather than independent visual SL. To account for basic letter knowledge, children completed a letter identification task. On each trial, children heard an auditory instruction (e.g., "Click on *b*") while five letters were simultaneously presented on screen. Children were instructed to select the target letter from the available options using the cursor.

Finally, reading proficiency was assessed using the TOWRE-2 (Test of Word Reading Efficiency), including both the Sight Word Reading Efficiency (word reading) and Phonemic Decoding Efficiency (nonword reading) subtests. Children first completed a short practice with five words to ensure task comprehension. They were then presented with a list of words ordered by increasing difficulty and were given 45 seconds to read aloud as many items as possible, skipping those they could not read. The same procedure was subsequently applied to nonwords: children first completed five practice trials, followed by a timed test in which they read as many nonwords as possible within 45 seconds. This task provided measures of reading fluency and decoding ability. The experiment was conducted online using the Children Helping Science platform (`https://childrenhelpingscience.com` by Scott and Schulz, 2017), a web-based tool for designing and administering behavioral experiments with children, allowing controlled stimulus presentation and response collection via standard web browsers.

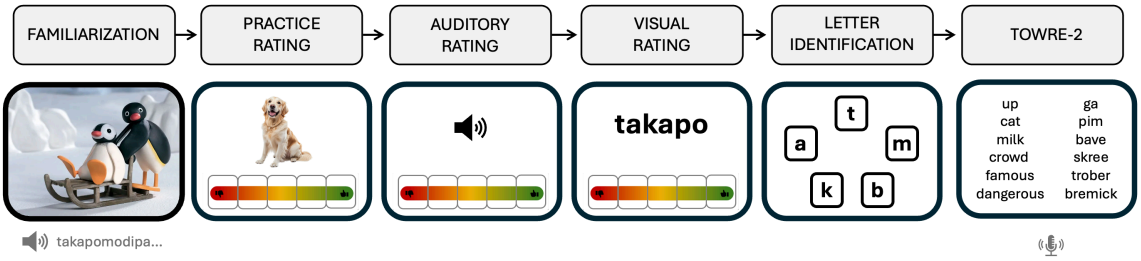


Figure 5.1: **Overview of the experimental design and procedure.** Children first completed a familiarization phase in which they were exposed to an artificial "penguin language" embedded in a cartoon narrative. This was followed by a rating practice phase to learn the use of a 5-point Likert scale. In the auditory rating task, children listened to trained words and foils and rated whether they belonged to the artificial language. The same procedure was then administered in the visual modality using written letter strings to assess auditory-to-visual transfer of previously learned regularities. Subsequently, children completed a letter identification task to assess knowledge of the stimulus letters. Finally, reading proficiency was measured using the TOWRE-2, which evaluates word and nonword reading efficiency under time constraints.

5.2.4 Behavioral Data Analysis

Behavioral analyses were conducted in R (version 4.3.2; R Core Team, 2023) within the RStudio integrated development environment (Posit team, 2023). Statistical analyses included Friedman tests, Wilcoxon signed-rank tests, repeated-measures analyses of variance (ANOVAs), one-sample and paired-samples t-tests, Pearson correlations, and partial correlations, depending on the structure of the data and the specific research question. Reliability of the rating tasks was assessed using Cronbach's alpha and split-half reliability analyses.

For the rating task, analyses focused on children's ability to discriminate between trained stimuli from the artificial "penguin language" and novel distractor items in both auditory and visual modalities. Ratings were first analyzed as raw Likert-scale responses relative to the midpoint of the scale. To account for individual differences in response tendencies, ratings were additionally z-score normalized separately within each participant and modality. Finally, difference scores were computed by subtracting distractor ratings from trained-stimulus ratings within each modality, providing a direct bias-corrected measure of auditory learning and auditory-to-visual transfer.

Reading ability was assessed using the Test of Word Reading Efficiency–Second Edition (TOWRE-2), including separate timed word and nonword reading subtests. Both raw accuracy scores and age-standardized scores were analyzed. Exploratory correlation analyses further examined relations between age, reading ability, auditory learning, and visual transfer performance indexed by rating-task difference scores.

5.3 Results

Behavioral results are organized into five sections. We first report performance on the practice and letter identification tasks to confirm that children understood the response procedure and were familiar with the visual stimuli used in the experiment. We then present analyses of the auditory and visual judgment tasks assessing SL and cross-modal transfer, followed by standardized reading performance on the TOWRE-2. Finally, exploratory analyses examining relations between age, reading ability, and SL performance are reported.

5.3.1 Practice Task

Children successfully used the 5-point response scale during the practice task (see the Appendix section, respectively Fig.A5.1).

Ratings differed significantly across stimulus categories, Friedman's $\chi^2(2) = 63.80$ ($p < .001$). Post-hoc Wilcoxon signed-rank tests revealed that animal stimuli received significantly higher ratings than both hybrid stimuli ($V = 595$, $p < .001$) and object stimuli ($V = 595$, $p < .001$). In contrast, ratings for hybrid and object stimuli did not significantly differ ($V = 21.5$, $p = .233$).

Descriptively, children assigned high ratings to animal stimuli ($M = 4.88$, $SD = 0.61$), whereas both hybrid ($M = 1.32$, $SD = 0.74$) and object stimuli ($M = 1.15$, $SD = 0.63$) received low ratings. Overall, these results indicate that children understood and appropriately used the response scale during the task.

5.3.2 Letter Identification

Children demonstrated near-ceiling performance in the letter identification task. Participants completed 12 trials each and correctly identified a mean of 11.9 letters out of 12 ($SD = 0.24$, range = 11–12). Thirty-two out of 34 children achieved perfect performance, while only two children made a single error. These results indicate that the letters used in the visual rating task were highly familiar to participants.

5.3.3 Rating Task

The auditory and visual rating tasks were designed to assess whether children could discriminate between stimuli from the artificial "penguin language" and novel distractor items across modalities. Importantly, children received auditory familiarization only; no visual familiarization phase was administered prior to the visual task. Performance in the visual modality therefore reflects cross-modal transfer of previously learned auditory regularities rather than independent visual SL. Specifically, we examined whether children would assign higher ratings to trained stimuli relative to distractors, reflecting successful auditory statistical learning (ASL) and potential auditory-to-visual transfer to orthographic forms.

Before examining task performance, we first assessed the reliability of the rating measures. Internal consistency was evaluated separately for the auditory and visual tasks using Cronbach's alpha and split-half reliability analyses. The auditory rating task showed moderate reliability (Cronbach's $\alpha = .62$, 95% CI [.39, .78]), whereas the visual rating task showed high reliability (Cronbach's $\alpha = .86$, 95% CI [.77, .92]). Split-half analyses yielded comparable results, with average split-half reliabilities of .62 for the auditory task and .86 for the visual task.

Additional split-half indices further supported these patterns. For the auditory task, the maximum split-half reliability (λ_4) was .84, Guttman's λ_6 was .76, Guttman's λ_2 was .66, and the minimum split-half reliability (β) was .38. For the visual task, the maximum split-half reliability (λ_4) was .91, Guttman's λ_6 was .87, Guttman's λ_2 was .86, and the minimum split-half reliability (β) was .77. Average inter-item correlations were .17 for the auditory task and .43 for the visual task. Together, these findings indicate that the rating tasks provided reliable measures of children's responses, particularly in the visual modality.

Before computing overall learning measures, we examined whether performance differed across the two testing blocks. Mean ratings were analyzed using a repeated-measures ANOVA with modality (auditory, visual), item type (stimulus, distractor), and block (Block 1, Block 2) as within-subject factors.

The analysis revealed significant main effects of modality ($F(1, 33) = 7.24$, $p = .011$, $\eta^2_G = .065$), and block ($F(1, 33) = 8.85$, $p = .005$, $\eta^2_G = .015$). The main effect of item type was not significant ($F(1, 33) = 3.59$, $p = .067$, $\eta^2_G = .012$). No significant interactions were observed. Specifically, the modality × item type interaction was not significant ($F(1, 33) = 1.86$, $p = .182$, $\eta^2_G = .003$); the modality × block interaction was not significant ($F(1, 33) = 2.94$, $p = .096$, $\eta^2_G = .006$); the item type × block interaction was not significant ($F(1, 33) = 0.11$, $p = .741$, $\eta^2_G < .001$); and the three-way interaction between modality, item type, and block was not significant ($F(1, 33)$

= 1.91, $p = .177$, $\eta^2_G = .004$).

Because no evidence suggested differential learning effects across testing blocks, ratings were collapsed across Block 1 and Block 2 for all subsequent analyses.

Given that the rating task employed a 5-point Likert scale, raw ratings were first evaluated relative to the scale midpoint (3), with higher values indicating stronger endorsement that an item belonged to the penguin language (see the Appendix section, respectively Fig.A5.2).

Descriptively, auditory trained stimuli received the highest ratings ($M = 3.94$, $SD = 0.73$), followed by auditory distractors ($M = 3.59$, $SD = 0.80$). In the visual modality, trained stimuli received mean ratings of $M = 3.30$ ($SD = 1.01$), whereas distractors received mean ratings of $M = 3.19$ ($SD = 0.93$).

One-sample t-tests against the scale midpoint revealed that auditory trained stimuli were rated significantly above the midpoint ($t(33) = 7.55$, $p < .001$), indicating that children reliably endorsed familiar auditory words as belonging to the artificial language. Auditory distractors were likewise rated above the midpoint ($t(33) = 4.33$, $p < .001$), suggesting an overall tendency toward positive responding in the auditory modality. In the visual modality, trained stimuli were also rated significantly above the midpoint ($t(33) = 1.70$, $p = .049$), although the effect was substantially weaker. Visual distractor ratings did not significantly differ from the midpoint ($t(33) = 1.22$, $p = .884$).

Direct comparisons between stimuli and distractors further supported auditory learning. In the auditory modality, trained stimuli received significantly higher ratings than distractors ($t(33) = 2.20$, $p = .035$), mean difference = 0.35 (95% CI [0.03, 0.67]). In contrast, the difference between visual stimuli and distractors was not significant ($t(33) = 0.85$, $p = .403$, mean difference = 0.10, 95% CI [-0.14, 0.35]). Comparisons across modalities further showed that auditory stimuli received significantly higher ratings than visual stimuli ($t(33) = 3.03$, $p = .005$, mean difference = 0.64, 95% CI [0.21, 1.08]), whereas auditory and visual distractor ratings did not significantly differ ($t(33) = 1.91$, $p = .065$, mean difference = 0.40, 95% CI [-0.03, 0.82]).

A repeated-measures ANOVA with modality (auditory, visual) and item type (stimulus, distractor) revealed significant main effects of modality ($F(1, 33) = 7.36$, $p = .011$, $\eta^2_G = .084$) and item type ($F(1, 33) = 4.15$, $p = .050$, $\eta^2_G = .017$), indicating overall higher ratings in the auditory modality and for trained stimuli relative to distractors. The modality × item type interaction was not significant ($F(1, 33) = 2.01$, $p = .165$, $\eta^2_G = .005$).

Inspection of individual response patterns revealed substantial variability in scale use across children, with some participants consistently using higher portions of the

response scale than others. To control for these response biases, ratings were z-score normalized separately within each participant and modality (see Fig. 5.2, Left panel). Three participants were excluded from the auditory z-score analyses because they showed no variance in auditory ratings, resulting in undefined z-scores.

Descriptively, auditory stimuli received positive mean z-scored ratings ($M = 0.13$, $SD = 0.30$), whereas auditory distractors received negative mean z-scored ratings (M = -0.13, $SD = 0.29$). A similar but weaker pattern was observed in the visual modality, with trained stimuli receiving slightly positive ratings ($M = 0.03$, $SD = 0.26$) and distractors receiving slightly negative ratings (M = -0.03, $SD = 0.26$).

One-sample t-tests against zero revealed that auditory stimuli were rated significantly above participants' modality-specific average response level ($t(30) = 2.50$, $p = .009$), whereas auditory distractors were rated significantly below average ($t(30)$ = -2.46, $p = .010$). In the visual modality, neither trained stimuli nor distractors significantly differed from zero (stimuli: $t(33) = 0.62$, $p = .270$; distractors: $t(33)$ = -0.65, $p = .260$).

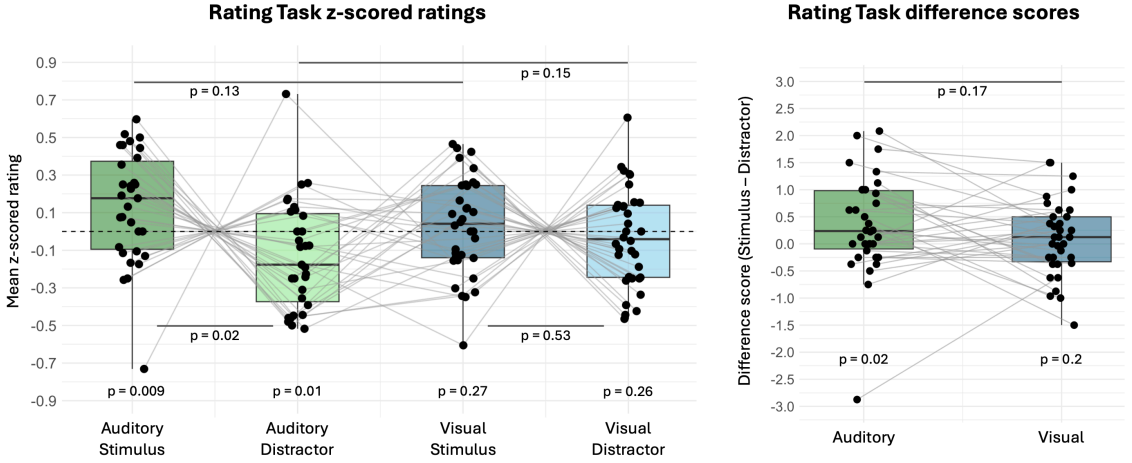


Figure 5.2: **Rating task performance across auditory and visual modalities. Left panel**: Mean z-scored ratings for trained stimuli and distractors in the auditory and visual rating tasks. Ratings were z-score normalized separately within each participant and modality to control for individual differences in response tendencies. Positive values indicate ratings above a participant's modality-specific average response level, whereas negative values indicate ratings below average. **Right panel**: Difference scores computed by subtracting distractor ratings from trained stimulus ratings within each modality. Positive values indicate greater endorsement of trained stimuli relative to distractors. Each point represents an individual child, lines connect responses from the same participant, boxplots depict the group distribution, and dashed horizontal lines indicate zero.

Direct comparisons between z-scored stimulus and distractor ratings again revealed a significant auditory learning effect ($t(30) = 2.48$, $p = .019$, mean difference = 0.26, 95% CI [0.05, 0.48]), whereas the visual difference remained non-significant ($t(33) = 0.63$, $p = .530$, mean difference = 0.06, 95% CI [-0.12, 0.24]).

To directly assess cross-modal transfer, modality-specific difference scores were computed by subtracting distractor ratings from trained-stimulus ratings within each modality (see Fig. 5.2, Right panel). Auditory learning difference scores and visual transfer difference scores were positively correlated across participants, although this association did not reach significance ($r(32) = .25$, $p = .162$, 95% CI [-0.10, 0.54]). This pattern remained similar after controlling for age using partial correlations ($r_p = .26$, $p = .139$). Corresponding analyses using z-scored difference scores likewise failed to reveal significant cross-modal associations ($r(29) = .13$, $p = .503$), and remained non-significant after controlling for age ($r_p = .13$, $p = .495$). Together, these findings suggest limited but directionally positive evidence for auditory-to-visual transfer at the individual level.

Finally, a transfer index was computed by subtracting visual transfer difference scores from auditory learning difference scores. Positive values therefore reflected relatively stronger auditory learning compared to visual transfer. The mean transfer index was positive ($M = 0.25$, $SD = 1.01$), although it did not significantly differ from zero ($t(33) = 1.42$, $p = .165$, 95% CI [-0.11, 0.60]). The transfer index was not significantly associated with age ($r(32) = -.20$, $p = .249$).

Overall, the rating-task results indicate robust ASL and weaker visual transfer effects. Although the direction of effects was generally comparable across modalities, evidence for reliable cross-modal transfer at the individual level was limited.

5.3.4 TOWRE-2 Performance

To assess children's reading abilities, participants completed the Test of Word Reading Efficiency–Second Edition (TOWRE-2),), which included separate timed Sight Word Reading Efficiency (word) and Phonemic Decoding Efficiency (nonword) reading subtests.

Raw accuracy scores revealed substantially higher performance for real words than nonwords. Across participants, mean word-reading accuracy exceeded nonword-reading accuracy by approximately 26 items. A paired-samples t-test confirmed that children read significantly more real words than nonwords ($t(32) = 15.38$, $p < .001$, 95% CI [22.32, 29.13]). This difference likely reflects the greater familiarity and lexical support associated with real-word reading relative to the phonological decoding demands of unfamiliar nonwords.

Importantly, performance across the two subtests was strongly positively correlated (r = .81, p < .001, 95% CI [.65, .90]), indicating that children with stronger word-reading abilities also tended to perform better on nonword decoding.

To account for age-related differences in reading development, raw scores were converted to age-normed standardized scores based on TOWRE-2 normative data (see Fig.5.3). Descriptively, age-normed scores were comparable across the word (M = 108.24, SD = 13.30, range = 70–130) and nonword (M = 110.00, SD = 12.20, range = 77–133) subtests.

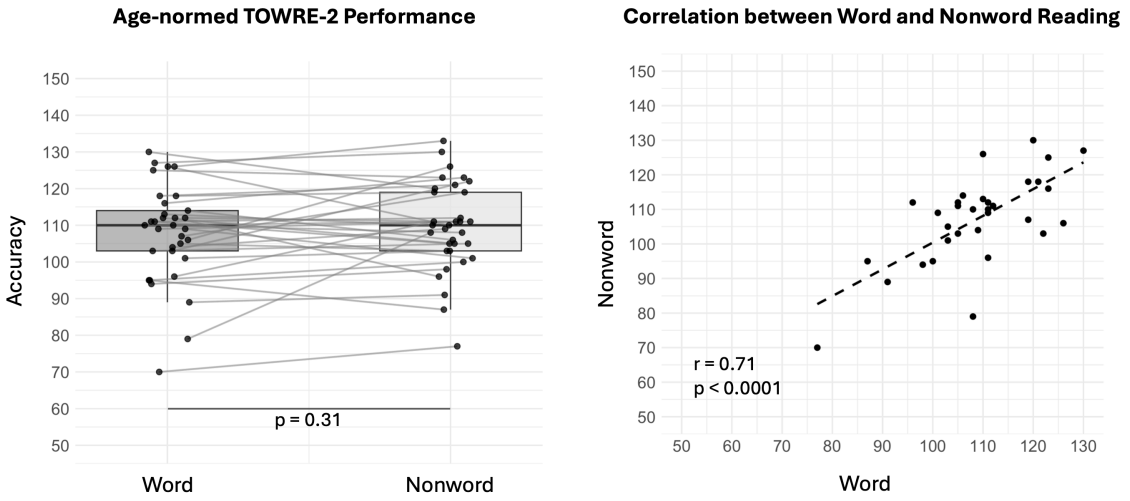


Figure 5.3: **TOWRE-2 age-normed reading performance. Left panel**: age-normed standardized scores for the word and nonword reading subtests of the TOWRE-2. Each point represents an individual child, grey lines connect scores from the same participant, and boxplots depict the group distribution. **Right panel**: Correlation between age-normed word and nonword reading scores across participants. Each point represents an individual child, and the dashed line indicates the linear regression fit.

A paired-samples t-test on age-normed scores revealed no significant difference between word and nonword performance ($t(32)$ = -1.04, p = .307, 95% CI [-5.21, 1.69]), suggesting that participants' decoding and sight-word reading abilities were broadly aligned relative to age expectations.

As with the raw scores, age-normed word and nonword performance were strongly positively correlated across participants (r = .71, p < .001, 95% CI [.49, .85]), indicating stable individual differences in reading ability across the two subtests.

Notably, substantial individual variability was observed across both word and nonword reading measures, despite age-normed performance remaining within the

typical range overall. This variability is consistent with the broad developmental heterogeneity expected during early reading acquisition.

5.3.5 Relations Between Age, Reading Ability, and Statistical Learning

To further investigate the factors associated with SL and cross-modal transfer, exploratory correlation analyses examined relations between age, reading ability, and rating-task performance. Rating-task performance was indexed using modality-specific difference scores computed by subtracting distractor ratings from trained-stimulus ratings. Positive values therefore reflected greater endorsement of trained items relative to distractors and provided a direct measure of learning while minimizing the influence of individual differences in overall response tendencies.

Reading ability was assessed using age-normed TOWRE-2 standard scores. Word and nonword reading scores were analyzed separately because they index partially distinct aspects of reading ability, with nonword decoding thought to more strongly reflect phonological and sublexical processing mechanisms potentially relevant for auditory learning and transfer. Given the relatively small sample size, all analyses were considered exploratory.

Exploratory correlation analyses examined associations between age and rating-task performance (see Fig. 5.4). In the auditory modality, no significant association was observed between age and auditory learning performance ($r(32) = -.04$, $p = .824$, 95% CI [-0.37, 0.30]). In the visual modality, age showed a weak positive association with visual transfer performance, although this relationship did not reach significance ($r(32) = .24$, $p = .170$, 95% CI [-0.11, 0.54]).

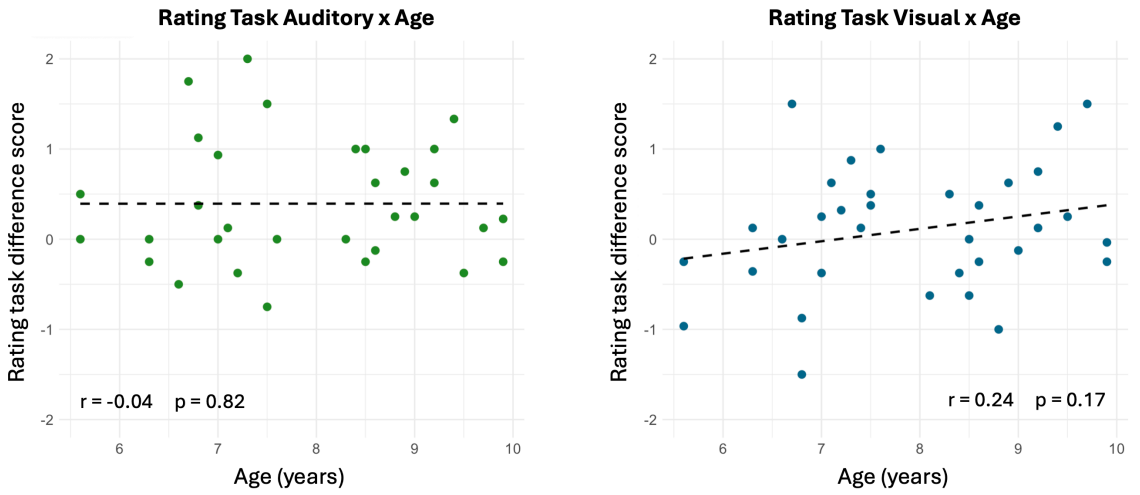


Figure 5.4: **Relations between age and rating-task performance.** Scatterplots show associations between age and modality-specific difference scores (trained stimuli minus distractors) in the auditory (left panel) and visual (right panel) rating tasks. Positive difference scores indicate greater endorsement of trained items relative to distractors. Each point represents an individual child, and dashed lines indicate linear regression fits.

Additional analyses examined whether auditory learning and visual transfer performance were related at the individual level (see Fig. 5.5). Auditory and visual difference scores were positively correlated across participants, although this association did not reach significance ($r(32) = .25$, $p = .162$, 95% CI [-0.10, 0.54]). This pattern remained similar after controlling for age using partial correlations ($r_p = .26$, $p = .139$). Corresponding analyses using z-scored difference scores likewise failed to reveal a reliable cross-modal association ($r(29) = .13$, $p = .503$), and remained non-significant after controlling for age ($r_p = .13$, $p = .495$). Together, these findings provide limited evidence that children who showed stronger auditory learning also exhibited stronger visual transfer performance.

To further characterize modality differences in learning performance, a transfer index was computed by subtracting visual transfer difference scores from auditory learning difference scores. Positive values therefore reflected relatively stronger auditory learning compared to visual transfer. The mean transfer index was positive ($M = 0.25$, $SD = 1.01$), although it did not significantly differ from zero ($t(33) = 1.42$, $p = .165$, 95% CI [-0.11, 0.60]). The transfer index was not significantly associated with age, $r(32) = -.20$, $p = .249$. In addition, the transfer index was not significantly associated with either age-normed word-reading scores ($r(32) = .20$, $p = .262$), or nonword-reading scores ($r(32) = -.01$, $p = .942$). These associations likewise

remained non-significant after controlling for age (word reading: $r_p = .16$, $p = .365$; nonword reading: $r_p = -.04$, $p = .834$).

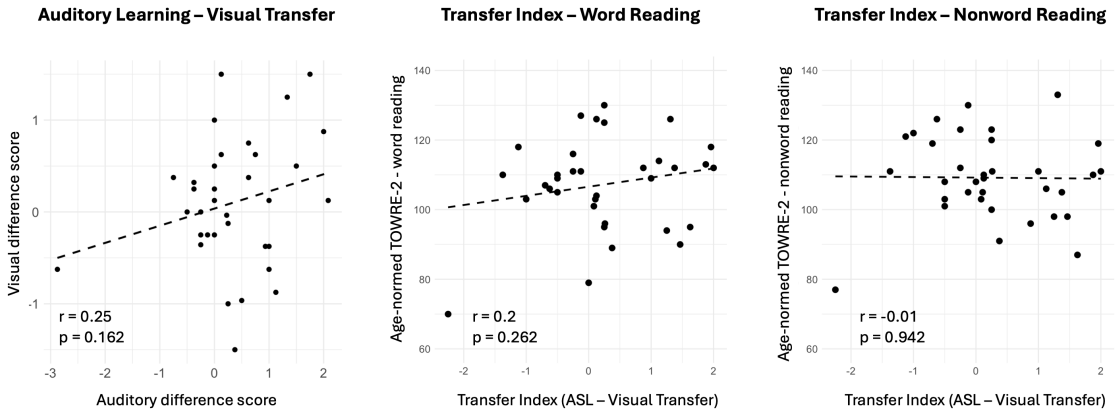


Figure 5.5: **Relations between auditory learning, visual transfer, and reading performance. Left panel**: Association between auditory learning difference scores and visual transfer difference scores. Difference scores were computed by subtracting distractor ratings from trained-stimulus ratings within each modality, such that positive values reflect greater endorsement of trained items relative to distractors. **Middle and right panels**: Associations between the transfer index (auditory learning minus visual transfer) and age-standardized TOWRE-2 word and nonword reading scores. Positive transfer-index values indicate relatively stronger auditory learning compared to visual transfer. Each point represents an individual child, and dashed lines indicate linear regression fits.

Exploratory correlation analyses next examined whether individual differences in auditory learning and visual transfer performance were related to standardized reading ability as measured by age-normed TOWRE-2 scores (see Fig. 5.6). Separate correlations were computed for word and nonword reading performance, as these tasks index partially distinct reading processes (i.e., the ability to read existing words vs. the ability to translate novel visual word forms into their phonemic representations).

In the auditory modality, auditory learning difference scores were significantly positively associated with age-normed word-reading scores ($r(32) = .51$, $p = .002$, 95% CI [.21, .72]), indicating that children with stronger sight-word reading abilities tended to show larger auditory learning effects. In contrast, auditory learning difference scores were not significantly related to age-normed nonword-reading scores ($r(32) = .22$, $p = .215$, 95% CI [-0.13, 0.52]).

In the visual modality, visual transfer difference scores were likewise significantly positively associated with age-normed word-reading scores ($r(32) = .39$, $p = .024$, 95% CI [.06, .64]), suggesting that stronger orthographic reading ability was linked to better visual transfer performance. Visual transfer scores also showed a positive but non-significant association with age-normed nonword-reading scores ($r(32) = .31$, $p = .079$, 95% CI [-0.04, 0.58]).

To determine whether developmental variability influenced these associations, additional partial correlations controlling for age were conducted. The overall pattern remained similar. Auditory learning difference scores continued to show a significant positive association with word-reading scores ($r_p = .51$, $p = .002$), but not with nonword-reading scores ($r_p = .22$, $p = .229$). Visual transfer difference scores also remained significantly associated with word-reading scores after controlling for age ($r_p = .46$, $p = .007$). In addition, the association between visual transfer and nonword reading became significant after controlling for age ($r_p = .35$, $p = .048$).

Overall, these findings suggest that individual differences in auditory learning and visual transfer were related to standardized reading ability, particularly word-reading performance. Children with stronger reading skills tended to exhibit larger auditory learning and visual transfer effects, even after accounting for developmental variability in age. In contrast, the relative discrepancy between auditory learning and visual transfer, as indexed by the transfer index, was not reliably associated with reading performance.

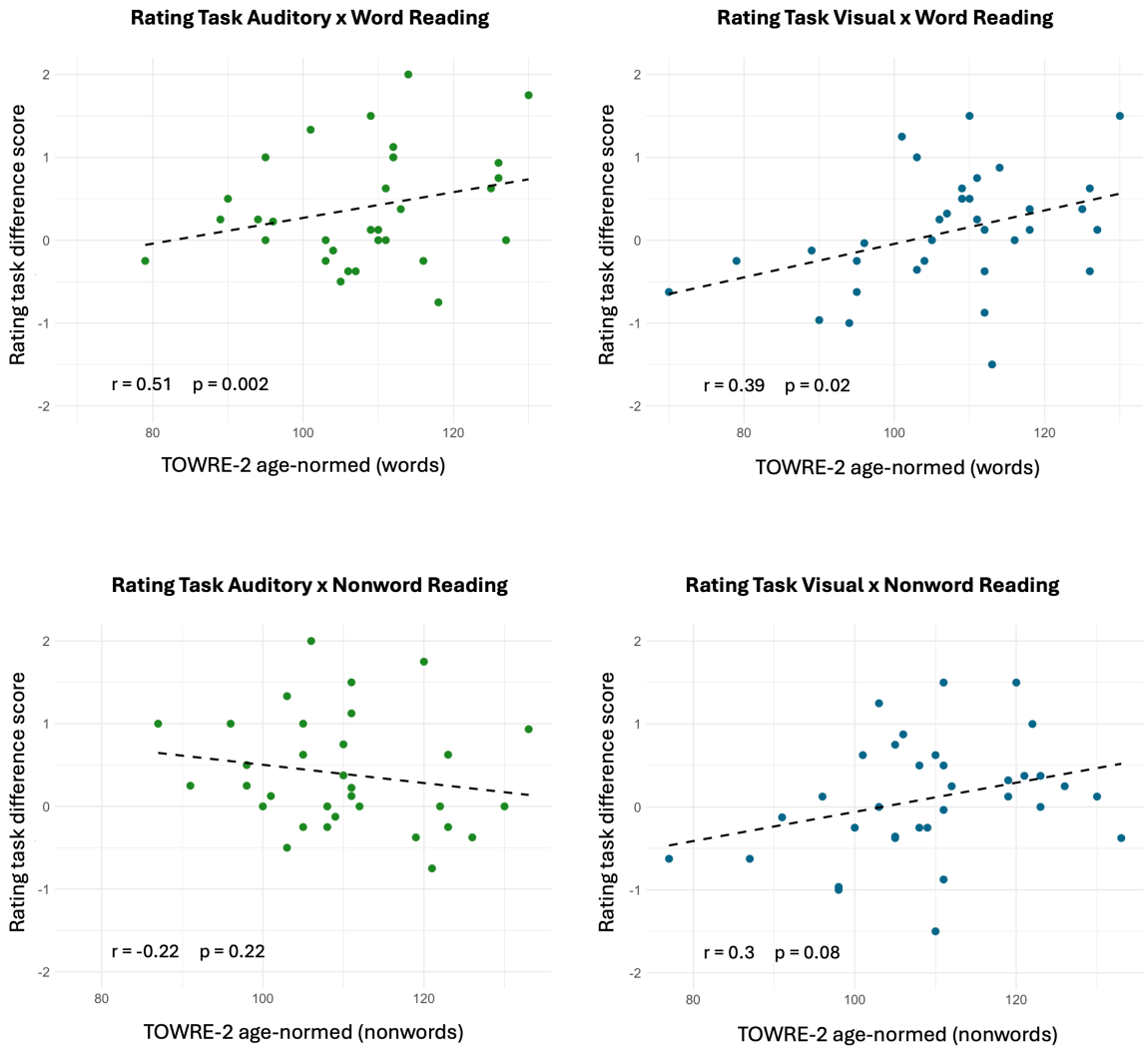


Figure 5.6: **Relations between rating-task performance and age-normed reading ability.** Scatterplots show associations between auditory learning (left panel) and visual transfer (right panel) difference scores and age-normed TOWRE-2 performance, with word reading perfomance on upper panels and nonword reading performance on lower panels. Difference scores were computed by subtracting distractor ratings from trained-stimulus ratings within each modality. Positive values therefore indicate greater endorsement of trained items relative to distractors. Each point represents an individual child, and dashed lines indicate linear regression fits.

5.4 Discussion

The present study investigated SL across auditory and visual modalities, cross-modal transfer, and their relation to literacy skills during early childhood. Specifically, we asked whether children aged 5–9 years would show reliable ASL, whether regularities acquired through auditory exposure would generalize to visually presented orthographic representations, and whether individual differences in learning performance would relate to standardized measures of reading ability. To address these questions, children were exposed to an artificial "penguin language" and subsequently completed auditory and visual judgment tasks together with standardized reading assessments (TOWRE-2). Overall, the findings revealed reliable ASL, weaker but directionally consistent evidence for visual transfer, substantial individual variability in learning performance, and significant associations between SL measures and literacy abilities.

Auditory and Visual Statistical Learning in Childhood

Consistent with extensive previous work demonstrating robust abilities in infancy and childhood (Saffran et al., 1996a; Kirkham et al., 2002; Fiser and Aslin, 2002), children in the present study successfully extracted regularities from the auditory speech stream. Across both raw and z-score normalized analyses, auditory trained items received significantly higher ratings than distractors, indicating successful learning of the artificial language structure. These findings align with evidence suggesting that children can segment continuous speech based on transitional probabilities and acquire sequential regularities from auditory input (Evans et al., 2009; Conway and Pisoni, 2008; Abreu et al., 2023). Importantly, auditory learning effects remained relatively stable across the sampled age range, consistent with proposals that ASL may emerge early and remain comparatively robust throughout childhood (Raviv and Arnon, 2018; Saffran et al., 1997).

Evidence for learning in the visual modality was weaker but followed a similar directional pattern. Although children did not significantly discriminate visual trained items from distractors at the group level, visual ratings were numerically biased toward trained orthographic strings, and visual difference scores were positively associated with reading ability. These findings suggest partial cross-modal transfer from previously learned auditory regularities to visual representations. Such transfer is theoretically important because reading acquisition requires children to connect phonological and orthographic representations across modalities (Garcia et al., 2019; Dehaene et al., 2015). The present findings therefore provide tentative support for

the idea that children can generalize sequential regularities acquired through auditory exposure to visually presented letter strings, even when transfer effects remain relatively modest.

The weaker visual effect likely reflects the additional cognitive demands associated with cross-modal generalization. Unlike the auditory task, which directly assessed learning within the familiarized modality, the visual task required children to map learned phonological representations onto novel orthographic forms. This process may place greater demands on orthographic knowledge, working memory, attentional control, and emerging phonological–orthographic integration abilities. Importantly, children demonstrated near-ceiling performance in the letter identification task, suggesting that the limited visual effect cannot simply be explained by unfamiliarity with the letters themselves. Rather, the findings are more consistent with proposals that SL operates under modality-specific constraints (Frost et al., 2015; Siegelman and Frost, 2015). At the same time, successful reading development depends critically on learning regularities linking orthography, phonology, and meaning across representational systems (Siegelman, 2020b). Similar to previous developmental work showing stronger learning effects within some modalities than others (Pavlidou and Bogaerts, 2019), the present results therefore suggest that learning and transfer may depend not only on shared statistical computations, but also on the representational and processing demands associated with integrating information across modalities.

Substantial individual variability was observed across both auditory learning and visual transfer performance. This variability is particularly noteworthy given the relatively high reliability estimates obtained for the judgment tasks, especially in the visual modality. Together, these findings support the idea that SL abilities vary meaningfully across children and may contribute differently to cognitive and literacy development. At the same time, the relatively modest group-level visual transfer effect highlights the possibility that cross-modal generalization remains fragile during the early stages of literacy acquisition and may continue to strengthen with reading experience and increasing integration between phonological and orthographic systems.

Cross-Modal Transfer between Auditory and Visual Statistical Learning

A central goal of the present study was to investigate whether statistical regularities acquired through auditory exposure would generalize to visual orthographic representations. Although auditory learning effects were stronger overall, children showed directionally positive response biases toward trained stimuli across both modalities. While the visual transfer effect did not reach significance at the group level, these

findings provide tentative evidence consistent with partial auditory-to-visual transfer of statistical regularities.

Interestingly, substantial individual variability was observed in visual transfer performance. Children who exhibited stronger visual transfer effects also tended to demonstrate better standardized reading abilities, particularly for age-standardized word reading performance. This pattern suggests that individual differences in cross-modal transfer may already be meaningfully related to literacy-related processes during early childhood, even when average transfer effects remain relatively modest.

Importantly, auditory learning and visual transfer difference scores were positively correlated across participants, although this association did not reach significance. This pattern suggests limited but directionally consistent evidence that children who showed stronger auditory learning also tended to exhibit stronger cross-modal transfer.

The present findings contribute to the broader literature examining whether SL generalizes across sensory modalities. Previous work has reported mixed evidence regarding cross-modal transfer (Mitchel and Weiss, 2011b; Conway and Christiansen, 2005). Some studies suggest that learning is constrained by modality-specific perceptual systems (Siegelman and Frost, 2015; Pavlidou and Bogaerts, 2019), whereas others propose the existence of more domain-general SL mechanisms (Frost et al., 2015). Our findings are broadly consistent with intermediate accounts proposing that shared computational principles operate within partially modality-specific representational systems (Frost et al., 2015). In the present context, children may have extracted sequential regularities from the auditory stream while only partially generalizing this knowledge to visually presented orthographic forms.

The relatively modest visual transfer effect is also consistent with developmental accounts suggesting that cross-modal integration processes continue to mature throughout early literacy acquisition. Learning to read requires the integration of phonological and orthographic regularities across modalities (Sawi and Rueckl, 2019), and successful reading acquisition likely depends on increasingly efficient mappings between auditory and visual information (Garcia et al., 2019). The present findings therefore raise the possibility that the ability to transfer statistical structure across modalities develops gradually alongside reading experience and orthographic knowledge.

At the same time, the weak average visual transfer effect may also reflect methodological and task-related factors. SL performance in children is known to vary substantially as a function of task structure, modality, and response demands (Isbilen and Christiansen, 2022; Pavlidou and Bogaerts, 2019). In the present study, the visual rating task required children not only to recognize previously learned regu-

larities, but also to generalize them to novel orthographic forms using a Likert-scale response procedure. These additional processing demands may have increased variability and reduced sensitivity to subtle transfer effects at the group level.

Statistical Learning, Literacy, and Developmental Factors

The present study additionally examined whether SL performance was related to literacy abilities measured using the TOWRE-2. In contrast to the relatively limited age-related effects observed across tasks, SL performance showed meaningful associations with standardized reading ability. Children with stronger age-standardized word reading scores exhibited larger ASL effects as well as stronger visual transfer performance. Moreover, visual transfer performance showed a positive association with nonword decoding ability after controlling for age.

These findings are broadly consistent with previous work reporting positive associations between SL and literacy-related skills (Arciuli and Simpson, 2012; Spencer et al., 2015; Qi et al., 2019). Sensitivity to sequential regularities has been proposed to support the acquisition of orthographic and phonological structure during reading development. At the same time, the literature linking SL and reading remains mixed, with some studies reporting weak or inconsistent associations depending on task demands, modality, and developmental stage (Schmalz et al., 2019; West et al., 2018). The present findings therefore contribute to growing evidence suggesting that SL may be related to literacy development, while also highlighting that these relations are likely influenced by the specific learning and reading processes being measured.

Interestingly, associations with SL were stronger for age-corrected word reading than for nonword decoding. This pattern may suggest that the present paradigm was more closely related to the acquisition or recognition of familiar orthographic and lexical representations than to phonological decoding processes alone. In particular, following auditory familiarization, children were required to recognize orthographic versions of previously learned auditory forms in the visual rating task. In this respect, the task may more closely resemble sight-word recognition processes, such as those assessed by the TOWRE-2 Sight-word Reading Efficiency subtest, rather than the phonological decoding demands emphasized in nonword reading. Successful performance therefore likely depended not only on sensitivity to sequential regularities, but also on the ability to link previously learned phonological representations with emerging orthographic representations.

In contrast, age itself showed relatively limited associations with task performance within the sampled range. Auditory learning performance was largely unrelated to age, whereas visual transfer showed only a weak positive trend. These findings sug-

gest that individual differences in SL and transfer may not simply reflect general developmental maturation during this period, but may instead relate more specifically to literacy-related abilities and cross-modal integration processes.

Importantly, the observed associations between SL and reading ability suggest that sensitivity to newly acquired sequential regularities may be meaningfully linked to literacy development during early childhood. Children with stronger reading abilities may more efficiently encode, retain, or generalize structured linguistic information across modalities. The association between visual transfer performance and standardized reading ability is particularly noteworthy in this context. Previous work has proposed that visual SL may contribute to orthographic acquisition and grapheme–phoneme mapping processes (Pavlidou and Bogaerts, 2019; Ktori et al., 2026). Although the present findings remain correlational, they are broadly consistent with this account and suggest that children who more effectively generalize auditory regularities to visual representations also tend to exhibit stronger reading skills.

At the same time, the relative discrepancy between auditory learning and visual transfer, indexed by the transfer measure, was not reliably associated with reading ability. This pattern may partly reflect methodological limitations of the present paradigm rather than the absence of a meaningful relationship between transfer and literacy. In particular, the relatively modest differences observed between trained stimuli and distractors in the rating tasks may have reduced sensitivity to individual differences in transfer performance. More generally, statistical learning effects in developmental populations are often small and highly variable across participants and tasks (Isbilen and Christiansen, 2022; Siegelman and Frost, 2017). Future studies using larger samples and more sensitive implicit or online measures may therefore be better positioned to determine whether individual differences in cross-modal transfer are reliably associated with literacy development.

Limitations and Future Directions

The present study employed a child-friendly rating procedure adapted from recent developmental SL work (Schuler et al., 2026), which may provide greater sensitivity to graded learning effects than binary forced-choice paradigms. Several limitations of the present study should be acknowledged.

First, the sample size was relatively small, which may limit the stability and generalizability of the observed associations between SL and literacy measures. SL studies in developmental populations often reveal substantial variability across participants (Siegelman and Frost, 2017), and larger samples will therefore be neces-

sary to more reliably characterize developmental trajectories and the specificity of learning–reading associations across modalities and tasks. Future studies with larger samples may additionally benefit from trial-level mixed-effects modeling approaches that simultaneously account for participant-level variability and developmental predictors such as age and reading ability.

Second, although the child-adapted Likert-scale procedure appeared generally well understood, it remains possible that some children did not fully or consistently apply the response procedure despite the practice phase. Young children may vary considerably in how they interpret and use graded response scales, potentially introducing additional measurement variability. Future studies could therefore complement rating-based measures with alternative approaches such as forced-choice paradigms, reaction-time indices, eye-tracking, or production-based tasks.

Production paradigms may be particularly informative for assessing individual differences in SL and generalization. Previous work suggests that tasks requiring active reproduction or generalization of learned structure may better capture abstract representations than offline familiarity judgments alone (Ren and Wang, 2024). Similarly, longitudinal artificial language studies indicate that generalization abilities may predict later language outcomes more strongly than segmentation performance (Ren and Wang, 2024). Future work could therefore investigate whether children are able not only to recognize learned regularities, but also to actively reproduce or generalize orthographic and phonological patterns across modalities.

Several broader theoretical questions also remain unresolved. Much of the developmental SL literature relies on cross-sectional designs or combines children across relatively broad age ranges, making it difficult to characterize developmental trajectories during the critical period of literacy acquisition. Longitudinal studies following children from prereading stages through later reading development would be particularly valuable for determining whether early SL abilities prospectively relate to later literacy outcomes. In addition, auditory-to-visual transfer remains comparatively underexplored despite its potential relevance for orthographic learning and grapheme–phoneme integration. The present findings suggest that children may show limited but meaningful transfer of learned auditory regularities to orthographic representations, highlighting the importance of further investigating how cross-modal generalization contributes to literacy acquisition. Future studies should therefore further examine how children generalize auditory statistical regularities to written representations using developmentally appropriate and ecologically valid paradigms.

More broadly, the relationship between laboratory-based SL tasks and natural language learning remains incompletely understood. Artificial language paradigms provide strong experimental control, but they likely capture only a subset of the pro-

cesses involved in real-world language and literacy acquisition. Combining controlled experimental paradigms with more naturalistic reading and language measures may therefore provide a more comprehensive understanding of how SL contributes to literacy development.

Overall, the present findings provide evidence that children can reliably extract statistical regularities from auditory input and show limited but directionally consistent transfer of this knowledge to visual orthographic representations. Importantly, individual differences in auditory learning and visual transfer performance were associated with standardized reading ability, particularly word-reading performance. Together, these findings support the possibility that SL and cross-modal integration processes may contribute to early literacy development. Further longitudinal and multimethod research will be important for clarifying how these mechanisms interact across development and how they support the acquisition of written language.

5.5 Conclusion

The present study investigated auditory SL, auditory-to-visual transfer, and their relation to literacy development in young children. Consistent with previous developmental research, children demonstrated ASL and showed weaker but directionally consistent evidence of visual transfer following auditory familiarization. Importantly, individual differences in auditory learning and visual transfer performance were positively associated with standardized reading ability, particularly age-standardized word-reading performance. These findings suggest that sensitivity to sequential regularities and the ability to generalize learned structure across modalities may be meaningfully related to early literacy development.

At the same time, age showed relatively limited associations with rating-task performance within the sampled developmental range, suggesting that individual differences in learning may not simply reflect general maturation alone. Together, the findings support the view that children can extract sequential regularities from structured input and show limited but potentially meaningful transfer of this knowledge across sensory modalities during the early stages of literacy acquisition.

More broadly, the present results contribute to ongoing debates regarding the extent to which SL reflects modality-general versus modality-specific mechanisms. The stronger auditory learning effects, together with the weaker but literacy-related visual transfer effects, are broadly consistent with accounts proposing shared computational principles operating within partially modality-specific systems. Importantly, auditory learning and visual transfer performance showed a directionally positive association across children, although this relationship did not reach significance, suggesting limited but potentially meaningful individual differences in cross-modal generalization. Future longitudinal and multimethod research will be important for clarifying how ASL and cross-modal transfer interact across development and how these processes contribute to the emergence of reading and orthographic knowledge.

Appendix

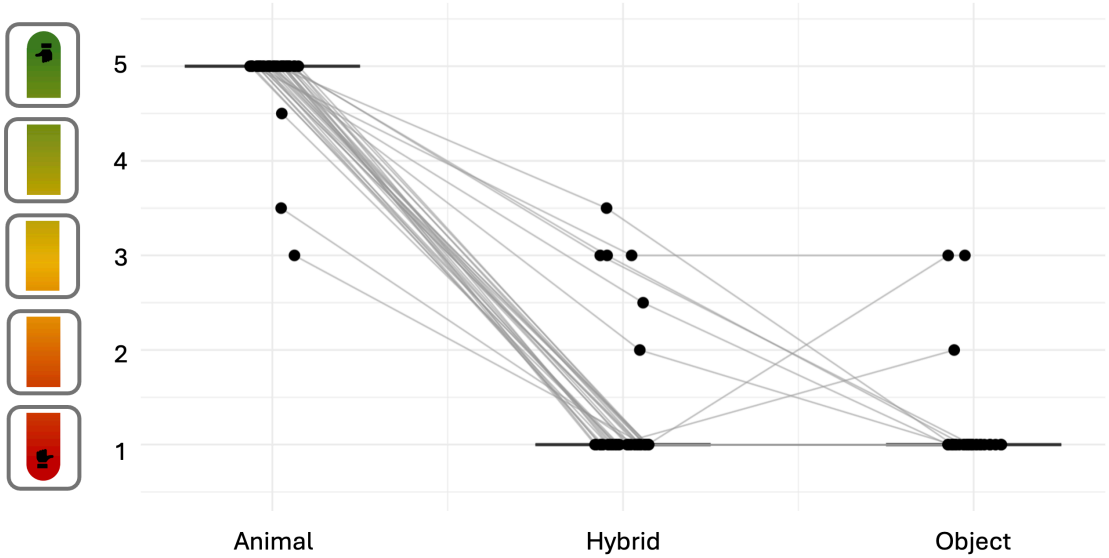


Figure A5.1: **Practice Task results.** Likert response scale adapted for the experiment. Stimuli consisted of two animal images, two object images, and two ambiguous animal–object hybrid images generated using artificial intelligence tools. Children were asked to judge whether each image depicted an animal, with higher ratings indicating stronger endorsement. The inclusion of hybrid stimuli encouraged use of intermediate points on the response scale and familiarized children with graded responses prior to the main experimental tasks.

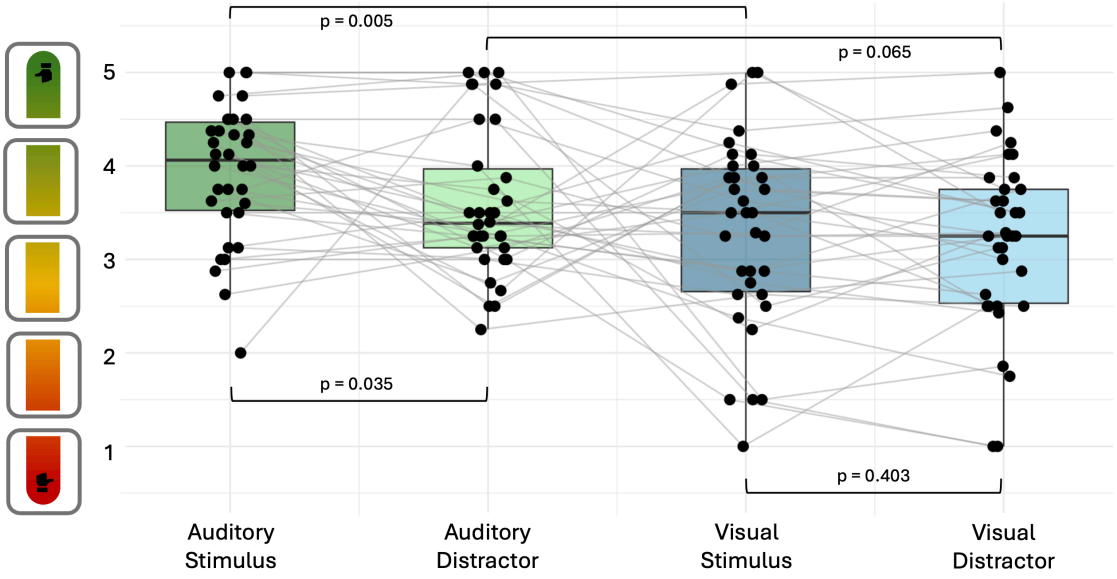


Figure A5.2: **Raw ratings in the Auditory and Visual Judgment Tasks.** Mean Likert ratings (1–5) assigned to trained stimuli and distractors in the auditory and visual modalities. Higher values indicate stronger endorsement that an item belonged to the artificial "penguin language." Each point represents an individual child, grey lines connect responses from the same participant, and boxplots depict the group distribution. Auditory trained stimuli received significantly higher ratings than the scale midpoint, whereas visual responses showed weaker but directionally similar patterns. Reported p-values indicate pairwise comparisons between conditions.

Chapter 6

General Discussion

The present thesis investigated statistical learning (SL) across auditory and visual modalities using both behavioral and neural methodologies in children, young adults, and older adults. Building on contemporary perspectives that conceptualize SL as a fundamental mechanism for extracting regularities from environmental input (Frost et al., 2019; Christiansen, 2019), the thesis addressed several unresolved questions concerning the nature of SL across modalities, tasks, and developmental stages. More specifically, the thesis examined whether SL reflects a largely domain-general learning mechanism or a set of partially dissociable processes constrained by modality and task demands; whether neural and behavioral measures capture shared or distinct aspects of learning; how SL varies across the lifespan; and how inter-individual variability can inform theoretical models of SL. Across all empirical chapters, particular emphasis was placed on the relationship between online neural sensitivity to regularities and offline behavioral expressions of learning, as well as on the broader question of whether SL should be conceptualized as a unitary or multicomponential construct (Frost et al., 2015; Siegelman and Frost, 2017; Bogaerts et al., 2022a).

Some of the theoretical issues discussed in this chapter were also addressed during a workshop that led to *"Time for Consensus: Open Challenges and Recommendations for Research on Individual Differences in Statistical Learning"* manuscript, currently under revision for publication in the *Perspectives on Psychological Science*. The pre-print can be found on PsyArXiv at `https://osf.io/preprints/psyarxiv/eh7fm_v1`.

The empirical chapters approached these questions through complementary paradigms and populations. Chapter 2 examined auditory SL in young adults using an EEG frequency-tagging paradigm based on fast periodic visual stimulation (FPVS). Neural responses revealed robust frequency-specific sensitivity to structured regularities despite limited evidence from explicit behavioral measures, suggesting a dissociation between neural entrainment and conscious behavioral access to learned structure. Chapter 3 extended this approach to older adults and demonstrated preserved neural sensitivity to statistical regularities despite increased variability in behavioral performance. Together, these findings suggested that online neural sensitivity to structured input may remain relatively preserved across adulthood and aging even when explicit behavioral expressions of learning appear weaker or less stable.

Chapter 4 investigated cross-modal SL and transfer in young adults using behavioral paradigms assessing auditory and visual learning. The findings revealed partial dissociations across modalities, with learning and transfer varying as a function of sensory domain and task demands. These results aligned with previous evidence suggesting that SL relies on both shared and modality-dependent computations rather than on a fully domain-general mechanism (Conway and Christiansen, 2005; Frost et al., 2015; Qi et al., 2019). Finally, Chapter 5 examined cross-modal SL in children during a developmental period characterized by ongoing literacy acquisition and sensory-cognitive maturation. The results highlighted substantial inter-individual variability in learning across modalities and suggested developmental asymmetries in the interaction between auditory and visual statistical structure. More broadly, the findings emphasized the importance of considering developmental constraints, attentional factors, and modality-specific processing demands when interpreting SL performance during childhood (Kidd et al., 2018; Forest et al., 2023; Arnon, 2020).

The findings of the present thesis revealed four recurring empirical themes that form the foundation of the present discussion. First, learning varied systematically across auditory and visual modalities, supporting the view that SL is shaped by modality-specific constraints despite relying on partially shared computational principles. Second, substantial differences emerged between neural and behavioral expressions of learning, with neural entrainment often revealing sensitivity to regularities that was only weakly reflected in explicit behavioral measures. Third, learning profiles differed across developmental stages, suggesting that the mechanisms supporting SL may mature and decline unevenly across the lifespan. Finally, marked inter-individual variability emerged across tasks and populations, reinforcing recent arguments that variability in SL is theoretically informative rather than mere experimental noise (Siegelman and Frost, 2015; Kidd et al., 2018; Bogaerts et al., 2022b).

Together, these findings support the view that SL is best conceptualized as a

multicomponential phenomenon emerging from partially dissociable learning processes that vary across modalities, tasks, representational formats, and developmental stages. Rather than reflecting a single unified cognitive ability, the present findings suggest that different paradigms may index distinct aspects of learning that differ in processing demands and accessibility to conscious report. In particular, neural entrainment and FPVS paradigms may primarily capture early online sensitivity to statistical regularities and perceptual binding processes, whereas behavioral tasks may additionally recruit memory retrieval, attentional control, decision-making, and metacognitive reflection (Batterink et al., 2015; Batterink and Paller, 2017; Siegelman and Frost, 2017). This interpretation is broadly consistent with recent theoretical frameworks proposing that SL involves multiple interacting computations distributed across modalities, domains, and representational levels (Thiessen et al., 2013; Conway, 2020; Bogaerts et al., 2022b).

The following sections develop these issues in greater detail by first discussing the empirical patterns emerging across modalities, developmental stages, and measurement levels, and then examining their broader theoretical and methodological implications for understanding the structure of SL.

6.1 Empirical Patterns Across Modalities, Measures, and Age

Before turning to the broader theoretical implications of the present findings, it is important to consider the main empirical patterns that emerged across the studies included in this thesis. Although the individual chapters employed different paradigms, populations, and measures, several consistent trends could be identified across the results. In particular, the findings revealed systematic differences related to (1) the relationship between neural and behavioral expressions of learning, (2) developmental and aging-related changes across the lifespan, and (3) modality-specific constraints on learning and cross-modal transfer. Importantly, these patterns did not emerge independently. Rather, they interacted in complex ways across tasks and populations, suggesting that SL may reflect the coordinated operation of multiple learning processes differing in representational format, accessibility to conscious report, and sensitivity to developmental and sensory constraints. The following sections examine these empirical regularities in greater detail before considering their broader theoretical and methodological implications.

Statistical Learning Across the Lifespan

A growing body of research suggests that statistical learning (SL) is present across the lifespan, from infancy to older adulthood, supporting the idea that sensitivity to environmental regularities constitutes a fundamental property of human cognition (Saffran et al., 1996a; Saffran and Kirkham, 2017; Frost et al., 2019). At the same time, developmental and aging studies increasingly indicate that SL cannot be understood as a static or unitary ability that remains unchanged across life stages. Rather, developmental differences appear to emerge as a function of changes in attentional control, working memory, sensory processing, executive functioning, literacy experience, and metacognitive access to learned representations (Kidd et al., 2018; Forest et al., 2023; Arnon, 2020). From this perspective, lifespan investigations provide important insight into the multicomponential nature of SL by revealing that different aspects of learning may mature, interact, and decline unevenly across development and aging.

Developmental studies have consistently shown that even infants are capable of extracting statistical regularities from auditory and visual input (Saffran et al., 1996a; Saffran and Kirkham, 2017). However, the mechanisms supporting learning during childhood may differ substantially from those operating in adulthood. Recent developmental frameworks emphasize that observed age-related differences in SL may partly reflect the maturation of domain-general cognitive systems rather than changes in learning computations themselves (Forest et al., 2023). Attention, working memory, processing speed, and executive control continue to develop throughout childhood and adolescence, and these systems may strongly influence how statistical regularities are encoded, maintained, and behaviorally expressed (Kidd et al., 2018, 2023). Importantly, this creates a major interpretative challenge: developmental differences observed in SL tasks may not necessarily reflect differences in learning per se, but differences in the cognitive demands required by specific paradigms (Arnon, 2020; Forest et al., 2023).

This issue becomes particularly relevant when considering cross-modal learning and modality-specific processing during literacy acquisition. Several studies have reported stronger associations between auditory statistical learning and language-related outcomes compared to visual learning measures (Qi et al., 2019; Isbilen et al., 2022). Auditory sequential processing may therefore provide an especially important scaffold for language acquisition and early reading development. At the same time, visual learning increasingly becomes shaped by orthographic experience and literacy-related exposure throughout development (Siegelman, 2020b; Lelonkiewicz et al., 2023). Such findings are consistent with proposals suggesting that SL reflects partially shared learning principles operating over modality-dependent representa-

tional systems (Frost et al., 2015; Conway, 2020). The developmental findings of the present thesis align with this broader literature by suggesting that children may show asymmetrical sensitivity to auditory and visual regularities and substantial variability in how learning transfers across modalities.

Importantly, developmental variability in SL has increasingly been interpreted as theoretically meaningful rather than as experimental noise. Research on individual differences has shown that children vary considerably in their sensitivity to statistical structure and that these differences can predict aspects of language and literacy development (Kidd, 2012; Kidd and Arciuli, 2015; Arciuli and Simpson, 2012). At the same time, recent large-scale reviews and consensus discussions have emphasized that these relations are often modest, task-dependent, and difficult to interpret due to issues of construct validity and measurement reliability (Bogaerts et al., 2022b; Boeve et al., 2024; Ren et al., 2023). The present findings are broadly compatible with this view, highlighting substantial variability across modalities and tasks during childhood rather than a single coherent developmental trajectory.

Studies of aging similarly challenge the notion of SL as a uniformly preserved or uniformly declining ability. While some forms of implicit and procedural learning appear relatively stable across older adulthood, age-related changes often emerge in tasks requiring explicit access, attentional control, episodic memory, or strategic retrieval (Wang et al., 2023; Cisneros et al., 2024). For example, Moreau et al. (2022) reported comparable neural entrainment to statistical regularities in children and adults despite limited behavioral differences, suggesting relative preservation of online neural sensitivity across development. Similarly, Wang et al. (2023) showed that older adults could maintain statistical learning abilities despite declines in behavioral pattern separation, indicating that implicit extraction of regularities and explicit representational precision may rely on partially distinct systems. Such findings support the idea that aging may differentially affect the accessibility and behavioral expression of learned representations rather than the extraction of regularities itself.

The results of the present thesis are broadly consistent with these perspectives. Neural frequency-tagging responses observed in older adults suggested preserved sensitivity to structured auditory input despite greater variability in explicit behavioral measures. Rather than reflecting a simple preservation-versus-decline distinction, these findings fit more naturally within multicomponential frameworks proposing that different SL measures rely on distinct combinations of perceptual, attentional, mnemonic, and metacognitive processes (Siegelman and Frost, 2017; Conway, 2020; Bogaerts et al., 2022b). Neural entrainment paradigms may therefore capture relatively automatic or online aspects of learning that remain comparatively robust with aging, whereas explicit behavioral tasks may become increasingly sensitive to

age-related variability in executive and memory systems.

Overall, evidence across development and aging suggests that SL is shaped by the interaction between core learning computations and broader cognitive systems that themselves change across the lifespan. Rather than supporting a single developmental trajectory, the available literature points toward partially independent trajectories across modalities, representational levels, and measurement types. The present findings contribute to this perspective by showing that developmental and aging-related differences in SL emerge most clearly when considering how learning is expressed across different modalities and across neural versus behavioral measures, rather than by treating SL as a single unified ability.

Modality Constraints and Cross-Modal Learning

One of the central debates in statistical learning (SL) research concerns the extent to which learning mechanisms are domain-general or constrained by sensory modality and representational format. Early demonstrations of SL across auditory, visual, and tactile domains were initially interpreted as evidence for a broadly domain-general mechanism capable of extracting regularities independently of stimulus modality (Saffran et al., 1996a; Conway and Christiansen, 2005). However, subsequent research increasingly challenged this interpretation, revealing substantial differences in learning performance across modalities, stimulus types, and task structures (Frost et al., 2015; Siegelman and Frost, 2017; Zhou et al., 2024). These findings have led to more nuanced theoretical accounts proposing that SL relies on partially shared computational principles operating over modality-constrained representational systems rather than on a single fully amodal mechanism (Frost et al., 2015; Conway, 2020).

A major source of evidence for modality constraints comes from studies reporting weak or inconsistent correlations between auditory and visual SL tasks. Although learners are generally capable of extracting regularities in both modalities, performance often varies substantially depending on whether information is presented sequentially in time or spatially in vision (Conway and Christiansen, 2005; Zhou et al., 2024). Frost et al. (2015) described this tension as the "paradox of statistical learning": SL appears sufficiently general to operate across domains and sensory systems, yet sufficiently constrained to produce strong modality and stimulus effects. This paradox has become one of the defining theoretical challenges in the field and has motivated increasing interest in identifying which aspects of SL are shared across modalities and which remain modality-specific.

Several theoretical frameworks have attempted to reconcile these apparently con-

tradictory findings. Conway and Christiansen (2005) proposed that while common computational principles may govern the extraction of regularities, learning remains shaped by modality-specific perceptual and representational constraints. Similarly, Siegelman et al. (2017) argued that SL should be conceptualized as a multidimensional construct varying across modality, domain, type of regularity, and task demands. More recently, Zhou et al. (2024) demonstrated that learning performance differs systematically not only across auditory and visual modalities but also across temporal versus spatial regularities within the same modality. Together, these findings suggest that modality effects are not peripheral methodological artifacts but rather reflect important properties of how regularities are represented and processed across sensory systems.

Evidence from language and literacy research further supports the existence of modality asymmetries. Auditory statistical learning has often shown stronger and more consistent associations with language acquisition and reading abilities than visual SL tasks (Frost et al., 2013b; Qi et al., 2019; Isbilen et al., 2022). Qi et al. (2019), for example, reported that auditory—but not visual—SL predicted reading performance, suggesting that sequential auditory processing may play a particularly important role in language-related learning. Similarly, meta-analytic work by Isbilen and Christiansen (2022) argued that auditory-linguistic SL paradigms may better capture ecologically relevant learning processes involved in natural language acquisition. These findings challenge strong domain-general accounts and instead support the idea that the functional relevance of SL may depend partly on the modality and representational characteristics of the learning environment.

At the same time, evidence for at least partial cross-modal transfer suggests that SL cannot be reduced to entirely independent modality-specific systems. Several studies have shown that prior exposure to regularities in one modality can influence processing in another modality, although transfer effects are often asymmetric and task-dependent (Conway and Christiansen, 2005; Qi et al., 2019). Such findings suggest that some higher-order computations involved in detecting structure, prediction, or chunking may generalize across sensory systems even when the representations themselves remain modality-bound. From this perspective, cross-modal learning may depend on interactions between shared learning principles and modality-specific encoding constraints rather than on a purely amodal learning mechanism.

The findings of the present thesis align with this broader literature. Cross-modal learning and transfer varied as a function of modality, developmental stage, and task demands, with auditory regularities often appearing more robustly learned or transferred than visual regularities. Importantly, however, the results did not support a fully modality-specific interpretation either. Instead, the observed patterns are

more consistent with frameworks proposing partially overlapping learning computations operating over modality-dependent representations (Frost et al., 2015; Conway, 2020). In this view, auditory and visual SL may rely on shared principles such as sensitivity to transitional probabilities, chunking, or predictive processing, while differing in the perceptual, attentional, and representational systems through which these computations are implemented.

Developmental findings provide additional support for this interpretation. During childhood, cross-modal learning may be particularly influenced by literacy acquisition, audiovisual integration, and ongoing maturation of attentional and executive systems (Forest et al., 2023; Kidd et al., 2018). Visual SL becomes increasingly shaped by orthographic and reading experience, whereas auditory sequential processing may provide an earlier and more developmentally stable basis for extracting regularities relevant to language learning (Qi et al., 2019; Siegelman, 2020b). Such developmental asymmetries further suggest that modality effects emerge through dynamic interactions between learning mechanisms and sensory-cognitive experience rather than through fixed modality-specific architectures alone.

Overall, the available evidence suggests that SL is neither fully domain-general nor entirely modality-specific. Rather, modality effects appear to emerge because common learning computations operate over sensory systems that differ in representational format, temporal dynamics, attentional constraints, and ecological function. The present findings contribute to this perspective by showing that modality-related dissociations become especially visible when learning is examined across different tasks, developmental stages, and levels of behavioral versus neural expression. These patterns therefore support conceptualizations of SL as a multicomponential phenomenon shaped by interactions between shared computational principles and modality-dependent representational constraints.

Neural and Behavioral Expressions of Statistical Learning

One of the most consistent and theoretically important patterns emerging from recent statistical learning (SL) research concerns the often weak or inconsistent relationship between neural and behavioral expressions of learning. Although SL has traditionally been assessed primarily through explicit offline behavioral tasks such as two-alternative forced-choice (2AFC) paradigms or familiarity judgments, increasing use of online neural and behavioral measures has revealed substantial dissociations across measurement types (Batterink and Paller, 2017; Batterink et al., 2015; Siegelman and Frost, 2017). These findings have challenged the assumption that different SL measures necessarily reflect redundant indices of a single underlying learning

process. Instead, accumulating evidence suggests that neural and behavioral tasks may capture partially dissociable aspects of learning that differ in representational accessibility, processing demands, and sensitivity to conscious report.

One important distinction emerging in the literature concerns the contrast between online and offline measures of SL (Batterink and Paller, 2017; Siegelman and Frost, 2017). Offline paradigms typically assess learning after exposure through explicit judgments requiring participants to consciously discriminate or recognize previously encountered structures. In contrast, online measures attempt to capture learning as it unfolds during exposure, often through neural entrainment, reaction times, target detection, eye movements, or electrophysiological responses. Importantly, online and offline measures frequently show weak correlations despite both revealing evidence of learning at the group level (Batterink et al., 2015; Franco et al., 2015). This pattern suggests that the behavioral expression of learning may depend not only on the extraction of regularities itself but also on additional processes such as memory retrieval, attentional control, explicit decision-making, and metacognitive access.

A growing number of neural entrainment and frequency-tagging studies have reinforced this interpretation. Neural frequency-tagging paradigms, including FPVS approaches, have repeatedly demonstrated sensitivity to structured regularities even in the absence of strong behavioral evidence of learning (Batterink and Paller, 2017; Moreau et al., 2022; Pinto et al., 2021). Pinto et al. (2021), for example, found that EEG frequency-tagging revealed reliable group-level neural discrimination of statistical structure despite limited correspondence with behavioral performance at the individual level. Similarly, Moreau et al. (2022) reported robust neural entrainment to auditory regularities in both children and adults despite relatively weak behavioral differences between groups. Such findings suggest that neural measures may capture forms of online sensitivity to regularities that are not necessarily accessible to explicit report.

Importantly, dissociations between neural and behavioral measures should not simply be interpreted as evidence of poor measurement sensitivity or task failure. Rather, several authors have argued that these dissociations may reflect the existence of partially independent learning components operating at different representational and processing levels (Batterink et al., 2015; Batterink and Paller, 2017; Ward et al., 2013). Batterink et al. (2015), for example, demonstrated that explicit recognition and implicit reaction-time measures of auditory segmentation could reveal distinct aspects of learning that did not necessarily correlate strongly with one another. Similarly, Ward et al. (2013) reported dissociations between neural repetition suppression and multivoxel pattern similarity, with different neural measures selectively

predicting implicit versus explicit expressions of memory. Together, these findings support the idea that learning can influence neural processing and implicit behavioral facilitation even when explicit conscious access to learned structure remains limited.

The distinction between reflection-based and processing-based measures discussed in the recent consensus paper on individual differences in SL is particularly relevant in this context (Bogaerts et al., 2026). Reflection-based tasks, such as 2AFC paradigms or familiarity judgments, require participants to consciously access and evaluate learned representations. Processing-based measures, by contrast, examine how learning influences ongoing behavior or neural activity without necessarily requiring explicit reflection on learned structure (Christiansen, 2019; Batterink and Paller, 2017). Importantly, these two classes of measures may rely on partially distinct cognitive and neural systems. Consequently, weak correlations between neural entrainment and explicit behavioral performance may not indicate absence of learning but differences in the level at which learning is being measured.

The findings of the present thesis align with this broader literature. Across both young and older adults, FPVS paradigms revealed robust neural sensitivity to structured auditory input despite relatively limited or variable behavioral evidence of learning. At the same time, behavioral cross-modal paradigms revealed learning effects that were not always accompanied by corresponding neural signatures. These patterns are broadly consistent with proposals suggesting that neural entrainment paradigms primarily capture early online sensitivity to statistical regularities and perceptual binding processes, whereas explicit behavioral tasks additionally recruit processes related to decision-making, memory retrieval, and metacognitive accessibility (Batterink et al., 2015; Batterink and Paller, 2017; Siegelman and Frost, 2017).

Recent work further suggests that even neural measures themselves may not reflect a single unified process. Different forms of neural entrainment, ERP responses, oscillatory synchronization, and repetition-based neural effects may index distinct aspects of regularity extraction and prediction (Schneider et al., 2022; Morris et al., 2025b). Schneider et al. (2022), for example, demonstrated dissociations between neural markers associated with distributional and transitional statistical learning, highlighting the possibility that different neural signatures reflect different computational components of learning. Similarly, Morris et al. (2025) reported that ERP indices of visual statistical learning varied substantially across individuals and were only weakly associated with explicit behavioral performance. Such findings reinforce the idea that neural-behavioral dissociations may emerge because different measures access different representational outputs of learning rather than because one measure is inherently more "valid" than another.

More broadly, the growing divergence between neural and behavioral findings

has important implications for how SL is conceptualized and measured. If different paradigms capture partially distinct aspects of learning, then low cross-task convergence may not necessarily undermine the existence of SL as a construct. Instead, it may indicate that SL involves multiple interacting processes differing in awareness, representational accessibility, temporal dynamics, and behavioral expression (Frost et al., 2015; Siegelman and Frost, 2017; Bogaerts et al., 2022b). From this perspective, neural and behavioral dissociations become theoretically informative rather than problematic. The present findings contribute to this interpretation by showing that online neural sensitivity to regularities can coexist with limited explicit behavioral access, particularly across different developmental stages and modalities. As discussed in the following sections, these dissociations provide important support for multicomponential and multi-level accounts of SL.

6.2 Theoretical and Methodological Implications

The empirical patterns discussed in the previous sections reveal a recurrent theme across contemporary statistical learning (SL) research: findings often vary substantially depending on the modality examined, the developmental stage considered, and the type of measure used to assess learning. Neural and behavioral indices do not always converge, cross-modal transfer is often partial and asymmetric, and inter-individual variability remains considerable across tasks and populations. While such dissociations have sometimes been interpreted as evidence of weak construct validity or unreliable measurement, an alternative possibility is that they reflect fundamental properties of SL itself. Increasingly, theoretical frameworks have proposed that SL may not constitute a single unified mechanism, but rather a multicomponential phenomenon emerging from multiple interacting processes differing in representational format, processing demands, and accessibility to conscious report (Frost et al., 2015; Siegelman and Frost, 2017; Conway, 2020; Bogaerts et al., 2022b).

The following sections examine the broader theoretical and methodological implications of this perspective. First, the discussion considers evidence supporting multicomponential accounts of SL and the idea that different tasks may recruit partially overlapping computations. The discussion then turns to the challenges posed by individual variability, task reliability, and construct validity, particularly in the context of neural and behavioral dissociations. Finally, the methodological implications of these issues are considered for future research on SL, including the need for multimethod approaches, improved measurement practices, and theoretically informed task selection.

Statistical Learning as a Multicomponential Construct

One of the central theoretical implications emerging from the present thesis concerns the nature of statistical learning (SL) itself. Historically, SL was often conceptualized as a relatively unitary mechanism allowing learners to extract regularities from sensory input across domains and modalities (Saffran et al., 1996a). Early demonstrations of learning across auditory, visual, tactile, and linguistic materials appeared to support the existence of a broadly domain-general learning system (Conway and Christiansen, 2005). However, accumulating evidence over the last decade has increasingly challenged this interpretation, suggesting instead that SL may involve multiple partially dissociable computations varying across modalities, tasks, representational formats, and levels of awareness (Frost et al., 2015; Siegelman and Frost, 2017; Conway, 2020; Bogaerts et al., 2022b). The findings of the present thesis are broadly consistent with this emerging perspective.

Several empirical tensions discussed throughout this thesis are difficult to reconcile with strong unitary accounts of SL. First, learning varied systematically across auditory and visual modalities, with cross-modal transfer often appearing partial, asymmetric, and task-dependent. Second, neural and behavioral measures frequently revealed weak correspondence, with neural entrainment sometimes emerging in the absence of robust explicit behavioral evidence and vice versa. Third, developmental and aging-related differences did not follow a single coherent trajectory across measures and modalities. Finally, substantial inter-individual variability emerged across tasks even within the same population. Taken together, these patterns suggest that SL cannot easily be reduced to a single homogeneous cognitive ability measured equivalently across paradigms.

The idea that SL may be multicomponential has gained substantial theoretical support in recent years. Frost et al. (2015) famously described the "paradox of statistical learning": SL appears sufficiently general to operate across domains and sensory modalities, yet sufficiently constrained to produce strong modality, stimulus, and task effects. To account for this tension, the authors proposed that SL may involve shared computational principles operating over modality-dependent representational systems. Similarly, Siegelman et al. (2017) argued that SL should be conceptualized as a multidimensional construct varying across modalities, domains, types of regularities, and task demands. Conway (2020) further proposed that different SL paradigms occupy distinct positions within a broader task space depending on the structure of the input, the nature of learning, and the degree of explicit awareness involved. Together, these frameworks increasingly move away from viewing SL as a single mechanism and toward conceptualizing it as a family of interacting learning processes.

The present findings align particularly well with frameworks emphasizing partially overlapping computations across tasks rather than complete independence or full equivalence between paradigms. For example, auditory and visual SL tasks may both involve sensitivity to transitional probabilities or chunking processes while differing substantially in perceptual encoding constraints, temporal dynamics, attentional requirements, and representational format (Conway and Christiansen, 2005; Frost et al., 2015; Zhou et al., 2024). Similarly, neural entrainment paradigms and explicit behavioral tasks may both reflect sensitivity to statistical structure while relying on different levels of representational accessibility and different supporting cognitive systems (Batterink et al., 2015; Batterink and Paller, 2017). Under this view, weak correlations across tasks do not necessarily indicate that SL does not exist as a construct. Rather, they may reflect the fact that different paradigms capture different components or outputs of learning.

Importantly, recent work on individual differences has further reinforced this interpretation. Reviews and consensus discussions increasingly emphasize that no single SL task is likely to provide full coverage of the construct (Bogaerts et al., 2022b, 2026). Canonical paradigms typically involve highly simplified regularities presented within specific modalities and domains, often emphasizing only one type of statistical structure, such as deterministic transitional probabilities (Isbilen and Christiansen, 2022). As argued in the recent consensus paper on individual differences in SL, the common assumption that performance on a single task reflects a general SL ability may therefore be overly simplistic (Bogaerts et al., 2026). Instead, different tasks may recruit distinct combinations of perceptual, attentional, mnemonic, predictive, and metacognitive processes. The present findings are consistent with this view, showing limited convergence across modalities and across neural versus behavioral measures despite evidence of learning within each paradigm.

The distinction between processing-based and reflection-based measures is particularly relevant in this context (Christiansen, 2019; Bogaerts et al., 2026). Processing-based paradigms assess how learning influences ongoing neural activity or behavior during exposure, whereas reflection-based tasks require conscious access to learned representations after learning has occurred. These two forms of measurement may therefore rely on partially distinct systems and may not necessarily converge strongly at the individual level. Importantly, such dissociations are increasingly viewed not as methodological failures but as evidence that SL unfolds across multiple representational and processing levels (Batterink et al., 2015; Batterink and Paller, 2017).

Relatedly, Thiessen et al. (2013) proposed the extraction-integration framework, according to which SL involves at least two interacting components: the extraction of statistical regularities from input and the integration of these regularities into

broader representational structures. This framework resonates strongly with the present findings. Neural entrainment paradigms may primarily capture early extraction and perceptual binding processes, whereas explicit behavioral tasks may additionally depend on integration, consolidation, memory retrieval, and metacognitive accessibility. Similar distinctions have also emerged in work emphasizing chunking and predictive facilitation as central mechanisms underlying SL (Isbilen and Christiansen, 2022; Dale et al., 2012). From this perspective, different tasks may not simply vary in sensitivity but may access qualitatively different outputs of learning.

Importantly, multicomponentiality does not imply that SL should be fragmented into entirely unrelated systems. The available evidence instead suggests the existence of partially shared learning principles interacting with modality-specific constraints, representational systems, and task-dependent processing demands (Frost et al., 2015; Conway, 2020). In this sense, the present findings support an intermediate position between strong domain-general and strong modality-specific accounts. SL may best be understood as a family of related computations that operate across modalities and domains while producing distinct behavioral and neural signatures depending on how learning is measured and expressed.

Overall, the empirical patterns emerging from the present thesis contribute to growing evidence that SL is unlikely to reflect a single unified mechanism. Rather, the findings support contemporary perspectives proposing that SL emerges from multiple interacting processes differing in modality dependence, representational accessibility, temporal dynamics, and cognitive demands. This interpretation provides a coherent framework for understanding why modality dissociations, developmental variability, and neural-behavioral divergences are so commonly observed across the literature. As discussed in the following sections, it also has important implications for how individual differences are interpreted and how SL should be measured methodologically.

Individual Differences and the Problem of Measurement

The growing recognition that statistical learning (SL) may be multicomponential has important implications for how individual variability is interpreted and measured. Early SL research primarily focused on demonstrating robust group-level effects, often treating inter-individual variability as experimental noise to be minimized (Frost et al., 2019). Over the last decade, however, increasing attention has been devoted to the study of individual differences in SL, motivated by the possibility that variability across learners may provide important insight into the nature of learning mechanisms and their relation to broader cognitive abilities (Siegelman and Frost,

2015; Kidd et al., 2018; Bogaerts et al., 2022b). At the same time, this emerging literature has revealed substantial methodological and theoretical challenges concerning construct validity, task reliability, and the interpretation of weak correlations across measures.

One important shift emphasized in recent theoretical discussions is the idea that variability in SL should not automatically be regarded as noise. Studies have shown that individual differences in SL can predict variability in language acquisition, reading, speech segmentation, and other cognitive domains (Arciuli and Simpson, 2012; Kidd, 2012; Frost et al., 2013b). Such findings motivated the proposal that SL may constitute an important individual cognitive ability (Siegelman and Frost, 2015). However, subsequent work revealed that correlations between SL and other cognitive functions are often modest, inconsistent, and highly dependent on the specific tasks employed (Pavlidou and Bogaerts, 2019; Schmalz et al., 2019; Boeve et al., 2024; Ren et al., 2023). These mixed findings have raised fundamental questions concerning what SL tasks actually measure and whether different paradigms truly assess the same underlying construct.

The present thesis aligns closely with these ongoing debates. Across modalities, developmental stages, and measurement types, substantial variability emerged both within and across tasks. Importantly, such variability was not random but appeared systematically related to modality constraints, task demands, and the distinction between neural and behavioral expressions of learning. These findings support recent arguments suggesting that individual differences in SL may reflect genuine heterogeneity in the computations and representational systems recruited by different paradigms rather than merely reflecting measurement error (Bogaerts et al., 2022b, 2026).

A major challenge in this context concerns construct validity. As emphasized in the recent consensus paper on individual differences in SL, the field still lacks a universally agreed-upon definition of SL itself (Bogaerts et al., 2026). Definitions range from narrow accounts emphasizing transitional probability tracking to broader frameworks encompassing chunking, prediction, procedural learning, and unsupervised extraction of environmental regularities (Christiansen, 2019; Thiessen et al., 2013; Conway, 2020). This theoretical ambiguity creates substantial difficulties for interpreting task performance. If different paradigms recruit partially distinct computations, then low convergence across tasks may not necessarily indicate poor measurement quality but insufficient overlap in the underlying processes being measured.

This issue is particularly evident when comparing different classes of SL measures. Reflection-based tasks, such as 2AFC paradigms and familiarity judgments, require

conscious access to learned representations and explicit decision-making. Processing-based measures, by contrast, assess how learning influences ongoing behavior or neural activity without requiring explicit reflection (Batterink and Paller, 2017; Christiansen, 2019; Bogaerts et al., 2026). These distinctions are critical because tasks differing in representational accessibility and cognitive demands may not be expected to correlate strongly even if they all reflect sensitivity to statistical structure. The present findings, particularly the dissociations observed between neural entrainment and explicit behavioral performance, are consistent with this interpretation.

Relatedly, recent work has increasingly emphasized the problem of task impurity in SL research. No task measures SL in isolation. Performance in SL paradigms is shaped not only by sensitivity to regularities but also by attention, working memory, perceptual encoding, executive control, memory retrieval, and metacognitive reflection (Janacsek and Nemeth, 2013; Lukács et al., 2024; Pedraza et al., 2024). Consequently, observed correlations between SL tasks and cognitive outcomes may sometimes reflect overlapping demands on these broader cognitive systems rather than SL-specific computations. This concern becomes particularly relevant in developmental and aging research, where changes in attention, memory, or sensory processing may strongly influence task performance independently of learning itself (Forest et al., 2023; Arnon, 2020).

The issue of reliability further complicates interpretation. Hedge et al. (2018) described the "reliability paradox," noting that paradigms producing highly robust group-level effects often show poor reliability for measuring individual differences. This paradox has become particularly relevant in SL research, where many canonical tasks were originally designed to demonstrate the existence of learning at the group level rather than to quantify stable inter-individual variability (Siegelman and Frost, 2017; Bogaerts et al., 2026). As a result, many commonly used SL paradigms exhibit weak internal consistency or test-retest reliability, limiting their suitability for correlational analyses and individual differences research (Arnon, 2020; Kalra et al., 2019). Importantly, low reliability attenuates observed correlations and may contribute substantially to inconsistent findings across studies (Hedge et al., 2018; Brysbaert, 2024).

Neural measures introduce additional methodological complexities. Although neural entrainment and FPVS paradigms often reveal robust group-level sensitivity to statistical structure, their sensitivity at the individual level can vary considerably (Pinto et al., 2021). Furthermore, neural signatures themselves may reflect multiple overlapping processes rather than a single unitary index of learning (Schneider et al., 2022; Morris et al., 2025b). Consequently, weak neural-behavioral correlations should not necessarily be interpreted as evidence that one measure failed to detect learning.

Rather, such dissociations may reflect the fact that neural and behavioral paradigms access different representational outputs or processing stages of learning.

The present findings broadly support this interpretation. Weak convergence across modalities and measures emerged alongside clear evidence of learning within individual paradigms. Such patterns are difficult to reconcile with views assuming that all SL tasks measure a single common ability. Instead, they are more consistent with multicomponential frameworks proposing that different tasks recruit partially overlapping combinations of perceptual, predictive, mnemonic, attentional, and metacognitive processes (Frost et al., 2015; Siegelman and Frost, 2017; Bogaerts et al., 2022b). Under this view, weak correlations across tasks may partly reflect genuine heterogeneity in learning processes rather than simple measurement failure.

Importantly, this interpretation does not imply that methodological rigor is unimportant. On the contrary, the present findings reinforce recent recommendations emphasizing the need for explicit theoretical definitions, careful task selection, reliability reporting, and multimethod approaches in SL research (Bogaerts et al., 2026). If SL is indeed multicomponential, then understanding individual variability will require moving beyond the assumption that a single paradigm can adequately capture the construct as a whole. Instead, future progress may depend on identifying how different tasks map onto distinct learning components and representational levels.

Methodological Implications for Statistical Learning Research

Beyond their theoretical implications, the findings of the present thesis also carry broader methodological consequences for the study of statistical learning (SL). Across chapters, substantial variability emerged depending on the modality examined, the population tested, and, critically, the type of measure used to assess learning. Neural and behavioral indices did not always converge, cross-modal transfer was often partial rather than complete, and individual differences varied considerably across tasks and developmental stages. Together, these findings reinforce a growing consensus in the field that no single paradigm is likely to provide a complete characterization of SL (Frost et al., 2019; Bogaerts et al., 2022b, 2026). Instead, progress in understanding SL may depend on integrating multiple complementary measures capable of capturing distinct components and representational levels of learning.

One major implication concerns the importance of multimethod approaches. Historically, many SL studies relied primarily on single offline behavioral tasks, particularly two-alternative forced-choice paradigms. While such measures have played a foundational role in demonstrating sensitivity to statistical structure, they provide only a limited window onto learning processes and may underestimate forms

of knowledge that remain inaccessible to conscious report (Batterink et al., 2015; Batterink and Paller, 2017). The present findings, together with previous evidence from neural entrainment, target detection, priming, and eye-tracking studies, suggest that different measures may reveal different expressions of learning rather than redundant indices of the same phenomenon (Lukics and Lukács, 2021; Batterink and Paller, 2017; Schneider et al., 2022). Combining neural and behavioral approaches therefore appears particularly important for capturing the multidimensional nature of SL.

More specifically, the present thesis highlights the value of online measures capable of tracking learning as it unfolds over time. Neural entrainment and FPVS paradigms revealed sensitivity to structured regularities even in situations where explicit behavioral evidence was weak or inconsistent. Similar conclusions have emerged from ERP, reaction-time, pupillometry, and eye-tracking studies demonstrating that online processing measures often detect learning processes that are not fully reflected in offline judgments (Batterink and Paller, 2017; Franco et al., 2015; Soares et al., 2020). Such findings support recent arguments that online and processing-based measures may provide improved access to implicit learning dynamics compared to traditional reflection-based tasks (Bogaerts et al., 2026). Importantly, however, online measures should not be viewed as inherently superior to offline tasks. Rather, different paradigms likely provide complementary information about distinct stages and representational outputs of learning.

The findings also reinforce the need for greater theoretical precision regarding what SL tasks are assumed to measure. As emphasized throughout the recent consensus literature, many canonical SL paradigms were originally developed to demonstrate group-level learning effects rather than to isolate specific computations or quantify stable individual differences (Siegelman and Frost, 2017; Bogaerts et al., 2026). Consequently, tasks often differ substantially in modality, stimulus structure, attentional demands, memory requirements, and levels of explicit access. Without clear theoretical specification, weak convergence across tasks may easily be misinterpreted either as evidence against SL as a construct or as simple measurement failure. The present findings instead support the view that different paradigms recruit partially overlapping learning components that vary in representational accessibility, processing demands, and domain specificity.

Relatedly, the present results further underscore the importance of reliability reporting and psychometric evaluation in SL research. The reliability paradox described by Hedge et al. (2018) is particularly relevant for SL paradigms, many of which produce robust group-level effects while showing limited reliability for individual differences analyses. Recent discussions have therefore emphasized the need to

systematically evaluate the psychometric properties of SL tasks, especially when correlational or developmental questions are central (Arnon, 2020; Brysbaert, 2024; Bogaerts et al., 2026). The variability observed across the present studies highlights how reliability, construct validity, and task impurity may substantially influence conclusions regarding modality effects, developmental trajectories, and neural-behavioral convergence.

Another important implication concerns the potential value of latent-variable and multi-task approaches. If SL reflects multiple partially dissociable computations, then single-task designs may provide only a fragmentary view of the underlying construct. Recent work has therefore advocated for large-scale studies combining multiple paradigms within the same participants to identify shared and unique sources of variance across tasks (Growns et al., 2020b; Zhou et al., 2024; Siegelman and Frost, 2015). Such approaches may help clarify which aspects of learning generalize across modalities and domains and which remain task- or representation-specific. The present findings strongly support this direction, particularly given the limited convergence observed across modalities and measures.

The thesis also speaks to ongoing debates regarding ecological validity in SL research. While highly controlled canonical paradigms remain essential for isolating specific learning mechanisms, they may capture only a subset of the computations involved in real-world learning environments (Frost et al., 2019; Bogaerts et al., 2026). More ecological paradigms examining sensitivity to naturally occurring regularities in language, reading, or complex sensory environments may therefore provide important complementary insight (Isbilen et al., 2022; Siegelman, 2020b; Nastase et al., 2020). At the same time, ecological paradigms introduce additional challenges related to attentional variability, prior knowledge, and uncontrolled exposure history. Future work will likely benefit from integrating controlled experimental paradigms with more naturalistic approaches capable of capturing learning under ecologically relevant conditions.

Overall, the present thesis supports the growing view that methodological progress in SL research may depend less on identifying a single "best" task and more on understanding how different paradigms map onto distinct components and representational levels of learning. Rather than treating variability across measures as a problem to eliminate, future research may benefit from leveraging such variability to better characterize the multidimensional nature of SL itself. Integrating neural and behavioral approaches, combining online and offline measures, improving reliability and construct validity, and adopting multimethod frameworks may therefore represent crucial steps toward a more comprehensive understanding of how humans extract and represent statistical structure across modalities and across the lifespan.

6.3 Toward a Multi-Level Framework of Statistical Learning

The preceding sections argued that statistical learning (SL) is unlikely to reflect a single, unitary cognitive ability and is better characterized as a multicomponential phenomenon. However, multicomponentiality alone does not explain why some measures converge whereas others dissociate, nor does it specify how different expressions of learning relate to one another. Across modalities, developmental stages, and measurement approaches, the literature reviewed throughout this thesis consistently revealed patterns of partial convergence and dissociation. Neural and behavioral measures often showed limited correspondence, learning effects varied across sensory modalities and age groups, and substantial individual variability emerged depending on the task used to assess learning. Rather than representing isolated inconsistencies, these findings suggest that different paradigms may provide access to distinct expressions of learning that vary in their processing demands, representational format, and accessibility to conscious awareness.

Recent theoretical accounts increasingly support such a perspective. Contemporary frameworks propose that SL involves multiple interacting computations distributed across domains, modalities, and representational systems rather than a single domain-general mechanism (Frost et al., 2015; Siegelman and Frost, 2017; Conway, 2020; Bogaerts et al., 2022b). Similarly, the recent consensus paper on individual differences in SL emphasized the distinction between processing-based and reflection-based measures, arguing that different tasks likely access different aspects of learning (Bogaerts et al., 2026). Together, these developments suggest that understanding SL may require moving beyond the question of whether learning occurred and instead considering how learned information is represented and expressed across different measurement contexts.

The aim of the present section is to develop this idea further by proposing a multi-level framework of SL. Rather than assuming that all measures provide equivalent access to a single underlying representation, the framework organizes common SL paradigms according to the degree of representational accessibility and processing demands they may involve. To motivate this proposal, the following sections review evidence for dissociations between neural and behavioral measures, between implicit and explicit behavioral expressions of learning, and between explicit recognition and controlled recall or production tasks. The present thesis provides examples of dissociations at each of these levels. Chapters 2 and 3 revealed robust neural entrainment despite relatively weak behavioral correspondence, whereas Chapters 4

and 5 demonstrated that learning and transfer effects varied across tasks, modalities, and developmental stages. Together with evidence from the broader literature, these recurring dissociations form the basis for a conceptual framework in which different measures may access partially overlapping representational outputs of learning operating at different levels of accessibility to conscious awareness.

Neural–Behavioural Dissociations

One of the most consistent patterns emerging from the statistical learning (SL) literature concerns the often limited correspondence between neural and behavioural measures of learning. Although both types of measures are frequently assumed to reflect the same underlying acquisition of regularities, a growing body of work suggests that neural sensitivity to statistical structure and behavioural expressions of learning do not always converge. Across EEG, MEG, and fMRI studies, neural indices often reveal robust learning-related responses even when behavioural evidence is weak, absent, or only partially related to neural performance. These findings raise important questions regarding what different measures capture and whether they provide access to the same underlying representations.

Evidence from EEG studies has been particularly influential in highlighting such dissociations. Using an auditory SL paradigm, Batterink et al. (2015) demonstrated that participants exhibited reaction-time facilitation to statistically predictable stimuli despite showing limited explicit recognition of the learned regularities. Similarly, Batterink and Paller (2017) showed that neural entrainment tracked the online emergence of statistical structure during exposure, whereas explicit behavioural measures were weaker and more variable. More recently, Pinto et al. (2021) reported robust frequency-tagging effects at the group level despite weak correspondence with behavioural measures and limited sensitivity at the individual level. Likewise, Morris et al. (2025a) found learning-related ERP responses even among participants who showed little explicit behavioural evidence of learning. Together, these findings suggest that neural measures may reveal sensitivity to statistical structure that is not necessarily expressed in conscious behavioural reports.

Similar patterns have been observed using MEG. Moser et al. (2021a) demonstrated that neural entrainment to structured tone sequences emerged early during exposure and increased over time, whereas the emergence of explicit knowledge varied substantially across participants. Neural tracking therefore appeared to reflect an earlier and more ubiquitous expression of learning than explicit behavioural performance. Relatedly, Maheu et al. (2019) showed that neural signatures of sequence learning could be detected across multiple timescales before learners demonstrated

consistent behavioural evidence of acquired knowledge.

Evidence from fMRI further supports this interpretation. In a seminal study, Turk-Browne et al. (2009) demonstrated learning-related activity in the striatum and medial temporal lobe during exposure to visual statistical regularities despite limited awareness of the underlying structure. Similarly, Rosenthal et al. (2016) reported learning-related responses within primary visual cortex and hippocampal networks during the acquisition and recognition of non-conscious visuospatial sequences. Together, these findings suggest that learning-related neural representations can emerge even when conscious behavioural access remains limited.

The findings of the present thesis are broadly consistent with this literature. In Chapters 2 and 3, neural entrainment revealed reliable sensitivity to statistically structured auditory input in both young and older adults, whereas behavioural performance showed greater variability and only limited correspondence with neural measures. Importantly, similar patterns have now been observed across EEG, MEG, and fMRI methodologies, suggesting that neural–behavioural dissociations are not isolated anomalies but a recurring feature of SL research.

Taken together, the available evidence suggests that neural measures often capture learning that is stronger, earlier, or qualitatively different from what is expressed behaviourally. Rather than indicating measurement failure, these dissociations may reflect the fact that neural and behavioural paradigms access different expressions of learned structure. This possibility provides the first motivation for distinguishing between early implicit neural sensitivity and later behavioural manifestations of learning, a distinction that forms a central component of the multi-level framework proposed below.

Implicit and Explicit Behavioural Expressions of Learning

Dissociations in the SL literature are not limited to comparisons between neural and behavioural measures. Even among behavioural tasks, evidence increasingly suggests that different paradigms may capture distinct expressions of learning. In particular, studies comparing online, processing-based measures with explicit judgment tasks often report only partial convergence, indicating that behavioural evidence of learning is itself heterogeneous.

One of the clearest examples comes from reaction-time paradigms. Using an auditory SL task, Batterink et al. (2015) found reliable reaction-time facilitation for statistically predictable elements despite relatively weak explicit recognition of the learned structure. Such findings suggest that statistical regularities can influence ongoing processing even when learners have limited conscious access to what has

been acquired.

Evidence from eye-tracking studies points in a similar direction. Zolnai et al. (2022) showed that anticipatory eye movements detected statistical regularities earlier and more sensitively than reaction times in a probabilistic sequence-learning task. Learning was evident in gaze behaviour even when corresponding behavioural measures showed weaker effects, suggesting that eye movements may provide access to implicit predictive knowledge not fully captured by conventional behavioural tasks. More broadly, Bays et al. (2015) demonstrated that visual statistical learning can produce dissociable outcomes across different behavioural measures, including detection, search, recognition, and associative tasks, indicating that no single behavioural measure provides a complete account of learning. Pupillometry provides further evidence that different behavioural measures may not always capture the same expression of learning. In a visual statistical learning paradigm, Zhang and Emberson (2019) found limited evidence that pupillary responses tracked learned regularities and argued that methodological constraints may severely limit the utility of pupillometry for assessing visual statistical learning.

Taken together, these findings suggest that implicit behavioural measures often capture consequences of learning that are not fully reflected in explicit judgments or recognition tasks. This conclusion is consistent with the distinction between processing-based and reflection-based measures emphasized in recent theoretical accounts of SL (Bogaerts et al., 2026). Whereas processing-based measures assess how learned regularities influence ongoing perception, attention, and behaviour, explicit measures require learners to consciously evaluate or report aspects of their acquired knowledge. Consequently, these paradigms may differ not only in sensitivity but also in the forms of representational access they require.

Importantly, the existence of dissociations between implicit and explicit behavioural measures does not imply that one type of task provides a more valid assessment of SL than the other. Rather, it suggests that different tasks may access different expressions of learning. Some paradigms may primarily capture facilitated processing and prediction-based behaviour, whereas others require conscious evaluation of learned structure. This distinction provides a second motivation for the framework proposed below and supports the idea that behavioural measures themselves may be organized according to differing levels of representational accessibility.

Explicit Recognition and Controlled Retrieval

A further distinction emerges within measures that are typically classified as explicit assessments of statistical learning. Although tasks such as familiarity judgments,

recognition ratings, and two-alternative forced-choice (2AFC) tests all require some degree of conscious access to learned structure, they may differ substantially from tasks that require participants to actively retrieve, generate, or reproduce learned information. Consequently, explicit measures themselves may not constitute a homogeneous category.

Recognition-based tasks are among the most commonly used measures in the SL literature. These paradigms require participants to evaluate presented stimuli and decide whether they appear familiar or consistent with previously encountered regularities. While such judgments necessitate some conscious access to learned representations, they can often be performed on the basis of familiarity or fluency without requiring detailed retrieval of the learned structure itself (Batterink et al., 2015; Bogaerts et al., 2026). As a result, recognition tasks may provide evidence that learned representations have become accessible to awareness while placing relatively limited demands on retrieval processes.

In contrast, recall, production, generation, and reproduction tasks require participants to actively reconstruct or produce information acquired during learning. Such paradigms place greater demands on memory retrieval, cognitive control, and the flexible use of learned representations. Evidence from the implicit learning literature suggests that successful recognition and successful production do not always coincide. Participants may demonstrate sensitivity to learned structure in recognition tasks while showing substantially weaker performance when asked to reproduce or generate that structure (Reber, 1989). Similarly, Destrebecqz et al. (2005) reported dissociations between implicit and explicit expressions of sequence knowledge, indicating that different tasks can reveal different forms of access to acquired representations.

Importantly, the distinction between recognition and controlled retrieval should not be interpreted as a categorical division between implicit and explicit learning. Rather, these tasks may differ in the degree of representational accessibility and cognitive control they require. Recognition tasks occupy an intermediate position in which learned representations become available for conscious evaluation, whereas recall and production tasks require deliberate retrieval and active deployment of those representations. Consequently, they may provide access to different expressions of the same underlying learning process.

Taken together, the available evidence suggests that explicit measures vary in the extent to which they require active access to learned representations. This distinction provides a third motivation for the framework proposed below and supports the idea that representational accessibility may vary even within tasks traditionally classified as explicit measures of statistical learning.

6.3.1 A Multi-Level Framework of Statistical Learning

Taken together, the evidence reviewed above suggests that dissociations in the statistical learning (SL) literature are neither isolated findings nor methodological anomalies. Rather, they appear to emerge systematically across different classes of measures. Neural indices do not always converge with behavioural performance, implicit behavioural measures frequently diverge from explicit judgments, and recognition-based tasks may not require the same degree of representational access as recall or production tasks. Viewed collectively, these observations suggest that different SL paradigms may access partially overlapping expressions of learning rather than a single, fully shared representation.

One possible interpretation of these findings is that measures differ in the degree of representational accessibility they require and in the processing operations involved in their expression. Figure 6.1 summarizes this proposal. Importantly, the framework should not be interpreted as a strict hierarchical sequence through which all learners necessarily progress, nor as a definitive architecture of SL. Rather, it is intended as a heuristic framework for organizing recurring dissociations observed across the literature and for conceptualizing how different measures may provide access to different expressions of learned structure.

At the lowest level of representational accessibility, early neural measures such as EEG, MEG, and fMRI may primarily capture online sensitivity to statistical regularities and the emergence of learning-related neural representations. These measures often reveal sensitivity to structured input even when behavioural evidence remains weak or absent, suggesting that learning can influence neural processing before it becomes available for behavioural report. At a second level, implicit behavioural measures such as reaction times, eye-tracking, and pupillometry may reflect facilitated processing, predictive allocation of attention, and increased perceptual fluency resulting from acquired regularities. Although these measures provide behavioural evidence of learning, they often do so without requiring participants to consciously evaluate or report the learned structure.

A further level is represented by explicit behavioural measures, including familiarity judgments, ratings, and two-alternative forced-choice tasks. Such paradigms require learners to consciously access and evaluate learned representations, but may still rely heavily on familiarity-based processes rather than deliberate retrieval. Finally, recall and production tasks may occupy a level characterized by the highest degree of representational accessibility, as they require learners to intentionally retrieve, reconstruct, or generate information derived from previous learning experiences. These tasks therefore place greater demands on memory retrieval, cognitive

control, and the flexible use of acquired representations.

Importantly, the proposed framework does not imply that these levels correspond to independent systems or discrete processing stages. Different measures likely recruit overlapping computations, and many tasks contain both implicit and explicit components. Moreover, factors such as attention, memory, motivation, task demands, and prior knowledge influence performance across all levels. Consequently, the framework should not be viewed as a taxonomy of learning mechanisms, but rather as a conceptual tool for understanding why different measures may yield partially divergent patterns of results despite all being sensitive to statistical structure.

A key advantage of this framework is that it provides a principled way of interpreting recurring dissociations across SL paradigms without assuming that some measures are inherently more valid than others. Instead, apparent inconsistencies may arise because different tasks access different expressions of learning. Neural measures may reveal the emergence of sensitivity to regularities, implicit behavioural measures may capture the influence of learning on ongoing processing, explicit judgments may reflect conscious access to learned representations, and recall or production tasks may index the ability to deliberately retrieve and flexibly use acquired knowledge.

Furthermore, the framework is compatible with contemporary views that statistical learning is inherently multicomponential. Rather than assuming a single endpoint of learning, it suggests that learning may generate multiple partially overlapping representational outcomes that differ in their accessibility and behavioural expression. From this perspective, the diversity of findings observed across the SL literature may be understood not as evidence against learning, but as evidence that learning can be expressed in multiple ways depending on how it is measured.

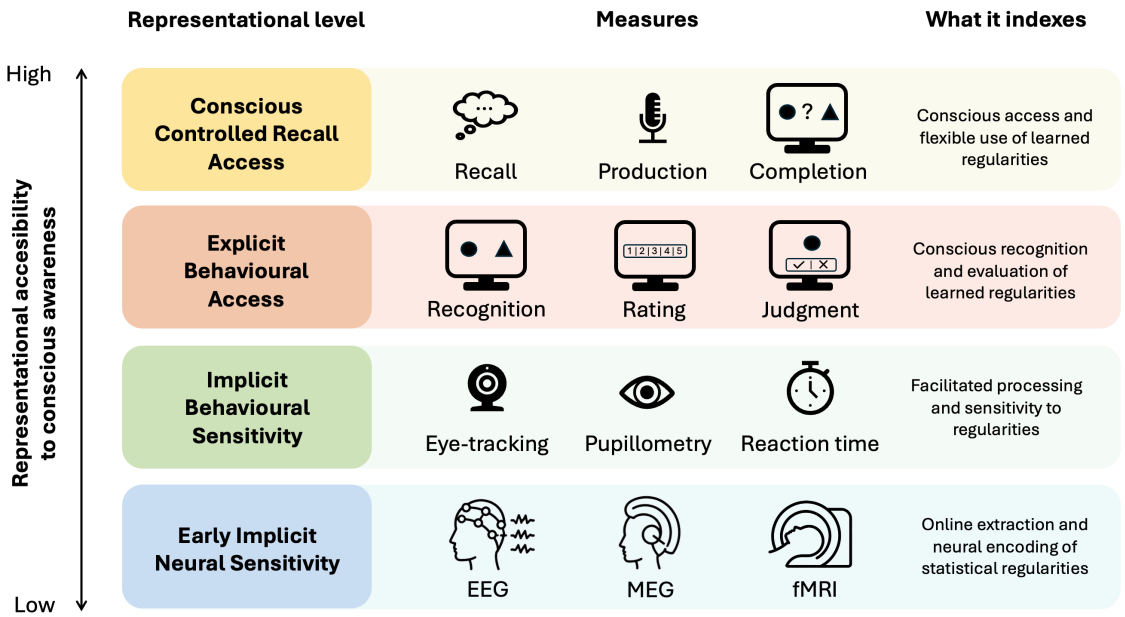


Figure 6.1: **Proposed multi-level framework of Statistical Learning.** The framework organizes commonly used statistical learning measures according to their presumed level of representational accessibility to conscious awareness. Early implicit neural measures (e.g., EEG, MEG, fMRI) are proposed to index online extraction, encoding, and perceptual binding of statistical regularities. Implicit behavioural measures (e.g., eye-tracking, pupillometry, reaction times) are proposed to reflect facilitated processing, predictive effects, and sensitivity to learned structure. Explicit behavioural measures (e.g., recognition, judgments, and ratings) require conscious access to learned representations and their evaluation. Finally, recall, production, and completion tasks involve controlled retrieval, reconstruction, and flexible use of learned representations. Importantly, the framework is intended as a conceptual heuristic for interpreting recurring dissociations across statistical learning paradigms rather than as a strict hierarchical model. Measures located at similar levels may show greater convergence because they rely on more similar forms of representational access, although such convergence is not assumed to be deterministic.

Relations Within and Across Representational Levels

The proposed framework may also help explain why correlations between statistical learning (SL) measures are often weak or inconsistent. If different paradigms access partially distinct expressions of learning, stronger convergence may be expected among measures that rely on similar forms of representational access than among measures that differ substantially in their processing demands and accessibility requirements. Importantly, this should not be interpreted as a deterministic prediction that measures located closer within the framework must necessarily correlate more

strongly. Rather, representational proximity may increase the probability of convergence because such measures are more likely to rely on overlapping computations and forms of access to learned structure.

Several findings reviewed throughout this thesis are consistent with this interpretation. Neural and behavioural measures frequently show limited correspondence, with learning-related neural responses often emerging in the absence of robust behavioural evidence (Batterink and Paller, 2017; Pinto et al., 2021; Moser et al., 2021a; Morris et al., 2025a). Similarly, implicit behavioural measures do not always converge with explicit recognition tasks. For example, Batterink et al. (2015) reported reaction-time facilitation despite weak explicit recognition, while Zhang and Emberson (2019) found limited correspondence between pupillary indices of learning and offline behavioural measures. These findings suggest that measures situated at different levels of representational accessibility may often capture distinct expressions of acquired knowledge.

At the same time, evidence also suggests that convergence can emerge among measures operating at similar levels. Studies employing multiple neural measures have often reported convergent evidence for sensitivity to statistical structure despite weak correspondence with behavioural performance (Maheu et al., 2019; Moser et al., 2021a). Such findings are consistent with the possibility that convergence may be more likely among measures relying on similar forms of representational access than among measures situated at more distant levels of the framework.

Importantly, the present proposal does not imply that representational accessibility is the sole determinant of convergence. Task demands, attentional allocation, memory requirements, stimulus characteristics, developmental factors, and individual differences are all likely to influence relations among measures. Nevertheless, organizing paradigms according to representational accessibility may provide a useful heuristic for understanding why some measures converge more reliably than others.

From this perspective, weak cross-task correlations should not necessarily be interpreted as evidence against SL as a coherent construct. Instead, they may partly reflect the fact that different tasks capture different expressions of learning. If representational accessibility constitutes a meaningful organizing principle, convergence should be more likely among measures relying on similar forms of access to learned structure than among measures that differ substantially in their processing demands and accessibility requirements. This interpretation offers a potential explanation for the recurring pattern of modest and inconsistent correlations observed throughout the SL literature.

Implications for Theories of Statistical Learning

The framework proposed here carries several broader implications for theories of statistical learning. First, it reinforces the view that SL is unlikely to reflect a single unified cognitive ability. Rather, the available evidence suggests that learning emerges through the interaction of multiple partially dissociable processes that vary in their modality dependence, representational format, temporal dynamics, and accessibility to awareness (Frost et al., 2015; Siegelman and Frost, 2017; Conway, 2020; Bogaerts et al., 2022b). In this respect, the present proposal extends existing multicomponential accounts by focusing not only on the mechanisms underlying learning, but also on the different ways in which learning may subsequently be expressed and measured.

Second, the framework offers a possible explanation for several longstanding debates within the SL literature. Questions concerning modality specificity, neural–behavioural convergence, developmental trajectories, and individual differences have often been addressed independently. However, many of these debates may reflect a common underlying issue: different studies frequently operationalize learning using measures that require different forms of representational access. Neural entrainment paradigms, online behavioural measures, explicit recognition tasks, and recall or production paradigms may therefore not constitute competing indices of a single outcome, but rather complementary windows onto different expressions of learning.

Third, the framework is compatible with contemporary theoretical perspectives emphasizing that learning may generate multiple representational consequences rather than a single stable endpoint. Statistical regularities may influence online neural processing, facilitate perceptual and attentional operations, support explicit recognition, and eventually become available for deliberate retrieval and flexible use. These expressions of learning may be related, but they need not emerge simultaneously nor be equally accessible across individuals, tasks, or developmental stages.

Finally, the present proposal has important methodological implications. If different measures access different expressions of learning, progress in SL research may depend less on identifying a single optimal task and more on understanding how different paradigms relate to one another. Integrating neural, implicit behavioural, explicit recognition, and controlled retrieval measures within the same experimental designs may therefore provide a more comprehensive account of how statistical regularities are extracted, represented, generalized, and made available for conscious report. Such an approach may be particularly important for understanding individual differences, developmental change, and the conditions under which learning transfers across modalities and domains.

Overall, the framework proposed here suggests that the diversity of findings ob-

served across the SL literature may not simply reflect measurement noise or methodological shortcomings. Instead, it may reflect the inherently multi-level nature of learning itself. Understanding how different measures map onto different expressions of acquired knowledge may therefore represent an important step toward a more comprehensive theory of statistical learning.

6.4 Limitations and Future Directions

Several limitations of the present thesis should be acknowledged. First, although the thesis adopted a multimethod and lifespan-oriented approach to statistical learning (SL), sample sizes within individual studies remained relatively modest, particularly for analyses focusing on inter-individual variability and neural-behavioral correlations. Recent methodological work has emphasized that studies investigating individual differences in SL require substantially larger samples than traditional group-level paradigms due to the combined effects of measurement noise, low task reliability, and attenuated correlations across tasks (Hedge et al., 2018; Brysbaert, 2024; Bogaerts et al., 2026). Consequently, some weak or absent correlations observed across measures in the present work should be interpreted cautiously, as they may partly reflect psychometric limitations in addition to genuine theoretical dissociations.

Second, the tasks employed throughout the thesis differ substantially in their processing demands, representational formats, and levels of explicitness. While this diversity was theoretically informative and allowed the investigation of multiple aspects of SL, it also complicates direct comparisons across paradigms. As highlighted in recent theoretical and methodological discussions, no single SL task can be assumed to provide exhaustive coverage of the construct (Siegelman and Frost, 2017; Bogaerts et al., 2022b; Frost et al., 2019). Behavioral tasks may recruit attentional control, memory retrieval, and metacognitive reflection in addition to learning itself, whereas neural measures may capture online sensitivity to regularities without necessarily indexing consciously accessible knowledge (Batterink et al., 2015; Batterink and Paller, 2017). From this perspective, weak convergence across measures may reflect both task impurity and genuine multicomponentiality.

Third, the present studies relied primarily on cross-sectional comparisons across age groups rather than longitudinal designs. Although the findings provide important insight into developmental and aging-related differences in SL, cross-sectional approaches cannot fully disentangle maturational changes from cohort effects or individual developmental trajectories. Longitudinal work following learners across child-

hood, adolescence, and aging would provide stronger evidence regarding how modality weighting, neural sensitivity, and explicit accessibility evolve over time (Forest et al., 2023; Kidd et al., 2018; Arnon, 2020).

Additional limitations concern the ecological validity of the paradigms employed. Most tasks used highly controlled artificial regularities designed to isolate specific learning computations. While such paradigms provide strong experimental control, they only partially approximate the complexity of naturalistic learning environments, where regularities are probabilistic, multimodal, hierarchical, and embedded within meaningful contexts (Frost et al., 2019; Nastase et al., 2020). Future research would benefit from integrating more ecological paradigms while maintaining sufficient experimental precision to isolate the computations underlying learning.

Several important directions for future research emerge from the present findings. First, future work should adopt latent-variable and multilevel modeling approaches capable of identifying shared and unique variance across different SL measures (Siegelman and Frost, 2017; Growns et al., 2020a; Kidd et al., 2023). Such approaches may help clarify whether different tasks reflect common underlying computations or largely distinct learning components. Large-scale collaborative datasets combining neural, behavioral, and developmental measures may be particularly valuable in this regard (Bogaerts et al., 2026).

Second, future studies should continue integrating multiple neural methodologies to better characterize the representational and temporal dynamics of SL. While EEG frequency-tagging and neural entrainment paradigms provide excellent temporal resolution, complementary approaches such as fMRI, MEG, pupillometry, and eye-tracking may help clarify how different neural systems contribute to extraction, chunking, prediction, and explicit access to learned structure (Schneider et al., 2022; Ward et al., 2013; Batterink and Paller, 2017). Multimodal neuroimaging approaches may be especially informative for investigating how modality-specific sensory systems interact with domain-general attentional and memory networks during learning.

Third, future work should further investigate the relationship between online implicit measures and offline explicit behavioral expressions of learning. The present thesis supports the idea that different measures may tap partially dissociable representational outputs differing in accessibility to conscious report. However, the precise relationships among neural entrainment, reaction-time facilitation, familiarity judgments, and explicit recognition remain insufficiently understood. Combining online continuous measures with computational modeling approaches may help clarify how representations evolve from early perceptual sensitivity to consciously accessible knowledge (Thiessen et al., 2013; Dale et al., 2012; Batterink et al., 2015).

Finally, future developmental work should investigate how literacy acquisition,

attentional maturation, and sensory experience shape modality-specific learning trajectories. Similarly, aging research should further examine the extent to which preserved implicit extraction mechanisms coexist with changes in explicit accessibility, attentional resources, and sensory integration processes (Wang et al., 2023; Moreau et al., 2022; Forest et al., 2023). Understanding how different components of SL change across the lifespan remains essential for developing comprehensive theories of learning and cognition.

Overall, the limitations of the present thesis reflect broader conceptual and methodological challenges currently facing the field of SL research. At the same time, they highlight promising directions for future work aimed at developing more reliable, theoretically precise, and ecologically valid approaches to understanding how humans acquire statistical structure from complex sensory environments.

Chapter 7

Conclusion

Statistical learning (SL) has often been described as a fundamental mechanism through which humans extract regularities from sensory input, supporting functions ranging from language acquisition to visual perception and adaptive behavior (Saffran et al., 1996a; Frost et al., 2019). Yet despite decades of research, important questions remain unresolved regarding the nature of SL, its underlying mechanisms, its developmental trajectory, and the extent to which different tasks and modalities reflect shared or distinct learning processes. The present thesis addressed these questions by investigating SL across auditory and visual modalities using both behavioral and neural methodologies in children, young adults, and older adults.

Across its empirical chapters, the thesis revealed a complex and heterogeneous picture of learning. Neural frequency-tagging paradigms demonstrated robust sensitivity to structured regularities even in the absence of strong explicit behavioral evidence, suggesting that neural entrainment may capture forms of online learning not readily accessible to conscious report. Cross-modal behavioral studies further showed that learning and transfer vary across sensory domains and developmental stages, supporting the view that SL is shaped by both shared computational principles and modality-dependent constraints. Developmental and aging findings additionally suggested that while some forms of implicit sensitivity to regularities remain relatively preserved across the lifespan, explicit accessibility, attentional resources, and the integration of multimodal information may vary substantially across age groups.

Taken together, these findings challenge strongly unitary conceptions of SL. Rather than reflecting a single domain-general ability uniformly expressed across tasks and modalities, the present thesis supports the view that SL is best conceptualized as a multicomponential phenomenon emerging from partially dissociable processes differing in representational format, processing demands, and accessibility to awareness. Neural entrainment, implicit behavioral facilitation, and explicit

recognition may therefore reflect different expressions or representational outputs of learning rather than interchangeable indices of the same underlying mechanism.

Building on this interpretation, the thesis proposed a multi-level framework of SL in which different measures capture distinct levels of representational accessibility. Within this framework, early neural sensitivity to regularities may index online extraction and perceptual binding processes, implicit behavioral measures may reflect facilitated processing and predictive sensitivity, and explicit behavioral tasks may require conscious access and metacognitive reflection on learned representations. Importantly, this framework does not imply fully independent learning systems, but rather partially overlapping processes interacting across multiple representational levels. Such an interpretation helps explain why correlations across SL tasks are often weak or inconsistent while still allowing for meaningful convergence within particular modalities, tasks, or levels of processing.

More broadly, the present thesis highlights the importance of integrating neural and behavioral methodologies within multimodal and lifespan-oriented approaches to SL research. The findings suggest that understanding learning requires moving beyond the search for a single optimal task or unified measure and instead examining how different paradigms map onto distinct components of learning. In this respect, variability across tasks, modalities, and individuals should not simply be treated as methodological noise, but rather as a valuable source of theoretical insight into the architecture of learning itself (Siegelman and Frost, 2015; Bogaerts et al., 2026).

Ultimately, the present work contributes to ongoing debates regarding the structure, measurement, and development of SL by emphasizing the need for theoretically informed, multilevel, and multimethod approaches. By integrating evidence across modalities, neural and behavioral measures, and developmental stages, the thesis provides a broader perspective on how humans acquire, represent, and generalize statistical structure in complex sensory environments. Rather than pointing toward a single unified mechanism, the present findings suggest that SL emerges from the dynamic interaction of multiple learning processes operating across different representational and cognitive levels. Understanding these interactions may constitute one of the central challenges for future research on learning and cognition.

References

Abreu, R., Postarnak, S., Vulchanov, V., Baggio, G., and Vulchanova, M. (2023). The association between statistical learning and language development during childhood: A scoping review. *Heliyon*, 9(7):e18693.

Altmann, G., Dienes, Z., and Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4):899.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., and Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior research methods*, 52(1):388–407.

Arciuli, J. (2017). Statistical learning: A potential mechanism of cognitive development. *Trends in Cognitive Sciences*, 21(8):573–575.

Arciuli, J. and Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, 14(3):464–473.

Arciuli, J. and Simpson, I. C. (2012). Statistical learning is lasting and consistent over time. *Neuroscience letters*, 517(2):133–135.

Arnon, I. (2020). Do current statistical learning tasks capture stable individual differences in children? an investigation of task reliability across modality. *Behavior research methods*, 52(1):68–81.

Aslin, R. N. (2017). Statistical learning: A powerful mechanism that operates by mere exposure. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1-2):e1373.

Aslin, R. N. and Saffran, J. R. (2025). Statistical learning: A core mechanism in a developmental hierarchy. *Current Opinion in Neurobiology*, 95:103124.

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., and Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: current status and future prospects. *Neuroscience & Biobehavioral Reviews*, 30(6):791–807.

Bánki, A., Brzozowska, A., Hoehl, S., and Köster, M. (2022). Neural entrainment vs. stimulus-tracking: a conceptual challenge for rhythmic perceptual stimulation in developmental neuroscience. *Frontiers in Psychology*, 13:878984.

Barillon, A., Schiltz, C., and Lochy, A. (2025). Tracking novel visual word learning via different methods with an original fpvs-eeg approach. *Frontiers in Human Neuroscience*, 19:1647925.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., Green, P., and Bolker, M. B. (2015). Package 'lme4'. *convergence*, 12(1):2.

Batikh, A., Barraud, A. B., Volland, F., Koun, E., Farnè, A., and Pélisson, D. (2025). Saccades adapt to visual and auditory stepping targets and display an asymmetrical pattern of cross-modal transfer. *Journal of Neurophysiology*.

Batterink, L. J. and Paller, K. A. (2017). Online neural monitoring of statistical learning. *cortex*, 90:31–45.

Batterink, L. J., Paller, K. A., and Reber, P. J. (2019). Understanding the neural bases of implicit and statistical learning. *Topics in cognitive science*, 11(3):482–503.

Batterink, L. J., Reber, P. J., Neville, H. J., and Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of memory and language*, 83:62–78.

Bays, B. C., Turk-Browne, N. B., and Seitz, A. R. (2015). Dissociable behavioural outcomes of visual statistical learning. *Visual Cognition*, 23(9–10):1072–1097.

Benitez, V. L., Bulgarelli, F., Byers-Heinlein, K., Saffran, J. R., and Weiss, D. J. (2020). Statistical learning of multiple speech streams: A challenge for monolingual infants. *Developmental science*, 23(2):e12896.

Bertels, J., Boursain, E., Destrebecqz, A., and Gaillard, V. (2015). Visual statistical learning in children and young adults: How implicit? *Frontiers in Psychology*, 5:1541.

Beyersmann, E., Wegener, S., Spencer, J., and Castles, A. (2023). Acquisition of orthographic forms via spoken complex word training. *Psychonomic Bulletin & Review*, 30(2):739–750.

Boeve, S., Zhou, H., and Bogaerts, L. (2024). A meta-analysis of 97 studies reveals that statistical learning and language ability are only weakly correlated. *L'Année Psychologique*, 124:283–316.

Bogaerts, L., Crepaldi, D., Zhou, H., Bautista, A., Brown, K., Chang, L.-Y., de Diego-Balaguer, R., Hofweber, J., Isbilen, E., Joanisse, M., et al. (2026). Time for consensus: Open challenges and recommendations for research on individual differences in statistical learning. *OSF*.

Bogaerts, L., Siegelman, N., Christiansen, M. H., and Frost, R. (2022a). Is there such a thing as a 'good statistical learner'? *Trends in cognitive sciences*, 26(1):25–37.

Bogaerts, L., Siegelman, N., Christiansen, M. H., and Frost, R. (2022b). Is there such a thing as a 'good statistical learner'? *Trends in Cognitive Sciences*, 26(1):25–37.

Bogaerts, L., Siegelman, N., and Frost, R. (2021). Statistical learning and language impairments: Toward more precise theoretical accounts. *Perspectives on Psychological science*, 16(2):319–337.

Boogert, N. J., Madden, J. R., Morand-Ferron, J., and Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1756).

Brainard, D. H. and Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, 10(4):433–436.

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? a tutorial of power analysis with reference tables. *Journal of cognition*, 2(1):16.

Brysbaert, M. (2024). Designing and evaluating tasks to measure individual differences in experimental psychology: A tutorial. *Cognitive Research: Principles and Implications*, 9(1):11.

Buiatti, M., Peña, M., and Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage*, 44(2):509–519.

Bulf, H., Johnson, S. P., and Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, 121(1):127–132.

Bulgarelli, F., Weiss, D. J., and Dennis, N. A. (2021). Cross-situational statistical learning in younger and older adults. *Aging, Neuropsychology, and Cognition*, 28(3):346–366.

Bulut, T. and Şenol, N. G. G. (2026). Association between statistical learning and morphosyntactic processing in turkish: A self-paced reading experiment. *Psikoloji Çalışmaları / Studies in Psychology*.

Cassar, M. and Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of educational psychology*, 89(4):631.

Chater, N. and Christiansen, M. H. (2018). Language acquisition as skill learning. *Current Opinion in Behavioral Sciences*, 21:205–208.

Chiarella, S. G., Simione, L., D'Angiò, M., Saracini, C., Raffone, A., and Di Pace, E. (2026). Implicit observational learning of second-order conditional repeated sequences presented in rapid serial visual presentation. *Consciousness and Cognition*, 137:103967.

Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., and Werker, J. F. (2020). Preverbal infants discover statistical word patterns at similar rates as adults: Evidence from neural entrainment. *Psychological Science*, 31(9):1161–1173.

Christiansen, M. H. (2019). Implicit statistical learning: A tale of two literatures. *Topics in cognitive science*, 11(3):468–481.

Christiansen, M. H. and Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and brain sciences*, 39:e62.

Cisneros, E., Karny, S., Ivry, R., and Tsay, J. S. (2024). Differential aging effects on implicit and explicit sensorimotor learning. *bioRxiv*.

Clayton, F. J. et al. (2020a). Auditory statistical learning and reading. *Journal of Experimental Psychology*.

Clayton, F. J., West, G., Sears, C., Hulme, C., and Lervåg, A. (2020b). A longitudinal study of early reading development: Letter-sound knowledge, phoneme awareness and ran, but not letter-sound integration, predict variations in reading development. *Scientific Studies of Reading*, 24(2):91–107.

Coll, M.-P., Murphy, J., Catmur, C., Bird, G., and Brewer, R. (2019). The importance of stimulus variability when studying face processing using fast periodic visual stimulation: A novel 'mixed-emotions' paradigm. *Cortex*, 117:182–195.

Conway, C. M. (2020). How does the brain learn environmental structure? ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews*, 112:279–299.

Conway, C. M., Bauernschmidt, A., Huang, S. S., and Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, 114(3):356–371.

Conway, C. M. and Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(1):24.

Conway, C. M. and Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological science*, 17(10):905–912.

Conway, C. M. and Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure. *Annals of the New York Academy of Sciences*, 1145:113–131.

Cortese, M. J. and Balota, D. A. (2012). Visual word recognition in skilled adult readers. *The Cambridge handbook of psycholinguistics*, pages 159–185.

Dale, R., Duran, N. D., and Morehead, J. (2012). Prediction during statistical learning, and implications for the implicit/explicit divide. *Advances in Cognitive Psychology*.

David, J., Koessler, L., and Rossion, B. (2023a). A robust neural index of automatic generalization across variable natural views of familiar face identities. *Visual Cognition*, 31(8):571–583.

David, J., Quenon, L., Hanseeuw, B. J., Ivanoiu, A., Oziat, J., Koessler, L., and Rossion, B. (2023b). Identification of new electrophysiological biomarkers of alzheimer's disease with fast periodic visual stimulation and electroencephalography. *Alzheimer's & Dementia*, 19:e082665.

Davidson, H. B. (2025). *Neural signatures of ADHD: A fast periodic visual stimulation EEG study*. Mississippi State University.

De Leeuw, J. R. (2015). jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior research methods*, 47:1–12.

De Rosa, M., Ktori, M., Vidal, Y., Bottini, R., and Crepaldi, D. (2022). Frequency-based neural discrimination in fast periodic visual stimulation. *Cortex*, 148:193–203.

De Rosa, M., Vignali, L., D'Urso, A., Ktori, M., Bottini, R., and Crepaldi, D. (2024). Selective neural entrainment reveals hierarchical tuning to linguistic regularities in reading. *Neurobiology of Language*, 5(2):528–552.

Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, 16(4):234–244.

Dehaene, S., Cohen, L., Sigman, M., and Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in cognitive sciences*, 9(7):335–341.

Delorme, A. and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21.

Destrebecqz, A. et al. (2005). The neural correlates of implicit and explicit sequence learning: Interacting networks revealed by the process dissociation procedure. *Learning & Memory*.

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature neuroscience*, 19(1):158–164.

Donohue, S. E., Weinhold, S., Schoenfeld, M. A., Quian Quiroga, R., and Hopf, J.-M. (2019). A neural hallmark of auditory implicit learning is altered in older adults. *PloS one*, 14(1):e0211468.

Duato, D., Giannelli, F., de Diego-Balaguer, R., and Pérez-Bellido, A. (2023). Cross-modal statistical learning is facilitated by modality predictability. *bioRxiv*.

Dzhelyova, M. and Rossion, B. (2014). Supra-additive contribution of shape and surface information to individual face discrimination as revealed by fast periodic visual stimulation. *Journal of Vision*, 14(14):15–15.

Dzhelyova, M., Schiltz, C., and Rossion, B. (2020). The relationship between the benton face recognition test and electrophysiological unfamiliar face individuation response as revealed by fast periodic stimulation. *Perception*, 49(2):210–221.

Ellis, E. M., Gonzalez, M. R., and Deák, G. O. (2014). Visual prediction in infancy: What is the association with later vocabulary? *Language Learning and Development*, 10(1):36–50.

Emberson, L. L. and Rubinstein, D. Y. (2016). Statistical learning is constrained to less abstract patterns in complex sensory input (but not the least). *Cognition*, 153:63–78.

Erickson, L. C. and Thiessen, E. D. (2015). Statistical learning of language: theory, validity, and predictions. *Developmental Review*, 37:66–108.

Evans, J. L., Saffran, J. R., and Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52(2):321–335.

Fairbanks, G. (1940). Voice and articulation drillbook. *(No Title)*.

Fiser, J. and Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological science*, 12(6):499–504.

Fiser, J. and Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3):458.

Fiser, J. and Aslin, R. N. (2005). Encoding multielement scenes: statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, 134(4):521.

Forest, T. A., Schlichting, M. L., Duncan, K. D., and Finn, A. S. (2023). Changes in statistical learning across development. *Nature Reviews Psychology*, 2:205–219.

Franco, A., Gaillard, V., Cleeremans, A., and Destrebecqz, A. (2015). Assessing segmentation processes by click detection: Online measure of statistical learning, or simple interference? *Behavior Research Methods*, 47:1393–1403.

Frost, R. (2012). Towards a universal model of reading. *Behavioral and brain sciences*, 35(5):263–279.

Frost, R., Armstrong, B. C., and Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, 145(12):1128.

Frost, R., Armstrong, B. C., Siegelman, N., and Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in cognitive sciences*, 19(3):117–125.

Frost, R., Siegelman, N., Narkiss, A., and Afek, L. (2009a). Statistical learning and reading: A model of language acquisition. *Psychological Science*, 24:1243–1252.

Frost, R., Siegelman, N., Narkiss, A., and Afek, L. (2013a). What predicts successful literacy acquisition in a second language? *Psychological science*, 24(7):1243–1252.

Frost, R., Siegelman, N., Narkiss, A., and Afek, L. (2013b). What predicts successful literacy acquisition in a second language? *Psychological Science*, 24(7):1243–1252.

Frost, S. J., Landi, N., Mencl, W. E., Sandak, R., Fulbright, R. K., Tejada, E. T., Jacobsen, L., Grigorenko, E. L., Constable, R. T., and Pugh, K. R. (2009b). Phonological awareness predicts activation patterns for print and speech. *Annals of dyslexia*, 59(1):78–97.

Gabay, Y. and Holt, L. L. (2015). Incidental learning of sound categories is impaired in developmental dyslexia. *cortex*, 73:131–143.

Garcia, R. B., Tomaino, A., and Cornoldi, C. (2019). Cross-modal working memory binding and learning of visual-phonological associations in children with reading difficulties. *Child Neuropsychology*, 25(8):1075–1094.

Ge, Y., Correia, S., Lee, Y.-W., Jin, Z., Rothman, J., and Rebuschat, P. (2025). Statistical learning of foreign language words in younger and older adults. *Bilingualism: Language and Cognition*, 28(3):716–727.

Gebhart, A. L., Aslin, R. N., and Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive science*, 33(6):1087–1116.

Glicksohn, A. and Cohen, A. (2013). The role of cross-modal associations in statistical learning. *Psychonomic Bulletin & Review*, 20(6):1161–1169.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological science*, 13(5):431–436.

Gómez, R. L. and Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in cognitive sciences*, 4(5):178–186.

Growns, B., Siegelman, N., Bogaerts, L., and Frost, R. (2020a). The latent structure of statistical learning performance and its relationship to language ability. *Open Mind*, 4:1–16.

Growns, B., Siegelman, N., and Martire, K. A. (2020b). The multi-faceted nature of visual statistical learning: Individual differences in learning conditional and distributional regularities across time and space. *Psychonomic Bulletin & Review*, 27(6):1291–1299.

Harm, M. W. and Seidenberg, M. S. (2004). Computing the meanings of words in reading. *Psychological Review*, 111(3):662–720.

Hedge, C., Powell, G., and Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior research methods*, 50(3):1166–1186.

Henry, M. J., Herrmann, B., Kunke, D., and Obleser, J. (2017). Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nature communications*, 8(1):15801.

Hermann, O. P., Petrini, K., Coulthard, E., and Stothart, G. (2025). Assessment of visual processing in alzheimer's disease and lewy body dementia using fast periodic visual stimulation. *Alzheimer's & Dementia*, 21:e108591.

Howard, J. H., Howard, D. V., Dennis, N. A., Yankovich, H., and Vaidya, C. J. (2004). Implicit spatial contextual learning in healthy aging. *Neuropsychology*, 18(1):124.

Huizeling, E., Wang, H., Holland, C., and Kessler, K. (2021). Changes in theta and alpha oscillatory signatures of attentional control in older and middle age. *European Journal of Neuroscience*, 54(1):4314–4337.

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., and Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological science*, 23(6):572–577.

Hupp, J. M. and Sloutsky, V. M. (2011). Learning to learn: From within-modality to cross-modality transfer during infancy. *Journal of Experimental Child Psychology*, 110(3):408–421.

Hupp, J. M., Sloutsky, V. M., and Culicover, P. W. (2009). Evidence for a domain-general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes*, 24(6):876–909.

Isbilen, E. S. and Christiansen, M. H. (2022). Statistical learning of language: A meta-analysis into 25 years of research. *Cognitive Science*, 46(9):e13198.

Isbilen, E. S., Laver, A., Siegelman, N., Magnuson, J. S., and Aslin, R. N. (2025). Finding words in a sea of text: Word search as a measure of sensitivity to statistical regularities in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Isbilen, E. S., McCauley, S. M., and Christiansen, M. H. (2022). Individual differences in artificial and natural language statistical learning. *Cognition*, 225:105123.

Isbilen, E. S., McCauley, S. M., Kidd, E., and Christiansen, M. H. (2020). Statistically induced chunking recall: A memory-based approach to statistical learning. *Cognitive science*, 44(7):e12848.

Jacques, C., Retter, T. L., and Rossion, B. (2016). A single glance at natural face images generate larger and qualitatively different category-selective spatio-temporal signatures than other ecologically-relevant categories in the human brain. *NeuroImage*, 137:21–33.

Janacsek, K. and Nemeth, D. (2013). Implicit sequence learning and working memory: Correlated or complicated? *Cortex*, 49(8):2001–2006.

Johannsen, L., Friedgen, E., Stephan, D. N., Batista, J., Schulze, D., Laurentius, T., Koch, I., and Bollheimer, L. C. (2022). Keeping in step with the young: Chronometric and kinematic data show intact procedural locomotor sequence learning in older adults. *Plos one*, 17(5):e0266733.

Johansson, T. (2009). Strengthening the case for stimulus-specificity in artificial grammar learning: No evidence for abstract representations with extended exposure. *Experimental Psychology*, 56(3):188–197.

Kalra, P. B., Gabrieli, J. D. E., and Finn, A. S. (2019). Evidence of stable individual differences in implicit learning. *Cognition*, 190:199–211.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., and Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, 116(3):321–340.

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, 48(1):171–184.

Kidd, E. and Arciuli, J. (2015). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, 87(1):184–193.

Kidd, E. and Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child development*, 87(1):184–193.

Kidd, E., Arciuli, J., Christiansen, M. H., Isbilen, E. S., Revius, K., and Smithson, M. (2020). Measuring children's auditory statistical learning via serial recall. *Journal of Experimental Child Psychology*, 200:104964.

Kidd, E., Arciuli, J., Christiansen, M. H., and Smithson, M. (2023). The sources and consequences of individual differences in statistical learning for language development. *Cognitive Development*, 66:101335.

Kidd, E., Donnelly, S., and Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in cognitive sciences*, 22(2):154–169.

Kimura, M., Schröger, E., and Czigler, I. (2011). Visual mismatch negativity and its importance in visual cognitive sciences. *Neuroreport*, 22(14):669–673.

Kirkham, N. Z., Slemmer, J. A., and Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83(2):B35–B42.

Kremláček, J., Kreegipuu, K., Tales, A., Astikainen, P., Poldver, N., Näätänen, R., and Stefanics, G. (2016). Visual mismatch negativity (vmmn): A review and meta-analysis of studies in psychiatric and neurological disorders. *Cortex*, 80:76–112.

Ktori, M., Lelonkiewicz, J. R., and Crepaldi, D. (2026). Statistical learning and reading: Visual regularities support affix detection in developing readers. *Child Development*, page aacaf014.

Lavi-Rotbain, O. and Arnon, I. (2021). Visual statistical learning is facilitated in zipfian distributions. *Cognition*, 206:104492.

Lavi-Rotbain, O. and Arnon, I. (2022). The learnability consequences of zipfian distributions in language. *Cognition*, 223:105038.

Lega, C., Di Caro, V., Strina, V., and Daini, R. (2023). Age-related differences in the statistical learning of target selection and distractor suppression. *Psychology and Aging*, 38(3):188.

Lelonkiewicz, J. R., Ktori, M., and Crepaldi, D. (2020). Morphemes as letter chunks: Discovering affixes through visual regularities. *Journal of Memory and language*, 115:104152.

Lelonkiewicz, J. R., Ktori, M., and Crepaldi, D. (2023). Morphemes as letter chunks: Linguistic information enhances the learning of visual regularities. *Journal of Memory and Language*, 130:104411.

Lengyel, G., Žalalytė, G., Pantelides, A., Ingram, J. N., Fiser, J., Lengyel, M., and Wolpert, D. M. (2019). Unimodal statistical learning produces multimodal object-like representations. *elife*, 8:e43942.

Li, S.-C., Lindenberger, U., and Sikström, S. (2001). Aging cognition: from neuromodulation to representation. *Trends in cognitive sciences*, 5(11):479–486.

Li, X., Zhao, X., Shi, W., Lu, Y., and Conway, C. M. (2018). Lack of cross-modal effects in dual-modality implicit statistical learning. *Frontiers in Psychology*, 9:146.

Ling, X., Su, J., Miao, Z., Jiang, Y., and Wu, X. (2025). How category frameworks shape the pre-activation of attentional templates: Evidence from high-resolution ssveps. *International Journal of Psychophysiology*, page 113246.

Liu-Shuang, J., Norcia, A. M., and Rossion, B. (2014). An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia*, 52:57–72.

Liu-Shuang, J., Torfs, K., and Rossion, B. (2016). An objective electrophysiological marker of face individualisation impairment in acquired prosopagnosia with fast periodic visual stimulation. *Neuropsychologia*, 83:100–113.

Lochy, A., Jacques, C., Maillard, L., Colnat-Coulbois, S., Rossion, B., and Jonas, J. (2018). Selective visual representation of letters and words in the left ventral occipito-temporal cortex with intracerebral recordings. *Proceedings of the National Academy of Sciences*, 115(32):E7595–E7604.

Lochy, A., Rossion, B., Ralph, M. L., Volfart, A., Hauk, O., and Schiltz, C. (2024). Linguistic and attentional factors–not statistical regularities–contribute to word-selective neural responses with fpvs-oddball paradigms. *Cortex*, 173:339–354.

Lochy, A., Van Belle, G., and Rossion, B. (2015). A robust index of lexical representation in the left occipito-temporal cortex as evidenced by eeg responses to fast periodic visual stimulation. *Neuropsychologia*, 66:18–31.

Lochy, A., Van Reybroeck, M., and Rossion, B. (2016). Left cortical specialization for visual letter strings predicts rudimentary knowledge of letter-sound association in preschoolers. *Proceedings of the National Academy of Sciences*, 113(30):8544–8549.

Lukács, Á., Ugrin, B. J., and Lukics, K. S. (2024). Statistical learning and individual differences in language abilities: A structural equation modelling study on the mediating roles of perceptual speed, working memory, and cognitive control. *SSRN Electronic Journal*.

Lukics, K. S. and Lukács, Á. (2021). Tracking statistical learning online: Word segmentation in a target detection task. *Acta Psychologica*, 215:103271.

Lukics, K. S. and Lukács, Á. (2022). Modality, presentation, domain and training effects in statistical learning. *Scientific Reports*, 12(1):20878.

Lum, J. A., Clark, G. M., Barhoun, P., Hill, A. T., Hyde, C., and Wilson, P. H. (2023). Neural basis of implicit motor sequence learning: Modulation of cortical power. *Psychophysiology*, 60(2):e14179.

Maheu, M., Dehaene, S., and Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans. *eLife*, 8.

Mainela-Arnold, E. and Evans, J. L. (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without sli? *Journal of child language*, 41(2):327–351.

Mao, X., Shenton, N., Puthusserypady, S., Lauritzen, M. J., and Benedek, K. (2025). Auditory steady state response can predict declining ef in healthy elderly individuals. *Frontiers in Aging Neuroscience*, 17:1516932.

Marchive, M., Rossion, B., and Lochy, A. (2025). Optimal word reading rate as evidenced by frequency-tagging electrophysiology. *Journal of Cognitive Neuroscience*, 37(5):988–1008.

Marcus, G. F., Fernandes, K. J., and Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological science*, 18(5):387–391.

Marcus, G. F., Vijayan, S., Bandi Rao, S., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398):77–80.

Mateus-Moreno, A., Lara-Díaz, M. F., Adrover-Roig, D., Aguilar-Mediavilla, E., and Jiménez-Fernández, G. (2025). Impaired visual and verbal statistical learning in children with dyslexia in a transparent orthography. *Annals of Dyslexia*.

Maye, J., Werker, J. F., and Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3):B101–B111.

Melby-Lervåg, M., Lyster, S.-A. H., and Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychological bulletin*, 138(2):322.

Misyak, J. B. and Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, 62(1):302–331.

Misyak, J. B., Christiansen, M. H., and Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in psychology*, 1:31.

Mitchel, A. D. and Weiss, D. J. (2011a). Learning across senses: cross-modal effects in multisensory statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5):1081.

Mitchel, A. D. and Weiss, D. J. (2011b). Learning across senses: cross-modal statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37:1081–1091.

Moreau, C. N., Joanisse, M. F., Mulgrew, J., and Batterink, L. J. (2022). No statistical learning advantage in children over adults: Evidence from behaviour and neural entrainment. *Developmental Cognitive Neuroscience*, 57:101154.

Morris, J. et al. (2025a). How individual differences shape erp responses to visual statistical learning. *bioRxiv*.

Morris, J., Kealey, E., Greene, F., Pulido, J., Rooney, T., Alzate, N., Moore, E., and Ramos-Marte, C. (2025b). How individual differences shape erp responses to visual statistical learning. *bioRxiv*.

Moser, J., Batterink, L., Hegner, Y. L., Schleger, F., Braun, C., Paller, K., and Preissl, H. (2021a). Dynamics of nonlinguistic statistical learning: From neural entrainment to the emergence of explicit knowledge. *NeuroImage*, 240.

Moser, J., Batterink, L., Hegner, Y. L., Schleger, F., Braun, C., Paller, K. A., and Preissl, H. (2021b). Dynamics of nonlinguistic statistical learning: From neural entrainment to the emergence of explicit knowledge. *NeuroImage*, 240:118378.

Murray, A. and Prashad, S. (2025). A review of the differential effects of age on online and offline learning of implicit motor sequences. *Journal of Motor Learning and Development*, 13(3):605–631.

Murray, C. A. and Shams, L. (2023). Crossmodal interactions in human learning and memory. *Frontiers in Human Neuroscience*, 17:1181760.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (mmn) in basic research of central auditory processing: a review. *Clinical neurophysiology*, 118(12):2544–2590.

Nagy, B., Kojouharova, P., Protzner, A. B., and Gaál, Z. A. (2024). Investigating the effect of contextual cueing with face stimuli on electrophysiological measures in younger and older adults. *Journal of Cognitive Neuroscience*, 36(5):776–799.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., and Chertkow, H. (2005). The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4):695–699.

Nastase, S. A., Goldstein, A., and Hasson, U. (2020). Keep it real: Rethinking the primacy of experimental control in cognitive neuroscience. *NeuroImage*, 222:117254.

Nava, E., Föcker, J., and Gori, M. (2020). Children can optimally integrate multisensory information after a short action-like mini game training. *Developmental science*, 23(1):e12840.

Neger, T. M., Rietveld, T., and Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in human neuroscience*, 8:628.

Nijhof, A. D., Catmur, C., Brewer, R., Coll, M.-P., Wiersema, J. R., and Bird, G. (2024). Differences in own-face but not own-name discrimination between autistic and neurotypical adults: A fast periodic visual stimulation-eeg study. *cortex*, 171:308–318.

Nissen, M. J. and Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive psychology*, 19(1):1–32.

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., and Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of vision*, 15(6):4–4.

Nurdal, V. (2022). *Using EEG frequency tagging to assess memory in healthy younger adults*. PhD thesis, University of Bath.

Orpella, J., Assaneo, M. F., Ripollés, P., Noejovich, L., López-Barroso, D., Diego-Balaguer, R. d., and Poeppel, D. (2022). Differential activation of a frontoparietal network explains population-level differences in statistical learning from speech. *PLoS biology*, 20(7):e3001712.

Pacton, S., Perruchet, P., Fayol, M., and Cleeremans, A. (2001). Implicit learning out of the lab: the case of orthographic regularities. *Journal of experimental psychology: General*, 130(3):401.

Palmer, S. D., Hutson, J., and Mattys, S. L. (2018). Statistical learning for speech segmentation: Age-related changes and underlying mechanisms. *Psychology and aging*, 33(7):1035.

Pavlidou, E. V. and Bogaerts, L. (2019). Implicit statistical learning across modalities and its relationship with reading in childhood. *Frontiers in Psychology*, 10:1834.

Pedraza, F., Farkas, B. C., Vékony, T., et al. (2024). Evidence for a competitive relationship between executive functions and statistical learning. *npj Science of Learning*, 9(1):30.

Peirce, J. W. (2007). Psychopy—psychophysics software in python. *Journal of neuroscience methods*, 162(1-2):8–13.

Perruchet, P. and Vinter, A. (1998). Parser: A model for word segmentation. *Journal of memory and language*, 39(2):246–263.

Pescuma, V. N., Ktori, M., Beyersmann, E., Sowman, P. F., Castles, A., and Crepaldi, D. (2022). Automatic morpheme identification across development: Magnetoencephalography (meg) evidence from fast periodic visual stimulation. *Frontiers in Psychology*, 13:932952.

Peykarjou, S., Hoehl, S., and Pauen, S. (2024). The development of global-level categorization: frequency tagging eeg responses. *Brain Sciences*, 14(6):541.

Pikulski, J. J. and Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The reading teacher*, 58(6):510–519.

Pinto, D., Prior, A., and Golumbic, E. Z. (2021). Assessing the sensitivity of eeg-based frequency-tagging as a metric for statistical learning. *Neurobiology of Language*.

Pinto, T. C., Machado, L., Bulgacov, T. M., Rodrigues-Júnior, A. L., Costa, M. L., Ximenes, R. C., and Sougey, E. B. (2019). Is the montreal cognitive assessment (moca) screening superior to the mini-mental state examination (mmse) in the detection of mild cognitive impairment (mci) and alzheimer's disease (ad) in the elderly? *International psychogeriatrics*, 31(4):491–504.

Posit team (2023). *RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA.

Powell, D. and Atkinson, L. (2021). Unraveling the links between rapid automatized naming (ran), phonological awareness, and reading. *Journal of Educational Psychology*, 113(4):706.

Prem, P., Saggu, S., Boadu, A., Saju, S., Nisbet, K., and Cummine, J. (2025). Neuroanatomical correlates of auditory and visual statistical learning: Cortical and subcortical volume predictors. *Neuroscience*.

Qi, Z., Sanchez Araujo, Y., Georgan, W. C., Gabrieli, J. D., and Arciuli, J. (2019). Hearing matters more than seeing: A cross-modality study of statistical learning and reading ability. *Scientific Studies of Reading*, 23(1):101–115.

Quentin, R., Fanuel, L., Kiss, M., Vernet, M., Vékony, T., Janacsek, K., Cohen, L. G., and Nemeth, D. (2021). Statistical learning occurs during practice while high-order rule learning during rest period. *NPJ Science of learning*, 6(1):14.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raviv, L. and Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities. *Developmental Science*, 21(4):e12593.

Rawal, A. and Tseng, P. (2021). The effect of visual statistical learning in rsvp: Implicit learning or stream location artifact? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(8):1246.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118(3):219–235.

Regan, D. (1966). Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and clinical neurophysiology*, 20(3):238–248.

Ren, J. and Wang, M. (2024). Contribution of statistical learning in learning to read across languages. *Plos one*, 19(3):e0298670.

Ren, J., Wang, M., and Arciuli, J. (2023). A meta-analysis on the correlations between statistical learning, language, and reading outcomes. *Developmental Psychology*, 59(9):1626–1645.

Retter, T. L. and Rossion, B. (2016). Uncovering the neural magnitude and spatiotemporal dynamics of natural image categorization in a fast visual stream. *Neuropsychologia*, 91:9–28.

Robinson, C. W. and Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child development*, 75(5):1387–1401.

Robinson, C. W. and Sloutsky, V. M. (2007). Visual processing speed: Effects of auditory input on visual processing. *Developmental science*, 10(6):734–740.

Rosenthal, C. R., Andrews, S. K., Antoniades, C. A., Kennard, C., and Soto, D. (2016). Learning and recognition of a non-conscious sequence of events in human primary visual cortex. *Current Biology*, 26(6):834–841.

Rosenthal, O., Fusi, S., and Hochstein, S. (2001). Forming classes by stimulus frequency: Behavior and theory. *Proceedings of the National Academy of Sciences*, 98(7):4265–4270.

Rossion, B. (2014). Understanding individual face discrimination by means of fast periodic visual stimulation. *Experimental brain research*, 232(6):1599–1621.

Rossion, B. (2020). Biomarkers of face perception in autism spectrum disorder: time to shift to fast periodic visual stimulation with electroencephalography? *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 5(3):258–260.

Rossion, B., Torfs, K., Jacques, C., and Liu-Shuang, J. (2015). Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. *Journal of vision*, 15(1):18–18.

Russo, S., Carnovalini, F., Calignano, G., Arfé, B., Rodà, A., and Valenza, E. (2024). Linking vestibular, tactile, and somatosensory rhythm perception to language development in infancy. *Cognition*, 243:105688.

Rutkowski, T. M., Narębski, S., Otake-Matsuura, M., and Komendziński, T. (2025). Early detection of cognitive impairment in elderly using a passive fpvs-eeg bci and machine learning–extended version. *arXiv preprint arXiv:2504.10973*.

Ruzza, C. and Crepaldi, D. (2026). Statistical learning of categories: discrepancies between neural response and behaviour. *Neuroscience*.

Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, 44(4):493–515.

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *science*, 274(5294):1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1):27–52.

Saffran, J. R. and Kirkham, N. Z. (2017). Infant statistical learning. *Annual Review of Psychology*.

Saffran, J. R., Newport, E. L., and Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of memory and language*, 35(4):606–621.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., and Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8(2):101–105.

Sáringer, S., Kaposvári, P., and Benyhe, A. (2024). Visual linguistic statistical learning is traceable through neural entrainment. *Psychophysiology*, 61(8):e14575.

Sawi, O. M. and Rueckl, J. (2019). Reading and the neurocognitive bases of statistical learning. *Scientific Studies of Reading*, 23(1):8–23.

Schevenels, K., Altvater-Mackensen, N., Zink, I., De Smedt, B., and Vandermosten, M. (2023). Aging effects and feasibility of statistical learning tasks across modalities. *Aging, Neuropsychology, and Cognition*, 30(2):201–230.

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., and Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience*, 29(1):37–51.

Schmalz, X., Moll, K., Mulatti, C., and Schulte-Körne, G. (2019). Is statistical learning ability related to reading ability, and if so, why? *Scientific Studies of Reading*, 23(1):64–76.

Schneider, J., Weng, Y.-L., Hu, A., and Qi, Z. (2022). Linking the neural basis of distributional statistical learning with transitional statistical learning: The paradox of attention. *Neuropsychologia*.

Schuler, K. D., Reeder, P. A., Getz, H., Kissinger, K., Aslin, R. N., and Newport, E. L. (2026). 6-to-8 year-old children adjust their linguistic generalizations based on distributional cues. *Journal of Cognition and Development*, pages 1–26.

Scott, K. and Schulz, L. (2017). Lookit (part 1): A new online platform for developmental research. *Open Mind*, 1(1):4–14.

Seidenberg, M. S. and McClelland, J. L. (1989). A distributed, developmental model of word recognition. *Psychological Review*, 96(4):523–568.

Shafto, C. L., Conway, C. M., Field, S. L., and Houston, D. M. (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy*, 17(3):247–271.

Siegelman, N. (2020a). Statistical learning abilities and their relation to language. *Language and Linguistics Compass*, 14(3):e12365.

Siegelman, N. (2020b). Statistical learning abilities and their relation to language. *Language and Linguistics Compass*, 14(3):e12365.

Siegelman, N., Bogaerts, L., Christiansen, M. H., and Frost, R. (2017a). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B*, 372(1711):20160059.

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., and Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, 177:198–213.

Siegelman, N., Bogaerts, L., and Frost, R. (2017b). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior research methods*, 49:418–432.

Siegelman, N. and Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of memory and language*, 81:105–120.

Siegelman, N. and Frost, R. (2017). Toward a theory of individual differences in statistical learning. *Psychonomic Bulletin & Review*, 24(3):969–995.

Singh, L., Steven Reznick, J., and Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental science*, 15(4):482–495.

Sklenar, A. M. and Mienaltowski, A. (2019). The impact of emotional faces on younger and older adults' attentional blink. *Cognition and Emotion*, 33(7):1436–1447.

Slone, L. K. and Johnson, S. P. (2018). When learning goes beyond statistics: Infants represent visual sequences in terms of chunks. *Cognition*, 178:92–102.

Soares, A., Gutiérrez-Domínguez, F.-J., Vasconcelos, M., et al. (2020). Not all words are equally acquired: Transitional probabilities and instructions affect the electrophysiological correlates of statistical learning. *Frontiers in Human Neuroscience*, 14:577991.

Solaja, O., Carr, J. W., and Crepaldi, D. (2026). Does morphology support novel word learning through statistical learning? *Language, Cognition and Neuroscience*, pages 1–15.

Spencer, M., Kaschak, M. P., Jones, J. L., and Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and writing*, 28(4):467–490.

Stacchi, L., Ramon, M., Lao, J., and Caldara, R. (2019). Neural representations of faces are tuned to eye movements. *Journal of Neuroscience*, 39(21):4113–4123.

Stefanics, G., Astikainen, P., and Czigler, I. (2015). Visual mismatch negativity (vmmn): a prediction error signal in the visual modality.

Stothart, G., Alderman, S., Hermann, O., Creavin, S., and Coulthard, E. J. (2025). A passive and objective measure of recognition memory in mild cognitive impairment using fastball memory assessment. *Brain Communications*, 7(5):fcaf279.

Stothart, G., Quadflieg, S., and Milton, A. (2017). A fast and implicit measure of semantic categorisation using steady state visual evoked potentials. *Neuropsychologia*, 102:11–18.

Stothart, G. and Smith, L. (2020). A new functional biomarker of cognitive impairment in alzheimer's disease: Neuroimaging/optimal neuroimaging measures for early detection. *Alzheimer's & Dementia*, 16:e041056.

Strzelczyk, D., Peylo, C., and Langer, N. (2025). Impact of aging on theta-phase gamma-amplitude coupling during learning: A multivariate analysis. *bioRxiv*, pages 2025–02.

Sun, L., Ren, J., and He, W. (2017). Neural correlates of facial expression processing during a detection task: An erp study. *PLoS One*, 12(3):e0174016.

Sznabel, D., Land, R., Kopp, B., and Kral, A. (2023). The relation between implicit statistical learning and proactivity as revealed by eeg. *Scientific Reports*, 13(1):15787.

Tang, Y., ZHAO, Y., ZENG, Q., ZHANG, Z., and WU, S. (2026). Cross-modal transfer of statistical learning under unimodal and multimodal learning conditions. *Acta Psychologica Sinica*, 58(4):590.

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., and Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC neuroscience*, 10(1):21.

The MathWorks Inc. (2021). Matlab (version r2021b).

Thiessen, E. D., Kronstein, A. T., and Hufnagle, D. G. (2013). The extraction and integration framework: a two-process account of statistical learning. *Psychological bulletin*, 139(4):792.

Tong, X., Leung, W. W. S., and Tong, X. (2019). Visual statistical learning and orthographic awareness in chinese children with and without developmental dyslexia. *Research in developmental disabilities*, 92:103443.

Torgeson, J. K., Wagner, R. K., and Rashotte, C. A. (1999). Test review: Test of word reading efficiency (towre). *Inc.: Austin, TX, USA*.

Turk-Browne, N. B., Isola, P. J., Scholl, B. J., and Treat, T. A. (2008). Multidimensional visual statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2):399.

Turk-Browne, N. B., Jungé, J. A., and Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, 134(4):552.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., and Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, 21(10):1934–1945.

Uno, K., Asano, M., Kadowaki, H., and Yokosawa, K. (2020). Grapheme-color associations can transfer to novel graphemes when synesthetic colors function as grapheme "discriminating markers". *Psychonomic Bulletin & Review*, 27(4):700–706.

Van der Donck, S., Dzhelyova, M., Vettori, S., Mahdi, S. S., Claes, P., Steyaert, J., and Boets, B. (2020). Rapid neural categorization of angry and fearful faces is specifically impaired in boys with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 61(9):1019–1029.

van der Wulp, I., Struiksma, M., Batterink, L., and Wijnen, F. (2025). Neural evidence for linguistic statistical learning is independent of rhythmic and cognitive abilities in neurotypical adults. *cortex*.

Varela, I. G., Orpella, J., Poeppel, D., Ripolles, P., and Assaneo, M. F. (2024). Syllabic rhythm and prior linguistic knowledge interact with individual differences to modulate phonological statistical learning. *Cognition*, 245:105737.

Verty, L. V., Mellah, S., Maltezos, S., Boujut, A., Lussier, M., Bherer, L., and Belleville, S. (2024). Youth-like brain activation linked with greater cognitive training gains in older adults: Insights from the actop study. *Cortex*, 176:221–233.

Vettori, S., Dzhelyova, M., Van der Donck, S., Jacques, C., Steyaert, J., Rossion, B., and Boets, B. (2019). Fast periodic visual stimulation eeg reveals reduced social bias in autism. *Journal of Vision*, 19(10):25a–25a.

Vidal, C., Content, A., and Chetail, F. (2017). Bacs: The brussels artificial character sets for studies in cognitive psychology and neuroscience. *Behavior research methods*, 49:2093–2112.

Vidal, Y., Viviani, E., Zoccolan, D., and Crepaldi, D. (2021). A general-purpose mechanism of visual feature association in visual word identification and beyond. *Current Biology*, 31(6):1261–1267.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., and Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron*, 55(1):143–156.

Volfart, A., Rice, G. E., Ralph, M. A. L., and Rossion, B. (2021). Implicit, automatic semantic word categorisation in the left occipito-temporal cortex as revealed by fast periodic visual stimulation. *NeuroImage*, 238:118228.

von Koss Torkildsen, J., Arciuli, J., and Wie, O. B. (2019a). Individual differences in statistical learning predict children's reading ability in a semi-transparent orthography. *Learning and individual differences*, 69:60–68.

von Koss Torkildsen, J., Arciuli, J., and Wie, O. B. (2019b). Individual differences in statistical learning predict children's reading ability in a semi-transparent orthography. *Learning and individual differences*, 69:60–68.

Wagner, R. K. and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological bulletin*, 101(2):192.

Walk, A. and Conway, C. (2011). Multisensory statistical learning: can cross-modal associations be acquired? In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 33.

Walker, M. E., Vibell, J. F., Dewald, A. D., and Sinnett, S. (2022). Ageing and selective inhibition of irrelevant information in an attention-demanding rapid serial visual presentation task. *Brain and Neuroscience Advances*, 6:23982128211073427.

Wang, H., Köhler, S., and Batterink, L. (2023). Separate but not independent: Behavioral pattern separation and statistical learning are differentially affected by aging. *Cognition*, 236:105564.

Ward, E. J., Chun, M., and Kuhl, B. A. (2013). Repetition suppression and multivoxel pattern similarity differentially track implicit and explicit visual memory. *Journal of Neuroscience*.

West, G. and Vadillo, M. A. (2018). Statistical learning is not stable across attention or modality and is not a predictor of linguistic abilities. *Journal of Memory and Language*, 98:266–284.

West, G., Vadillo, M. A., and Shanks, D. R. (2018). The procedural learning deficit hypothesis of language learning disorders. *Developmental Science*, 21:e12552.

Yildirim, I. and Jacobs, R. A. (2015). Learning multisensory representations for auditory-visual transfer of sequence category knowledge: a probabilistic language of thought approach. *Psychonomic bulletin & review*, 22(3):673–686.

Zając-Lamparska, L., Zabielska-Mendyk, E., Zapała, D., and Augustynowicz, P. (2025). Cognitive training and retest learning effects on theta and alpha power in older and young adults: A perspective on the crunch hypothesis and the stac-r model. *International Journal of Clinical and Health Psychology*, 25(2):100568.

Zhang, F. and Emberson, L. L. (2019). Opposing timing constraints severely limit the use of pupillometry to investigate visual statistical learning. *Frontiers in Psychology*, 10.

Zhou, H., van der Ham, S., De Boer, B., Bogaerts, L., and Raviv, L. (2024). Modality and stimulus effects on distributional statistical learning: Sound vs. sight, time vs. space. *Journal of Memory and Language*, 138:104531.

Zolnai, T., David, D., Pesthy, O., Németh, M., Kiss, M., Nagy, M., and Németh, D. (2022). Measuring statistical learning by eye-tracking. *Experimental Results*, 3.

Supplementary Materials

Table S1. Word stimuli - Experiment 1

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
1	polso	FPVS	1733	6.238799	1.45
2	salto	FPVS	3010	6.478566	1.00
3	crema	FPVS	1539	6.187239	1.00
4	mosca	FPVS	1902	6.279211	1.40
5	pasta	FPVS	1267	6.102777	1.00
6	ombra	FPVS	2707	6.432488	1.45
7	dieta	FPVS	1138	6.056142	1.00
8	scala	FPVS	2514	6.400365	1.00
9	conte	FPVS	2118	6.325926	1.00
10	legno	FPVS	2299	6.361539	1.00
11	forno	FPVS	2299	6.361539	1.00
12	gusto	FPVS	2528	6.402777	1.00
13	epoca	FPVS	2418	6.383456	1.60
14	miele	FPVS	2064	6.314710	1.45
15	bosco	FPVS	2248	6.351796	1.15
16	fiato	FPVS	1815	6.258877	1.00
1	sposo	Memory - Stim. vs. Foil - Foil B1	1734	6.239049	1.35
2	crisi	Memory - Stim. vs. Foil - Foil B1	3171	6.501196	1.40
3	bimbo	Memory - Stim. vs. Foil - Foil B1	1699	6.230193	1.35
4	costo	Memory - Stim. vs. Foil - Foil B1	1797	6.254548	1.00
5	ansia	Memory - Stim. vs. Foil - Foil B1	1275	6.105510	1.50
6	volto	Memory - Stim. vs. Foil - Foil B1	2714	6.433610	1.00
7	verme	Memory - Stim. vs. Foil - Foil B1	1140	6.056905	1.00
8	furto	Memory - Stim. vs. Foil - Foil B1	2634	6.420616	1.10
9	cassa	Memory - Stim. vs. Foil - Foil B1	2067	6.315340	1.00
10	torre	Memory - Stim. vs. Foil - Foil B1	2255	6.353147	1.00
11	delia	Memory - Stim. vs. Foil - Foil B1	1594	6.202488	1.00
12	bugia	Memory - Stim. vs. Foil - Foil B1	2553	6.407051	1.50
13	ferro	Memory - Stim. vs. Foil - Foil B1	2238	6.349860	1.00
14	folla	Memory - Stim. vs. Foil - Foil B1	2031	6.307710	1.00
15	gomma	Memory - Stim. vs. Foil - Foil B1	1768	6.247482	1.40
16	nozze	Memory - Stim. vs. Foil - Foil B1	1807	6.256958	1.40
1	banco	Memory - Stim. vs. Foil - Foil B2	1681	6.225568	1.00
2	copia	Memory - Stim. vs. Foil - Foil B2	3045	6.483587	1.00
3	villa	Memory - Stim. vs. Foil - Foil B2	1433	6.156246	1.00
4	tocco	Memory - Stim. vs. Foil - Foil B2	1767	6.247237	1.00
5	fango	Memory - Stim. vs. Foil - Foil B2	1322	6.121231	1.00
6	socio	Memory - Stim. vs. Foil - Foil B2	2715	6.433770	1.00
7	molla	Memory - Stim. vs. Foil - Foil B2	1121	6.049606	1.00

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
8	sparo	Memory - Stim. vs. Foil - Foil B2	2688	6.429429	1.00
9	sorta	Memory - Stim. vs. Foil - Foil B2	2178	6.338058	1.00
10	dosso	Memory - Stim. vs. Foil - Foil B2	2362	6.373280	1.00
11	monte	Memory - Stim. vs. Foil - Foil B2	1599	6.203848	1.00
12	ladro	Memory - Stim. vs. Foil - Foil B2	3079	6.488410	1.20
13	prete	Memory - Stim. vs. Foil - Foil B2	2656	6.424228	1.00
14	vasca	Memory - Stim. vs. Foil - Foil B2	2017	6.304706	1.25
15	barba	Memory - Stim. vs. Foil - Foil B2	1832	6.262925	1.05
16	disco	Memory - Stim. vs. Foil - Foil B2	2133	6.328991	1.00
1	massa	Memory - Foil vs. Foil - Item	1741	6.240799	1.00
2	dente	Memory - Foil vs. Foil - Item	1036	6.015360	1.00
3	danza	Memory - Foil vs. Foil - Item	1774	6.248954	1.00
4	soldo	Memory - Foil vs. Foil - Item	1167	6.067071	1.00
5	merce	Memory - Foil vs. Foil - Item	1375	6.138303	1.00
6	pasto	Memory - Foil vs. Foil - Item	1758	6.245019	1.00
7	cella	Memory - Foil vs. Foil - Item	3320	6.521138	1.00
8	preda	Memory - Foil vs. Foil - Item	1170	6.068186	1.05
9	fiore	Memory - Foil vs. Foil - Item	1764	6.246499	1.10
10	spesa	Memory - Foil vs. Foil - Item	1517	6.180986	1.15
11	frigo	Memory - Foil vs. Foil - Item	1475	6.168792	1.20
12	norma	Memory - Foil vs. Foil - Item	991	5.996074	1.25
13	succo	Memory - Foil vs. Foil - Item	1789	6.252610	1.30
14	ciclo	Memory - Foil vs. Foil - Item	1064	6.026942	1.35
15	talpa	Memory - Foil vs. Foil - Item	937	5.971740	1.40
16	drago	Memory - Foil vs. Foil - Item	1382	6.140508	1.45
1	suono	Memory - Foil vs. Foil - Foil B1	3678	6.565612	1.00
2	posta	Memory - Foil vs. Foil - Foil B1	3277	6.515476	1.00
3	mazza	Memory - Foil vs. Foil - Foil B1	1256	6.098990	1.00
4	fetta	Memory - Foil vs. Foil - Foil B1	918	5.962843	1.00
5	spina	Memory - Foil vs. Foil - Foil B1	1374	6.137987	1.00
6	fonte	Memory - Foil vs. Foil - Foil B1	2904	6.462997	1.00
7	prato	Memory - Foil vs. Foil - Foil B1	944	5.974972	1.00
8	multa	Memory - Foil vs. Foil - Foil B1	989	5.995196	1.05
9	ruolo	Memory - Foil vs. Foil - Foil B1	3314	6.520353	1.10
10	lotta	Memory - Foil vs. Foil - Foil B1	3533	6.548144	1.00
11	mensa	Memory - Foil vs. Foil - Foil B1	1076	6.031812	1.20
12	circo	Memory - Foil vs. Foil - Foil B1	1405	6.147676	1.25
13	curva	Memory - Foil vs. Foil - Foil B1	979	5.990783	1.30
14	mucca	Memory - Foil vs. Foil - Foil B1	1155	6.062582	1.30
15	cuoco	Memory - Foil vs. Foil - Foil B1	980	5.991226	1.35
16	pugno	Memory - Foil vs. Foil - Foil B1	3267	6.514149	1.45
1	pollo	Memory - Foil vs. Foil - Foil B2	3761	6.575303	1.00
2	tasse	Memory - Foil vs. Foil - Foil B2	1832	6.262925	1.00
3	valle	Memory - Foil vs. Foil - Foil B2	1327	6.122871	1.00
4	canto	Memory - Foil vs. Foil - Foil B2	1764	6.246499	1.00

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
5	botta	Memory - Foil vs. Foil - Foil B2	1040	6.017033	1.00
6	grano	Memory - Foil vs. Foil - Foil B2	925	5.966142	1.00
7	gesto	Memory - Foil vs. Foil - Foil B2	1757	6.244772	1.00
8	animo	Memory - Foil vs. Foil - Foil B2	945	5.975432	1.10
9	somma	Memory - Foil vs. Foil - Foil B2	1269	6.103462	1.10
10	sorte	Memory - Foil vs. Foil - Foil B2	1264	6.101747	1.00
11	pesca	Memory - Foil vs. Foil - Foil B2	1795	6.254064	1.20
12	zuppa	Memory - Foil vs. Foil - Foil B2	1807	6.256958	1.30
13	quota	Memory - Foil vs. Foil - Foil B2	1099	6.040998	1.30
14	scudo	Memory - Foil vs. Foil - Foil B2	1155	6.062582	1.35
15	falla	Memory - Foil vs. Foil - Foil B2	2255	6.353147	1.00
16	zaino	Memory - Foil vs. Foil - Foil B2	1036	6.015360	1.55
1	pacco	Discrimination - Foil B1	1754	6.244030	1.00
2	pizza	Discrimination - Foil B1	2889	6.460748	1.00
3	tenda	Discrimination - Foil B1	1534	6.185825	1.00
4	botte	Discrimination - Foil B1	1590	6.201397	1.00
5	guaio	Discrimination - Foil B1	1261	6.100715	1.50
6	petto	Discrimination - Foil B1	2782	6.444357	1.00
7	trono	Discrimination - Foil B1	1122	6.049993	1.00
8	gioia	Discrimination - Foil B1	3090	6.489958	1.10
9	penna	Discrimination - Foil B1	2134	6.329194	1.00
10	museo	Discrimination - Foil B1	2023	6.305996	1.75
11	targa	Discrimination - Foil B1	1574	6.197005	1.00
12	patto	Discrimination - Foil B1	2640	6.421604	1.00
13	retta	Discrimination - Foil B1	2389	6.378216	1.00
14	morso	Discrimination - Foil B1	1987	6.298198	1.00
15	frase	Discrimination - Foil B1	2118	6.325926	1.10
16	busta	Discrimination - Foil B1	1844	6.265761	1.15
1	salsa	Discrimination - Foil B2	1702	6.230960	1.15
2	banda	Discrimination - Foil B2	3238	6.510277	1.00
3	reato	Discrimination - Foil B2	1520	6.181844	1.00
4	gonna	Discrimination - Foil B2	1239	6.093071	1.00
5	tigre	Discrimination - Foil B2	1355	6.131939	1.40
6	firma	Discrimination - Foil B2	2716	6.433930	1.30
7	pozzo	Discrimination - Foil B2	1150	6.060698	1.10
8	regno	Discrimination - Foil B2	2830	6.451786	1.00
9	torto	Discrimination - Foil B2	1924	6.284205	1.00
10	metro	Discrimination - Foil B2	2155	6.333447	1.00
11	ritmo	Discrimination - Foil B2	1622	6.210051	1.35
12	abito	Discrimination - Foil B2	2642	6.421933	1.00
13	becco	Discrimination - Foil B2	2442	6.387746	1.30
14	palco	Discrimination - Foil B2	1977	6.296007	1.10
15	mappa	Discrimination - Foil B2	2460	6.390935	1.20
16	croce	Discrimination - Foil B2	1849	6.266937	1.60
1	burro	Discrimination - Foil B3	1677	6.224533	1.10

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
2	tazza	Discrimination - Foil B3	2843	6.453777	1.00
3	leone	Discrimination - Foil B3	1527	6.183839	1.25
4	virus	Discrimination - Foil B3	2372	6.375115	1.40
5	suolo	Discrimination - Foil B3	1259	6.100026	1.00
6	tasca	Discrimination - Foil B3	2791	6.445760	1.00
7	acido	Discrimination - Foil B3	1155	6.062582	1.15
8	punta	Discrimination - Foil B3	2787	6.445137	1.00
9	vetro	Discrimination - Foil B3	2181	6.338656	1.00
10	umore	Discrimination - Foil B3	2169	6.336260	1.15
11	ruota	Discrimination - Foil B3	1591	6.201670	1.50
12	tomba	Discrimination - Foil B3	2655	6.424065	1.30
13	corda	Discrimination - Foil B3	2454	6.389875	1.00
14	sfida	Discrimination - Foil B3	2110	6.324282	1.20
15	sonno	Discrimination - Foil B3	2672	6.426836	1.00
16	carro	Discrimination - Foil B3	1863	6.270213	1.00

Table S2. Pseudofont stimuli - Experiment 1

ID	Stimulus	BACS-2	Condition	$\log_{10}$ frequency
1	bslfd	ᕿᓚᑐᕿᕿ	FPVS	3.45
2	cnldq	ᒋᓱᑐᕿᕗ	FPVS	3.15
3	efrdq	ᓅᕿᓇᕿᕗ	FPVS	3.40
4	gvpid	ᓴᓕᕼᑎᕿ	FPVS	3.20
5	gziou	ᓴᓇᑎᐃᔾ	FPVS	3.00
6	ndrom	ᓱᕿᓇᐃᓂ	FPVS	3.95
7	oldsh	ᐃᑐᕿᓚᒐ	FPVS	3.95
8	qginh	ᕗᓴᑎᓱᒐ	FPVS	3.95
9	qrabc	ᕗᓇᓴᕿᒋ	FPVS	3.00
10	rbfgd	ᓇᕿᕿᓴᕿ	FPVS	3.65
11	tmqiz	ᖅᓂᕗᑎᓇ	FPVS	3.60
12	trbep	ᖅᓇᕿᓅᕼ	FPVS	2.95
13	vaifm	ᓕᓴᑎᕿᓂ	FPVS	3.00
14	venzm	ᓕᓅᓱᓇᓂ	FPVS	2.90
15	zmqlb	ᓇᓂᕗᑐᕿ	FPVS	3.90
16	zuqnb	ᓇᔾᕗᓱᕿ	FPVS	3.75
1	aclrz	ᓴᒋᑐᓇᓇ	Memory - Stim. vs. Foil - Foil B1	3.00
2	bepst	ᕿᓅᕼᓚᖅ	Memory - Stim. vs. Foil - Foil B1	3.00
3	chmnz	ᒋᒐᓂᓱᓇ	Memory - Stim. vs. Foil - Foil B1	3.00
4	hipsu	ᒐᑎᕼᓚᔾ	Memory - Stim. vs. Foil - Foil B1	3.95
5	cmnrv	ᒋᓂᓱᓇᓕ	Memory - Stim. vs. Foil - Foil B1	3.05
6	bhpqt	ᕿᒐᕼᕗᖅ	Memory - Stim. vs. Foil - Foil B1	3.85
7	afioz	ᓴᕿᑎᐃᓇ	Memory - Stim. vs. Foil - Foil B1	3.00
8	efoqz	ᓅᕿᐃᕗᓇ	Memory - Stim. vs. Foil - Foil B1	3.65
9	abefg	ᓴᕿᓅᕿᓴ	Memory - Stim. vs. Foil - Foil B1	3.00
10	bersz	ᕿᓅᓇᓚᓇ	Memory - Stim. vs. Foil - Foil B1	3.95
11	efimt	ᓅᕿᑎᓂᖅ	Memory - Stim. vs. Foil - Foil B1	3.00
12	ghmuz	ᓴᒐᓂᔾᓇ	Memory - Stim. vs. Foil - Foil B1	3.65
13	bgpsv	ᕿᓴᕼᓚᓕ	Memory - Stim. vs. Foil - Foil B1	3.75
14	biquv	ᕿᑎᕗᔾᓕ	Memory - Stim. vs. Foil - Foil B1	3.00
15	afhqu	ᓴᕿᒐᕗᔾ	Memory - Stim. vs. Foil - Foil B1	3.85
16	mpqsz	ᓂᕼᕗᓚᓇ	Memory - Stim. vs. Foil - Foil B1	3.90
1	bhqsu	ᕿᒐᕗᓚᔾ	Memory - Stim. vs. Foil - Foil B2	3.65
2	ginqs	ᓴᑎᓱᕗᓚ	Memory - Stim. vs. Foil - Foil B2	3.00
3	aeipz	ᓴᓅᑎᕼᓇ	Memory - Stim. vs. Foil - Foil B2	3.05
4	clmoz	ᒋᑐᓂᐃᓇ	Memory - Stim. vs. Foil - Foil B2	3.00
5	fhtuv	ᕿᒐᖅᔾᓕ	Memory - Stim. vs. Foil - Foil B2	3.45
6	cdfnu	ᒋᕿᕿᓱᔾ	Memory - Stim. vs. Foil - Foil B2	3.65
7	abchi	ᓴᕿᒋᒐᑎ	Memory - Stim. vs. Foil - Foil B2	3.85
8	glpvz	ᓴᑐᕼᓕᓇ	Memory - Stim. vs. Foil - Foil B2	3.90
9	bcmnu	ᕿᒋᓂᓱᔾ	Memory - Stim. vs. Foil - Foil B2	3.45
10	ghpqr	ᓴᒐᕼᕗᓇ	Memory - Stim. vs. Foil - Foil B2	3.70

ID	Stimulus	BACS-2	Condition	log10FREQcount
11	mpsuv	ᖘᖷᓀᓕᔭ	Memory - Stim. vs. Foil - Foil B2	3.55
12	hirtz	ᒉᑎᖇᓂᖁ	Memory - Stim. vs. Foil - Foil B2	2.95
13	dgios	ᖴᔋᑎᐅᓀ	Memory - Stim. vs. Foil - Foil B2	2.90
14	blntv	ᖅᑐᔨᓂᔭ	Memory - Stim. vs. Foil - Foil B2	3.00
15	egpqt	ᔰᔋᖷᕫᓂ	Memory - Stim. vs. Foil - Foil B2	3.70
16	mpqtv	ᖘᖷᕫᓂᔭ	Memory - Stim. vs. Foil - Foil B2	3.60
1	bdfpt	ᖅᖴᖁᖷᓂ	Memory - Foil vs. Foil - Item	3.50
2	bcfor	ᖅᑕᖁᐅᖇ	Memory - Foil vs. Foil - Item	3.00
3	efstv	ᔰᖁᓀᓂᔭ	Memory - Foil vs. Foil - Item	3.10
4	adnoq	ᔪᖴᔨᐅᕫ	Memory - Foil vs. Foil - Item	3.00
5	abdlo	ᔪᖅᖴᑐᐅ	Memory - Foil vs. Foil - Item	2.85
6	dhilz	ᖴᒉᑎᑐᖁ	Memory - Foil vs. Foil - Item	3.00
7	achlo	ᔪᑕᒉᑐᐅ	Memory - Foil vs. Foil - Item	2.85
8	abdfm	ᔪᖅᖴᖁᖘ	Memory - Foil vs. Foil - Item	3.20
9	hlmsz	ᒉᑐᖘᓀᖁ	Memory - Foil vs. Foil - Item	3.80
10	bhntz	ᖅᒉᔨᓂᖁ	Memory - Foil vs. Foil - Item	3.00
11	fhnrz	ᖁᒉᔨᖇᖁ	Memory - Foil vs. Foil - Item	3.50
12	cdehm	ᑕᖴᔰᒉᖘ	Memory - Foil vs. Foil - Item	3.00
13	eglpz	ᔰᔋᑐᖷᖁ	Memory - Foil vs. Foil - Item	3.60
14	bchrv	ᖅᑕᒉᖇᔭ	Memory - Foil vs. Foil - Item	3.60
15	bclqv	ᖅᑕᑐᕫᔭ	Memory - Foil vs. Foil - Item	3.75
16	iloqu	ᐃᑐᐅᕫᕆ	Memory - Foil vs. Foil - Item	3.00
1	chiln	ᑕᒉᐃᑐᔨ	Memory - Foil vs. Foil - Foil B1	2.70
2	ahisz	ᔪᒉᐃᓀᖁ	Memory - Foil vs. Foil - Foil B1	3.00
3	bcefr	ᖅᑕᔰᖁᖇ	Memory - Foil vs. Foil - Foil B1	3.05
4	abefs	ᔪᖅᔰᖁᓀ	Memory - Foil vs. Foil - Foil B1	3.00
5	goprz	ᔋᐅᖷᖇᖁ	Memory - Foil vs. Foil - Foil B1	3.00
6	almoz	ᔪᑐᖘᐅᖁ	Memory - Foil vs. Foil - Foil B1	3.00
7	adguv	ᔪᖴᔋᑎᔭ	Memory - Foil vs. Foil - Foil B1	3.00
8	abmpu	ᔪᖅᖘᖷᑎᕆ	Memory - Foil vs. Foil - Foil B1	3.00
9	bfhpz	ᖅᖁᒉᖷᖁ	Memory - Foil vs. Foil - Foil B1	3.90
10	demnz	ᖴᔰᖘᔨᖁ	Memory - Foil vs. Foil - Foil B1	3.00
11	alnuv	ᔪᑐᔨᑎᔭ	Memory - Foil vs. Foil - Foil B1	3.00
12	egmou	ᔰᔋᖘᐅᕆ	Memory - Foil vs. Foil - Foil B1	3.00
13	cfhrv	ᑕᖁᒉᖇᔭ	Memory - Foil vs. Foil - Foil B1	3.55
14	fimpu	ᖁᐃᖘᖷᑎᕆ	Memory - Foil vs. Foil - Foil B1	3.00
15	ahnqz	ᔪᒉᔨᕫᖁ	Memory - Foil vs. Foil - Foil B1	3.85
16	fosuz	ᖁᐅᓀᑎᖁ	Memory - Foil vs. Foil - Foil B1	3.00
1	abqsu	ᔪᖅᕫᓀᑎ	Memory - Foil vs. Foil - Foil B2	3.00
2	fgmvz	ᖁᔋᖘᔨᖁ	Memory - Foil vs. Foil - Foil B2	3.95
3	gmprv	ᔋᖘᖷᖇᔭ	Memory - Foil vs. Foil - Foil B2	3.05
4	bcgmr	ᖅᑕᔋᖘᖇ	Memory - Foil vs. Foil - Foil B2	3.75
5	imstv	ᐃᖘᓀᓂᔭ	Memory - Foil vs. Foil - Foil B2	3.00
6	chilt	ᑕᒉᐃᑐᓂ	Memory - Foil vs. Foil - Foil B2	2.75
7	anosz	ᔪᔨᐅᓀᖁ	Memory - Foil vs. Foil - Foil B2	2.90

ID	Stimulus	BACS-2	Condition	log10FREQcount
8	cdhms	ᓀᕈᒉᔨᓪ	Memory - Foil vs. Foil - Foil B2	3.15
9	ahitz	ᖶᒉᐱᓂᖁ	Memory - Foil vs. Foil - Foil B2	3.95
10	prtvz	ᕃᕮᓂᔦᖁ	Memory - Foil vs. Foil - Foil B2	3.00
11	cglos	ᓀᔭᑐᐅᓪ	Memory - Foil vs. Foil - Foil B2	2.90
12	amotv	ᖶᔝᐅᓂᔦ	Memory - Foil vs. Foil - Foil B2	2.95
13	bnops	ᖇᔨᐅᕃᓪ	Memory - Foil vs. Foil - Foil B2	2.95
14	dilsu	ᕈᐱᑐᓪᓗ	Memory - Foil vs. Foil - Foil B2	3.00
15	delnz	ᕈᔋᑐᔨᖁ	Memory - Foil vs. Foil - Foil B2	3.00
16	cfgmu	ᓀᖁᔭᔝᓗ	Memory - Foil vs. Foil - Foil B2	3.50
1	efiru	ᔋᖁᐱᕮᓗ	Discrimination - Foil B1	3.00
2	cdegq	ᓀᕈᔋᔭᕐ	Discrimination - Foil B1	3.20
3	abrtv	ᖶᖇᕮᓂᔦ	Discrimination - Foil B1	3.00
4	dhisu	ᕈᒉᐱᓪᓗ	Discrimination - Foil B1	2.95
5	afstv	ᖶᖁᓪᓂᔦ	Discrimination - Foil B1	3.25
6	dgnos	ᕈᔭᔨᐅᓪ	Discrimination - Foil B1	2.90
7	bfhqs	ᖇᖁᒉᕐᓪ	Discrimination - Foil B1	3.90
8	acdls	ᖶᓀᕈᑐᓪ	Discrimination - Foil B1	2.95
9	ehmnv	ᔋᒉᔝᔨᔦ	Discrimination - Foil B1	3.00
10	adntv	ᖶᕈᔨᓂᔦ	Discrimination - Foil B1	3.00
11	imnqz	ᐱᔝᔨᕐᖁ	Discrimination - Foil B1	3.80
12	bcmpu	ᖇᓀᔝᔭᓗ	Discrimination - Foil B1	3.10
13	chimz	ᓀᒉᐱᔝᖁ	Discrimination - Foil B1	2.90
14	cdrtz	ᓀᕈᕮᓂᖁ	Discrimination - Foil B1	3.00
15	cdisv	ᓀᕈᐱᓪᔦ	Discrimination - Foil B1	3.00
16	abgmu	ᖶᖇᔭᔝᓗ	Discrimination - Foil B1	3.25
1	bdert	ᖇᕈᔋᕮᓂ	Discrimination - Foil B2	2.90
2	degoz	ᕈᔋᔭᐅᖁ	Discrimination - Foil B2	2.95
3	epstz	ᔋᔭᓪᓂᖁ	Discrimination - Foil B2	3.00
4	ademp	ᖶᕈᔋᔝᔭ	Discrimination - Foil B2	2.95
5	nptuz	ᔨᔭᓂᓗᖁ	Discrimination - Foil B2	3.50
6	ilstu	ᐱᑐᓪᓂᓗ	Discrimination - Foil B2	2.95
7	chnor	ᓀᒉᔨᐅᕮ	Discrimination - Foil B2	2.90
8	cfisv	ᓀᖁᐱᓪᔦ	Discrimination - Foil B2	3.00
9	fmpsz	ᖁᔝᔭᓪᖁ	Discrimination - Foil B2	3.90
10	adegv	ᖶᕈᔋᔭᔦ	Discrimination - Foil B2	3.00
11	denpr	ᕈᔋᔨᔭᕮ	Discrimination - Foil B2	3.90
12	deimr	ᕈᔋᐱᔝᕮ	Discrimination - Foil B2	3.00
13	einrz	ᔋᐱᔨᕮᖁ	Discrimination - Foil B2	2.95
14	emuvz	ᔋᔝᓂᔦᖁ	Discrimination - Foil B2	3.60
15	inotu	ᐱᔨᐅᓂᓂᓗ	Discrimination - Foil B2	3.00
16	cnuvz	ᓀᔨᓂᔦᖁ	Discrimination - Foil B2	2.70
1	hmnqr	ᒐᔝᔨᕐᕮ	Discrimination - Foil B3	3.90
2	efgos	ᔋᖁᔭᐅᓪ	Discrimination - Foil B3	3.00
3	befgt	ᖇᔋᖁᔭᓂ	Discrimination - Foil B3	3.00
4	aeiuz	ᖶᔋᐱᓂᖁ	Discrimination - Foil B3	3.00

ID	Stimulus	BACS-2	Condition	log10FREQcount
5	deqtz	𐤓𐤔𐤉𐤊𐤋	Discrimination - Foil B3	3.00
6	abchp	𐤃𐤒𐤏𐤋𐤌	Discrimination - Foil B3	3.00
7	efhpq	𐤔𐤕𐤋𐤌𐤉	Discrimination - Foil B3	3.90
8	coptz	𐤏𐤁𐤌𐤊𐤋	Discrimination - Foil B3	2.95
9	behms	𐤒𐤔𐤋𐤎𐤓	Discrimination - Foil B3	3.00
10	flnot	𐤕𐤁𐤍𐤁𐤊	Discrimination - Foil B3	2.90
11	einru	𐤔𐤊𐤍𐤃𐤅	Discrimination - Foil B3	3.00
12	cdfln	𐤏𐤓𐤕𐤁𐤍	Discrimination - Foil B3	3.55
13	gilst	𐤂𐤊𐤁𐤓𐤊	Discrimination - Foil B3	2.90
14	airst	𐤃𐤊𐤃𐤓𐤊	Discrimination - Foil B3	2.85
15	iotuv	𐤊𐤁𐤊𐤅𐤄	Discrimination - Foil B3	3.00
16	cinpz	𐤏𐤊𐤍𐤌𐤋	Discrimination - Foil B3	2.95

Table S3. Word stimuli - Experiment 2

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
1	suono	FPVS - Stimuli - Set 1	3678	6.565612	1.00
2	posta	FPVS - Stimuli - Set 1	3277	6.515476	1.00
3	fetta	FPVS - Stimuli - Set 1	918	5.962843	1.00
4	somma	FPVS - Stimuli - Set 1	1269	6.103462	1.10
5	zuppa	FPVS - Stimuli - Set 1	1807	6.256958	1.30
6	bimbo	FPVS - Stimuli - Set 1	1699	6.230193	1.35
7	costo	FPVS - Stimuli - Set 1	1797	6.254548	1.00
8	gomma	FPVS - Stimuli - Set 1	1768	6.247482	1.40
9	frigo	FPVS - Stimuli - Set 1	1475	6.168792	1.20
10	spina	FPVS - Stimuli - Set 1	1374	6.137987	1.00
11	torre	FPVS - Stimuli - Set 1	2255	6.353147	1.00
12	dosso	FPVS - Stimuli - Set 1	2362	6.373280	1.00
13	mazza	FPVS - Stimuli - Set 1	1256	6.098990	1.00
14	prato	FPVS - Stimuli - Set 1	944	5.974972	1.00
15	mucca	FPVS - Stimuli - Set 1	1155	6.062582	1.30
16	pollo	FPVS - Stimuli - Set 1	3761	6.575303	1.00
1	polso	FPVS - Stimuli - Set 2	1733	6.238799	1.45
2	salto	FPVS - Stimuli - Set 2	3010	6.478566	1.00
3	crema	FPVS - Stimuli - Set 2	1539	6.187239	1.00
4	mosca	FPVS - Stimuli - Set 2	1902	6.279211	1.40
5	pasta	FPVS - Stimuli - Set 2	1267	6.102777	1.00
6	ombra	FPVS - Stimuli - Set 2	2707	6.432488	1.45
7	dieta	FPVS - Stimuli - Set 2	1138	6.056142	1.00
8	scala	FPVS - Stimuli - Set 2	2514	6.400365	1.00
9	conte	FPVS - Stimuli - Set 2	2118	6.325926	1.00
10	legno	FPVS - Stimuli - Set 2	2299	6.361539	1.00
11	forno	FPVS - Stimuli - Set 2	2299	6.361539	1.00
12	gusto	FPVS - Stimuli - Set 2	2528	6.402777	1.00
13	epoca	FPVS - Stimuli - Set 2	2418	6.383456	1.60
14	miele	FPVS - Stimuli - Set 2	2064	6.314710	1.45
15	bosco	FPVS - Stimuli - Set 2	2248	6.351796	1.15
16	fiato	FPVS - Stimuli - Set 2	1815	6.258877	1.00
1	sposo	Memory - Stim. vs. Foil - Foil - Set 1	1734	6.239049	1.35
2	ansia	Memory - Stim. vs. Foil - Foil - Set 1	1275	6.105510	1.50
3	furto	Memory - Stim. vs. Foil - Foil - Set 1	2634	6.420616	1.10
4	delia	Memory - Stim. vs. Foil - Foil - Set 1	1594	6.202488	1.00
5	bugia	Memory - Stim. vs. Foil - Foil - Set 1	2553	6.407051	1.50
6	copia	Memory - Stim. vs. Foil - Foil - Set 1	3045	6.483587	1.00
7	socio	Memory - Stim. vs. Foil - Foil - Set 1	2715	6.433770	1.00
8	sparo	Memory - Stim. vs. Foil - Foil - Set 1	2688	6.429429	1.00
9	ladro	Memory - Stim. vs. Foil - Foil - Set 1	3079	6.488410	1.20
10	monte	Memory - Stim. vs. Foil - Foil - Set 1	1599	6.203848	1.00

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
11	fango	Memory - Stim. vs. Foil - Foil - Set 1	1322	6.121231	1.00
12	morso	Memory - Stim. vs. Foil - Foil - Set 1	1987	6.298198	1.00
13	barba	Memory - Stim. vs. Foil - Foil - Set 1	1832	6.262925	1.05
14	villa	Memory - Stim. vs. Foil - Foil - Set 1	1433	6.156246	1.00
15	mensa	Memory - Stim. vs. Foil - Foil - Set 1	1076	6.031812	1.20
16	falla	Memory - Stim. vs. Foil - Foil - Set 1	2255	6.353147	1.00
1	cuoco	Memory - Stim. vs. Foil - Foil - Set 2	980	5.991226	1.35
2	valle	Memory - Stim. vs. Foil - Foil - Set 2	1327	6.122871	1.00
3	crisi	Memory - Stim. vs. Foil - Foil - Set 2	3171	6.501196	1.40
4	prete	Memory - Stim. vs. Foil - Foil - Set 2	2656	6.424228	1.00
5	circo	Memory - Stim. vs. Foil - Foil - Set 2	1405	6.147676	1.25
6	pacco	Memory - Stim. vs. Foil - Foil - Set 2	1754	6.244030	1.00
7	trono	Memory - Stim. vs. Foil - Foil - Set 2	1122	6.049993	1.00
8	penna	Memory - Stim. vs. Foil - Foil - Set 2	2134	6.329194	1.00
9	targa	Memory - Stim. vs. Foil - Foil - Set 2	1574	6.197005	1.00
10	frase	Memory - Stim. vs. Foil - Foil - Set 2	2118	6.325926	1.10
11	salsa	Memory - Stim. vs. Foil - Foil - Set 2	1702	6.230960	1.15
12	ritmo	Memory - Stim. vs. Foil - Foil - Set 2	1622	6.210051	1.35
13	becco	Memory - Stim. vs. Foil - Foil - Set 2	2442	6.387746	1.30
14	tazza	Memory - Stim. vs. Foil - Foil - Set 2	2843	6.453777	1.00
15	ruota	Memory - Stim. vs. Foil - Foil - Set 2	1591	6.201670	1.50
16	tomba	Memory - Stim. vs. Foil - Foil - Set 2	2655	6.424065	1.30
1	massa	Memory - Foil vs. Foil - Stimuli	1741	6.240799	1.00
2	dente	Memory - Foil vs. Foil - Stimuli	1036	6.015360	1.00
3	danza	Memory - Foil vs. Foil - Stimuli	1774	6.248954	1.00
4	soldo	Memory - Foil vs. Foil - Stimuli	1167	6.067071	1.00
5	merce	Memory - Foil vs. Foil - Stimuli	1375	6.138303	1.00
6	pasto	Memory - Foil vs. Foil - Stimuli	1758	6.245019	1.00
7	cella	Memory - Foil vs. Foil - Stimuli	3320	6.521138	1.00
8	preda	Memory - Foil vs. Foil - Stimuli	1170	6.068186	1.05
9	fiore	Memory - Foil vs. Foil - Stimuli	1764	6.246499	1.10
10	spesa	Memory - Foil vs. Foil - Stimuli	1517	6.180986	1.15
11	norma	Memory - Foil vs. Foil - Stimuli	991	5.996074	1.25
12	succo	Memory - Foil vs. Foil - Stimuli	1789	6.252610	1.30
13	ciclo	Memory - Foil vs. Foil - Stimuli	1064	6.026942	1.35
14	talpa	Memory - Foil vs. Foil - Stimuli	937	5.971740	1.40
15	drago	Memory - Foil vs. Foil - Stimuli	1382	6.140508	1.45
16	banco	Memory - Foil vs. Foil - Stimuli	1681	6.225568	1.00
1	ruolo	Memory - Foil vs. Foil - Foil	3314	6.520353	1.10
2	lotta	Memory - Foil vs. Foil - Foil	3533	6.548144	1.00
3	croce	Memory - Foil vs. Foil - Foil	1849	6.266937	1.60
4	curva	Memory - Foil vs. Foil - Foil	979	5.990783	1.30
5	pugno	Memory - Foil vs. Foil - Foil	3267	6.514149	1.45
6	botta	Memory - Foil vs. Foil - Foil	1040	6.017033	1.00
7	grano	Memory - Foil vs. Foil - Foil	925	5.966142	1.00

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
8	disco	Memory - Foil vs. Foil - Foil	2133	6.328991	1.00
9	animo	Memory - Foil vs. Foil - Foil	945	5.975432	1.10
10	sorte	Memory - Foil vs. Foil - Foil	1264	6.101747	1.00
11	scudo	Memory - Foil vs. Foil - Foil	1155	6.062582	1.35
12	multa	Memory - Foil vs. Foil - Foil	989	5.995196	1.05
13	quota	Memory - Foil vs. Foil - Foil	1099	6.040998	1.30
14	zaino	Memory - Foil vs. Foil - Foil	1036	6.015360	1.55
15	ferro	Memory - Foil vs. Foil - Foil	2238	6.349860	1.00
16	vasca	Memory - Foil vs. Foil - Foil	2017	6.304706	1.25
1	cuoco	Discrimination - Foil - Set 1	980	5.991226	1.35
2	valle	Discrimination - Foil - Set 1	1327	6.122871	1.00
3	crisi	Discrimination - Foil - Set 1	3171	6.501196	1.40
4	prete	Discrimination - Foil - Set 1	2656	6.424228	1.00
5	circo	Discrimination - Foil - Set 1	1405	6.147676	1.25
6	pacco	Discrimination - Foil - Set 1	1754	6.244030	1.00
7	trono	Discrimination - Foil - Set 1	1122	6.049993	1.00
8	penna	Discrimination - Foil - Set 1	2134	6.329194	1.00
9	targa	Discrimination - Foil - Set 1	1574	6.197005	1.00
10	frase	Discrimination - Foil - Set 1	2118	6.325926	1.10
11	salsa	Discrimination - Foil - Set 1	1702	6.230960	1.15
12	ritmo	Discrimination - Foil - Set 1	1622	6.210051	1.35
13	becco	Discrimination - Foil - Set 1	2442	6.387746	1.30
14	tazza	Discrimination - Foil - Set 1	2843	6.453777	1.00
15	ruota	Discrimination - Foil - Set 1	1591	6.201670	1.50
16	tomba	Discrimination - Foil - Set 1	2655	6.424065	1.30
1	sposo	Discrimination - Foil - Set 2	1734	6.239049	1.35
2	ansia	Discrimination - Foil - Set 2	1275	6.105510	1.50
3	furto	Discrimination - Foil - Set 2	2634	6.420616	1.10
4	delia	Discrimination - Foil - Set 2	1594	6.202488	1.00
5	bugia	Discrimination - Foil - Set 2	2553	6.407051	1.50
6	copia	Discrimination - Foil - Set 2	3045	6.483587	1.00
7	socio	Discrimination - Foil - Set 2	2715	6.433770	1.00
8	sparo	Discrimination - Foil - Set 2	2688	6.429429	1.00
9	ladro	Discrimination - Foil - Set 2	3079	6.488410	1.20
10	monte	Discrimination - Foil - Set 2	1599	6.203848	1.00
11	fango	Discrimination - Foil - Set 2	1322	6.121231	1.00
12	morso	Discrimination - Foil - Set 2	1987	6.298198	1.00
13	barba	Discrimination - Foil - Set 2	1832	6.262925	1.05
14	villa	Discrimination - Foil - Set 2	1433	6.156246	1.00
15	mensa	Discrimination - Foil - Set 2	1076	6.031812	1.20
16	falla	Discrimination - Foil - Set 2	2255	6.353147	1.00

Table S4. Pseudofont stimuli - Experiment 2

ID	Stimulus	BACS-2	Condition	$\log_{10}$ frequency
1	abqsu	⊬ᖘᘐᓀᓚ	FPVS - Stimuli - Set 1	3.00
2	fgmvz	ᖁᕽᖇᔑᖆ	FPVS - Stimuli - Set 1	3.95
3	gmprv	ᕽᖇᕮᕘᔑ	FPVS - Stimuli - Set 1	3.05
4	bcgmr	ᘐᑕᕽᖇᕘ	FPVS - Stimuli - Set 1	3.75
5	imstv	∩ᖇᓀᔓᔑ	FPVS - Stimuli - Set 1	3.00
6	chilt	ᑕᒉ∩ᑐᔓ	FPVS - Stimuli - Set 1	2.75
7	anosz	⊬ᔑᐅᓀᖆ	FPVS - Stimuli - Set 1	2.90
8	cdhms	ᑕᖘᒉᖇᓀ	FPVS - Stimuli - Set 1	3.15
9	ahitz	⊬ᒉ∩ᔓᖆ	FPVS - Stimuli - Set 1	2.95
10	prtvz	ᕮᕘᔓᔑᖆ	FPVS - Stimuli - Set 1	3.00
11	cglos	ᑕᕽᑐᐅᓀ	FPVS - Stimuli - Set 1	2.90
12	amotv	⊬ᖇᐅᔓᔑ	FPVS - Stimuli - Set 1	2.95
13	bnops	ᘐᔑᐅᕮᓀ	FPVS - Stimuli - Set 1	2.95
14	dilsu	ᖘ∩ᑐᓀᓚ	FPVS - Stimuli - Set 1	3.00
15	delnz	ᖘᓄᑐᔑᖆ	FPVS - Stimuli - Set 1	3.00
16	cfgmu	ᑕᖁᕽᖇᓚ	FPVS - Stimuli - Set 1	3.50
1	bslfd	ᘐᓀᑐᖁᖘ	FPVS - Stimuli - Set 2	3.45
2	cnldq	ᑕᔑᑐᖘᘏ	FPVS - Stimuli - Set 2	3.15
3	efrdq	ᓄᖁᕈᖘᘏ	FPVS - Stimuli - Set 2	3.40
4	gvpid	ᕽᔑᕮ∩ᖘ	FPVS - Stimuli - Set 2	3.20
5	gziou	ᕽᖆ∩ᐅᓚ	FPVS - Stimuli - Set 2	3.00
6	ndrom	ᔑᖘᕈᐅᖇ	FPVS - Stimuli - Set 2	2.95
7	oldsh	ᐅᑐᖘᓀᑕ	FPVS - Stimuli - Set 2	2.95
8	qginh	ᘏᕽ∩ᔑᑕ	FPVS - Stimuli - Set 2	2.95
9	qrabc	ᘏᕈᓄᘐᑕ	FPVS - Stimuli - Set 2	3.00
10	rbfgd	ᕈᘐᖁᕽᖘ	FPVS - Stimuli - Set 2	3.65
11	tmqiz	ᔓᖇᘏ∩ᖆ	FPVS - Stimuli - Set 2	3.60
12	trbep	ᔓᕈᘐᓄᕮ	FPVS - Stimuli - Set 2	3.95
13	vaifm	ᔑᓀ∩ᖁᖇ	FPVS - Stimuli - Set 2	3.00
14	venzm	ᔑᓄᔑᖆᖇ	FPVS - Stimuli - Set 2	3.90
15	zmqlb	ᖆᖇᘏᑐᘐ	FPVS - Stimuli - Set 2	3.90
16	zuqnb	ᖆᓚᘏᔑᘐ	FPVS - Stimuli - Set 2	3.75
1	befgt	ᘐᓄᖁᕽᔓ	Memory - Stim. vs. Foil - Set 1	3.00
2	adegv	⊬ᖘᓄᕽᔑ	Memory - Stim. vs. Foil - Set 1	3.00
3	bcmnu	ᘐᑕᖇᔑᓚ	Memory - Stim. vs. Foil - Set 1	3.45
4	bdert	ᘐᖘᓄᕈᔓ	Memory - Stim. vs. Foil - Set 1	2.90
5	cfisv	ᑕᖁ∩ᒉᔑ	Memory - Stim. vs. Foil - Set 1	3.00
6	clmoz	ᑕᑐᖇᐅᖆ	Memory - Stim. vs. Foil - Set 1	3.00
7	deqtz	ᖘᓄᘏᔓᖆ	Memory - Stim. vs. Foil - Set 1	3.00
8	epstz	ᓄᕮᒉᔓᖆ	Memory - Stim. vs. Foil - Set 1	3.00
9	efiru	ᓄᖁ∩ᕈᓚ	Memory - Stim. vs. Foil - Set 1	3.00
10	egpqt	ᓄᖁᕮᘏᔓ	Memory - Stim. vs. Foil - Set 1	3.70

ID	Stimulus	BACS-2	Condition	$\log_{10}$ frequency
11	fmpsz	ᖁᖘᒪᓀᖅ	Memory - Stim. vs. Foil - Set 1	3.90
12	ghpqr	ᕐᑎᒪᕮᕆ	Memory - Stim. vs. Foil - Set 1	3.70
13	ginqs	ᕐᑎᓂᕮᓀ	Memory - Stim. vs. Foil - Set 1	3.00
14	imnqz	ᐃᖘᓂᕮᖅ	Memory - Stim. vs. Foil - Set 1	3.80
15	mpqtv	ᖘᒪᕮᐲᓗ	Memory - Stim. vs. Foil - Set 1	3.60
16	nptuz	ᓂᒪᐲᑎᖅ	Memory - Stim. vs. Foil - Set 1	3.50
1	aclrz	ᖃᑕᑐᕆᖅ	Memory - Stim. vs. Foil - Set 2	3.00
2	aeipz	ᖃᓄᐃᒪᖅ	Memory - Stim. vs. Foil - Set 2	3.05
3	chmnz	ᑕᑎᖘᓂᖅ	Memory - Stim. vs. Foil - Set 2	3.00
4	hipsu	ᑌᐃᒪᓀᔪ	Memory - Stim. vs. Foil - Set 2	2.95
5	cmnrv	ᑕᖘᓂᕆᓗ	Memory - Stim. vs. Foil - Set 2	3.05
6	bhpqt	ᖵᑎᒪᕮᐲ	Memory - Stim. vs. Foil - Set 2	3.85
7	afioz	ᖃᖁᐃᓄᖅ	Memory - Stim. vs. Foil - Set 2	3.00
8	efoqz	ᓄᖁᓄᕮᖅ	Memory - Stim. vs. Foil - Set 2	3.65
9	abefg	ᖃᖵᓄᖁᕐᑐ	Memory - Stim. vs. Foil - Set 2	3.00
10	bersz	ᖵᓄᕆᓀᖅ	Memory - Stim. vs. Foil - Set 2	2.95
11	efimt	ᓄᖁᐃᖘᐲ	Memory - Stim. vs. Foil - Set 2	3.00
12	ghmuz	ᕐᑎᖘᔪᖅ	Memory - Stim. vs. Foil - Set 2	3.65
13	bgpsv	ᖵᕐᒪᓀᓗ	Memory - Stim. vs. Foil - Set 2	3.75
14	biquv	ᖵᐃᕮᔪᓗ	Memory - Stim. vs. Foil - Set 2	3.00
15	afhqu	ᖃᖁᑎᕮᔪ	Memory - Stim. vs. Foil - Set 2	3.85
16	mpqsz	ᖘᒪᕮᓀᖅ	Memory - Stim. vs. Foil - Set 2	3.90
1	bdfpt	ᖵᖠᖁᒪᐲ	Memory - Foil vs. Foil - Stimuli	3.50
2	bcfor	ᖵᑕᖁᓄᕆ	Memory - Foil vs. Foil - Stimuli	3.00
3	efstv	ᓄᖁᓀᐲᓗ	Memory - Foil vs. Foil - Stimuli	3.10
4	adnoq	ᖃᖠᓂᓄᕮ	Memory - Foil vs. Foil - Stimuli	3.00
5	abdlo	ᖃᖵᖠᑐᓄ	Memory - Foil vs. Foil - Stimuli	2.85
6	dhilz	ᖠᑌᐃᑐᖅ	Memory - Foil vs. Foil - Stimuli	3.00
7	achlo	ᖃᑕᑌᑐᓄ	Memory - Foil vs. Foil - Stimuli	2.85
8	abdfm	ᖃᖵᖠᖁᖘ	Memory - Foil vs. Foil - Stimuli	3.20
9	hlmsz	ᑌᑐᖘᓀᖅ	Memory - Foil vs. Foil - Stimuli	3.80
10	bhntz	ᖵᑌᓂᐲᖅ	Memory - Foil vs. Foil - Stimuli	3.00
11	fhnrz	ᖁᑌᓂᕆᖅ	Memory - Foil vs. Foil - Stimuli	3.50
12	cdehm	ᑕᖠᓄᑌᖘ	Memory - Foil vs. Foil - Stimuli	3.00
13	eglpz	ᓄᕐᑐᒪᖅ	Memory - Foil vs. Foil - Stimuli	3.60
14	bchrv	ᖵᑕᑌᕆᓗ	Memory - Foil vs. Foil - Stimuli	3.60
15	bclqv	ᖵᑕᑐᖠᓗ	Memory - Foil vs. Foil - Stimuli	3.75
16	iloqu	ᐃᑐᓄᖠᔪ	Memory - Foil vs. Foil - Stimuli	3.00
1	degoz	ᖠᓄᕐᓄᖅ	Memory - Foil vs. Foil - Foil	2.95
2	ahisz	ᖃᑌᐃᓀᖅ	Memory - Foil vs. Foil - Foil	3.00
3	ilstu	ᐃᑐᓀᐲᔪ	Memory - Foil vs. Foil - Foil	2.95
4	abefs	ᖃᖵᓄᖁᓀ	Memory - Foil vs. Foil - Foil	3.00
5	goprz	ᕐᓄᒪᕆᖅ	Memory - Foil vs. Foil - Foil	3.00
6	almoz	ᖃᑐᖘᓄᖅ	Memory - Foil vs. Foil - Foil	3.00
7	adguv	ᖃᖠᕐᔪᓗ	Memory - Foil vs. Foil - Foil	3.00

ID	Stimulus	BACS-2	Condition	$\log_{10}$ frequency
8	abmpu	ᓯᕿᓂᕼᓕ	Memory - Foil vs. Foil - Foil	3.00
9	bfhpz	ᕿᖁᒐᕼᓇ	Memory - Foil vs. Foil - Foil	3.90
10	demnz	ᕿᓅᓂᓱᓇ	Memory - Foil vs. Foil - Foil	3.00
11	alnuv	ᓯᑐᓱᓕᓘ	Memory - Foil vs. Foil - Foil	3.00
12	egmou	ᓅᓭᓂᐃᓕ	Memory - Foil vs. Foil - Foil	3.00
13	cfhrv	ᒋᖁᒐᒪᓘ	Memory - Foil vs. Foil - Foil	3.55
14	fimpu	ᖁᐱᓂᕼᓕ	Memory - Foil vs. Foil - Foil	3.00
15	ahnqz	ᓯᒐᓱᕿᓇ	Memory - Foil vs. Foil - Foil	3.85
16	fosuz	ᖁᐃᓚᓕᓇ	Memory - Foil vs. Foil - Foil	3.00
1	aclrz	ᓯᒋᑐᒪᓇ	Discrimination - Foil - Set 1	3.00
2	aeipz	ᓯᓅᐱᕼᓇ	Discrimination - Foil - Set 1	3.05
3	chmnz	ᒋᒐᓂᓱᓇ	Discrimination - Foil - Set 1	3.00
4	hipsu	ᒐᐱᕼᓚᓕ	Discrimination - Foil - Set 1	3.95
5	cmnrv	ᒋᓂᓱᒪᓘ	Discrimination - Foil - Set 1	3.05
6	bhpqt	ᕿᒐᕼᕿᕝ	Discrimination - Foil - Set 1	3.85
7	afioz	ᓯᖁᐱᐃᓇ	Discrimination - Foil - Set 1	3.00
8	efoqz	ᓅᖁᐃᕿᓇ	Discrimination - Foil - Set 1	3.65
9	abefg	ᓯᕿᓅᖁᓭ	Discrimination - Foil - Set 1	3.00
10	bersz	ᕿᓅᒪᓚᓇ	Discrimination - Foil - Set 1	2.95
11	efimt	ᓅᖁᐱᓂᕝ	Discrimination - Foil - Set 1	3.00
12	ghmuz	ᓭᒐᓂᓕᓇ	Discrimination - Foil - Set 1	3.65
13	bgpsv	ᕿᓭᕼᓚᓘ	Discrimination - Foil - Set 1	3.75
14	biquv	ᕿᐱᕿᓕᓘ	Discrimination - Foil - Set 1	3.00
15	afhqu	ᓯᖁᒐᕿᓕ	Discrimination - Foil - Set 1	3.85
16	mpqsz	ᓂᕼᕿᓚᓇ	Discrimination - Foil - Set 1	3.90
1	befgt	ᕿᓅᖁᓭᕝ	Discrimination - Foil - Set 2	3.00
2	adegv	ᓯᕿᓅᓭᓘ	Discrimination - Foil - Set 2	3.00
3	bcmnu	ᕿᒋᓂᓱᓕ	Discrimination - Foil - Set 2	3.45
4	bdert	ᕿᕿᓅᒪᕝ	Discrimination - Foil - Set 2	2.90
5	cfisv	ᒋᖁᐱᓚᓘ	Discrimination - Foil - Set 2	3.00
6	clmoz	ᒋᑐᓂᐃᓇ	Discrimination - Foil - Set 2	3.00
7	deqtz	ᕿᓅᕿᕝᓇ	Discrimination - Foil - Set 2	3.00
8	epstz	ᓅᕼᓚᕝᓇ	Discrimination - Foil - Set 2	3.00
9	efiru	ᓅᖁᐱᒪᓕ	Discrimination - Foil - Set 2	3.00
10	egpqt	ᓅᓭᕼᕿᕝ	Discrimination - Foil - Set 2	3.70
11	fmpsz	ᖁᓂᕼᓚᓇ	Discrimination - Foil - Set 2	3.90
12	ghpqr	ᓭᒐᕼᕿᒪ	Discrimination - Foil - Set 2	3.70
13	ginqs	ᓭᐱᓱᕿᓚ	Discrimination - Foil - Set 2	3.00
14	imnqz	ᐱᓂᓱᕿᓇ	Discrimination - Foil - Set 2	3.80
15	mpqtv	ᓂᕼᕿᕝᓘ	Discrimination - Foil - Set 2	3.60
16	nptuz	ᓱᕼᕝᓕᓇ	Discrimination - Foil - Set 2	3.50

Table S5. Word stimuli

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
1	suono	FPVS - Stimuli - Set 1	3678	6.565612	1.00
2	posta	FPVS - Stimuli - Set 1	3277	6.515476	1.00
3	fetta	FPVS - Stimuli - Set 1	918	5.962843	1.00
4	somma	FPVS - Stimuli - Set 1	1269	6.103462	1.10
5	zuppa	FPVS - Stimuli - Set 1	1807	6.256958	1.30
6	bimbo	FPVS - Stimuli - Set 1	1699	6.230193	1.35
7	costo	FPVS - Stimuli - Set 1	1797	6.254548	1.00
8	gomma	FPVS - Stimuli - Set 1	1768	6.247482	1.40
9	frigo	FPVS - Stimuli - Set 1	1475	6.168792	1.20
10	spina	FPVS - Stimuli - Set 1	1374	6.137987	1.00
11	torre	FPVS - Stimuli - Set 1	2255	6.353147	1.00
12	dosso	FPVS - Stimuli - Set 1	2362	6.373280	1.00
13	mazza	FPVS - Stimuli - Set 1	1256	6.098990	1.00
14	prato	FPVS - Stimuli - Set 1	944	5.974972	1.00
15	mucca	FPVS - Stimuli - Set 1	1155	6.062582	1.30
16	pollo	FPVS - Stimuli - Set 1	3761	6.575303	1.00
1	polso	FPVS - Stimuli - Set 2	1733	6.238799	1.45
2	salto	FPVS - Stimuli - Set 2	3010	6.478566	1.00
3	crema	FPVS - Stimuli - Set 2	1539	6.187239	1.00
4	mosca	FPVS - Stimuli - Set 2	1902	6.279211	1.40
5	pasta	FPVS - Stimuli - Set 2	1267	6.102777	1.00
6	ombra	FPVS - Stimuli - Set 2	2707	6.432488	1.45
7	dieta	FPVS - Stimuli - Set 2	1138	6.056142	1.00
8	scala	FPVS - Stimuli - Set 2	2514	6.400365	1.00
9	conte	FPVS - Stimuli - Set 2	2118	6.325926	1.00
10	legno	FPVS - Stimuli - Set 2	2299	6.361539	1.00
11	forno	FPVS - Stimuli - Set 2	2299	6.361539	1.00
12	gusto	FPVS - Stimuli - Set 2	2528	6.402777	1.00
13	epoca	FPVS - Stimuli - Set 2	2418	6.383456	1.60
14	miele	FPVS - Stimuli - Set 2	2064	6.314710	1.45
15	bosco	FPVS - Stimuli - Set 2	2248	6.351796	1.15
16	fiato	FPVS - Stimuli - Set 2	1815	6.258877	1.00
1	sposo	Memory - Stim. vs. Foil - Foil - Set 1	1734	6.239049	1.35
2	ansia	Memory - Stim. vs. Foil - Foil - Set 1	1275	6.105510	1.50
3	furto	Memory - Stim. vs. Foil - Foil - Set 1	2634	6.420616	1.10
4	delia	Memory - Stim. vs. Foil - Foil - Set 1	1594	6.202488	1.00
5	bugia	Memory - Stim. vs. Foil - Foil - Set 1	2553	6.407051	1.50
6	copia	Memory - Stim. vs. Foil - Foil - Set 1	3045	6.483587	1.00
7	socio	Memory - Stim. vs. Foil - Foil - Set 1	2715	6.433770	1.00
8	sparo	Memory - Stim. vs. Foil - Foil - Set 1	2688	6.429429	1.00
9	ladro	Memory - Stim. vs. Foil - Foil - Set 1	3079	6.488410	1.20
10	monte	Memory - Stim. vs. Foil - Foil - Set 1	1599	6.203848	1.00

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
11	fango	Memory - Stim. vs. Foil - Foil - Set 1	1322	6.121231	1.00
12	morso	Memory - Stim. vs. Foil - Foil - Set 1	1987	6.298198	1.00
13	barba	Memory - Stim. vs. Foil - Foil - Set 1	1832	6.262925	1.05
14	villa	Memory - Stim. vs. Foil - Foil - Set 1	1433	6.156246	1.00
15	mensa	Memory - Stim. vs. Foil - Foil - Set 1	1076	6.031812	1.20
16	falla	Memory - Stim. vs. Foil - Foil - Set 1	2255	6.353147	1.00
1	cuoco	Memory - Stim. vs. Foil - Foil - Set 2	980	5.991226	1.35
2	valle	Memory - Stim. vs. Foil - Foil - Set 2	1327	6.122871	1.00
3	crisi	Memory - Stim. vs. Foil - Foil - Set 2	3171	6.501196	1.40
4	prete	Memory - Stim. vs. Foil - Foil - Set 2	2656	6.424228	1.00
5	circo	Memory - Stim. vs. Foil - Foil - Set 2	1405	6.147676	1.25
6	pacco	Memory - Stim. vs. Foil - Foil - Set 2	1754	6.244030	1.00
7	trono	Memory - Stim. vs. Foil - Foil - Set 2	1122	6.049993	1.00
8	penna	Memory - Stim. vs. Foil - Foil - Set 2	2134	6.329194	1.00
9	targa	Memory - Stim. vs. Foil - Foil - Set 2	1574	6.197005	1.00
10	frase	Memory - Stim. vs. Foil - Foil - Set 2	2118	6.325926	1.10
11	salsa	Memory - Stim. vs. Foil - Foil - Set 2	1702	6.230960	1.15
12	ritmo	Memory - Stim. vs. Foil - Foil - Set 2	1622	6.210051	1.35
13	becco	Memory - Stim. vs. Foil - Foil - Set 2	2442	6.387746	1.30
14	tazza	Memory - Stim. vs. Foil - Foil - Set 2	2843	6.453777	1.00
15	ruota	Memory - Stim. vs. Foil - Foil - Set 2	1591	6.201670	1.50
16	tomba	Memory - Stim. vs. Foil - Foil - Set 2	2655	6.424065	1.30
1	massa	Memory - Foil vs. Foil - Stimuli	1741	6.240799	1.00
2	dente	Memory - Foil vs. Foil - Stimuli	1036	6.015360	1.00
3	danza	Memory - Foil vs. Foil - Stimuli	1774	6.248954	1.00
4	soldo	Memory - Foil vs. Foil - Stimuli	1167	6.067071	1.00
5	merce	Memory - Foil vs. Foil - Stimuli	1375	6.138303	1.00
6	pasto	Memory - Foil vs. Foil - Stimuli	1758	6.245019	1.00
7	cella	Memory - Foil vs. Foil - Stimuli	3320	6.521138	1.00
8	preda	Memory - Foil vs. Foil - Stimuli	1170	6.068186	1.05
9	fiore	Memory - Foil vs. Foil - Stimuli	1764	6.246499	1.10
10	spesa	Memory - Foil vs. Foil - Stimuli	1517	6.180986	1.15
11	norma	Memory - Foil vs. Foil - Stimuli	991	5.996074	1.25
12	succo	Memory - Foil vs. Foil - Stimuli	1789	6.252610	1.30
13	ciclo	Memory - Foil vs. Foil - Stimuli	1064	6.026942	1.35
14	talpa	Memory - Foil vs. Foil - Stimuli	937	5.971740	1.40
15	drago	Memory - Foil vs. Foil - Stimuli	1382	6.140508	1.45
16	banco	Memory - Foil vs. Foil - Stimuli	1681	6.225568	1.00
1	ruolo	Memory - Foil vs. Foil - Foil	3314	6.520353	1.10
2	lotta	Memory - Foil vs. Foil - Foil	3533	6.548144	1.00
3	croce	Memory - Foil vs. Foil - Foil	1849	6.266937	1.60
4	curva	Memory - Foil vs. Foil - Foil	979	5.990783	1.30
5	pugno	Memory - Foil vs. Foil - Foil	3267	6.514149	1.45
6	botta	Memory - Foil vs. Foil - Foil	1040	6.017033	1.00
7	grano	Memory - Foil vs. Foil - Foil	925	5.966142	1.00

ID	Stimulus	Condition	FREQcount	Zipf	OLD20
8	disco	Memory - Foil vs. Foil - Foil	2133	6.328991	1.00
9	animo	Memory - Foil vs. Foil - Foil	945	5.975432	1.10
10	sorte	Memory - Foil vs. Foil - Foil	1264	6.101747	1.00
11	scudo	Memory - Foil vs. Foil - Foil	1155	6.062582	1.35
12	multa	Memory - Foil vs. Foil - Foil	989	5.995196	1.05
13	quota	Memory - Foil vs. Foil - Foil	1099	6.040998	1.30
14	zaino	Memory - Foil vs. Foil - Foil	1036	6.015360	1.55
15	ferro	Memory - Foil vs. Foil - Foil	2238	6.349860	1.00
16	vasca	Memory - Foil vs. Foil - Foil	2017	6.304706	1.25
1	cuoco	Discrimination - Foil - Set 1	980	5.991226	1.35
2	valle	Discrimination - Foil - Set 1	1327	6.122871	1.00
3	crisi	Discrimination - Foil - Set 1	3171	6.501196	1.40
4	prete	Discrimination - Foil - Set 1	2656	6.424228	1.00
5	circo	Discrimination - Foil - Set 1	1405	6.147676	1.25
6	pacco	Discrimination - Foil - Set 1	1754	6.244030	1.00
7	trono	Discrimination - Foil - Set 1	1122	6.049993	1.00
8	penna	Discrimination - Foil - Set 1	2134	6.329194	1.00
9	targa	Discrimination - Foil - Set 1	1574	6.197005	1.00
10	frase	Discrimination - Foil - Set 1	2118	6.325926	1.10
11	salsa	Discrimination - Foil - Set 1	1702	6.230960	1.15
12	ritmo	Discrimination - Foil - Set 1	1622	6.210051	1.35
13	becco	Discrimination - Foil - Set 1	2442	6.387746	1.30
14	tazza	Discrimination - Foil - Set 1	2843	6.453777	1.00
15	ruota	Discrimination - Foil - Set 1	1591	6.201670	1.50
16	tomba	Discrimination - Foil - Set 1	2655	6.424065	1.30
1	sposo	Discrimination - Foil - Set 2	1734	6.239049	1.35
2	ansia	Discrimination - Foil - Set 2	1275	6.105510	1.50
3	furto	Discrimination - Foil - Set 2	2634	6.420616	1.10
4	delia	Discrimination - Foil - Set 2	1594	6.202488	1.00
5	bugia	Discrimination - Foil - Set 2	2553	6.407051	1.50
6	copia	Discrimination - Foil - Set 2	3045	6.483587	1.00
7	socio	Discrimination - Foil - Set 2	2715	6.433770	1.00
8	sparo	Discrimination - Foil - Set 2	2688	6.429429	1.00
9	ladro	Discrimination - Foil - Set 2	3079	6.488410	1.20
10	monte	Discrimination - Foil - Set 2	1599	6.203848	1.00
11	fango	Discrimination - Foil - Set 2	1322	6.121231	1.00
12	morso	Discrimination - Foil - Set 2	1987	6.298198	1.00
13	barba	Discrimination - Foil - Set 2	1832	6.262925	1.05
14	villa	Discrimination - Foil - Set 2	1433	6.156246	1.00
15	mensa	Discrimination - Foil - Set 2	1076	6.031812	1.20
16	falla	Discrimination - Foil - Set 2	2255	6.353147	1.00

Table S6. Pseudofont stimuli

ID	Stimulus	BACS-2	Condition	$\log_{10}$ frequency
1	abqsu	𐬀𐬗𐬛𐬋𐬎	FPVS - Stimuli - Set 1	3.00
2	fgmvz	𐬆𐬌𐬙𐬎𐬨	FPVS - Stimuli - Set 1	3.95
3	gmprv	𐬌𐬙𐬊𐬉𐬎	FPVS - Stimuli - Set 1	3.05
4	bcgmr	𐬗𐬁𐬌𐬙𐬉	FPVS - Stimuli - Set 1	3.75
5	imstv	𐬁𐬙𐬋𐬥𐬎	FPVS - Stimuli - Set 1	3.00
6	chilt	𐬁𐬗𐬁𐬌𐬥	FPVS - Stimuli - Set 1	2.75
7	anosz	𐬀𐬎𐬚𐬋𐬨	FPVS - Stimuli - Set 1	2.90
8	cdhms	𐬁𐬗𐬗𐬙𐬋	FPVS - Stimuli - Set 1	3.15
9	ahitz	𐬀𐬗𐬁𐬥𐬨	FPVS - Stimuli - Set 1	2.95
10	prtvz	𐬊𐬉𐬥𐬎𐬨	FPVS - Stimuli - Set 1	3.00
11	cglos	𐬁𐬌𐬋𐬚𐬋	FPVS - Stimuli - Set 1	2.90
12	amotv	𐬀𐬙𐬚𐬥𐬎	FPVS - Stimuli - Set 1	2.95
13	bnops	𐬗𐬎𐬚𐬊𐬋	FPVS - Stimuli - Set 1	2.95
14	dilsu	𐬗𐬁𐬋𐬋𐬎	FPVS - Stimuli - Set 1	3.00
15	delnz	𐬗𐬛𐬋𐬎𐬨	FPVS - Stimuli - Set 1	3.00
16	cfgmu	𐬁𐬆𐬌𐬙𐬎	FPVS - Stimuli - Set 1	3.50
1	bslfd	𐬗𐬋𐬋𐬆𐬗	FPVS - Stimuli - Set 2	3.45
2	cnldq	𐬁𐬎𐬋𐬗𐬛	FPVS - Stimuli - Set 2	3.15
3	efrdq	𐬛𐬆𐬉𐬗𐬛	FPVS - Stimuli - Set 2	3.40
4	gvpid	𐬌𐬎𐬊𐬁𐬗	FPVS - Stimuli - Set 2	3.20
5	gziou	𐬌𐬨𐬁𐬚𐬎	FPVS - Stimuli - Set 2	3.00
6	ndrom	𐬎𐬗𐬉𐬚𐬙	FPVS - Stimuli - Set 2	3.95
7	oldsh	𐬚𐬋𐬗𐬋𐬗	FPVS - Stimuli - Set 2	2.95
8	qginh	𐬛𐬌𐬁𐬎𐬗	FPVS - Stimuli - Set 2	2.95
9	qrabc	𐬛𐬉𐬀𐬗𐬁	FPVS - Stimuli - Set 2	3.00
10	rbfgd	𐬉𐬗𐬆𐬌𐬗	FPVS - Stimuli - Set 2	3.65
11	tmqiz	𐬥𐬙𐬛𐬁𐬨	FPVS - Stimuli - Set 2	3.60
12	trbep	𐬥𐬉𐬗𐬛𐬊	FPVS - Stimuli - Set 2	2.95
13	vaifm	𐬎𐬀𐬁𐬆𐬙	FPVS - Stimuli - Set 2	3.00
14	venzm	𐬎𐬛𐬎𐬨𐬙	FPVS - Stimuli - Set 2	2.90
15	zmqlb	𐬨𐬙𐬛𐬋𐬗	FPVS - Stimuli - Set 2	3.90
16	zuqnb	𐬨𐬎𐬛𐬎𐬗	FPVS - Stimuli - Set 2	3.75
1	befgt	𐬗𐬛𐬆𐬌𐬥	Memory - Stim. vs. Foil - Set 1	3.00
2	adegv	𐬀𐬗𐬛𐬌𐬎	Memory - Stim. vs. Foil - Set 1	3.00
3	bcmnu	𐬗𐬁𐬙𐬎𐬨	Memory - Stim. vs. Foil - Set 1	3.45
4	bdert	𐬗𐬗𐬛𐬉𐬥	Memory - Stim. vs. Foil - Set 1	2.90
5	cfisv	𐬁𐬆𐬁𐬋𐬎	Memory - Stim. vs. Foil - Set 1	3.00
6	clmoz	𐬁𐬋𐬙𐬚𐬨	Memory - Stim. vs. Foil - Set 1	3.00
7	deqtz	𐬗𐬛𐬛𐬥𐬨	Memory - Stim. vs. Foil - Set 1	3.00
8	epstz	𐬛𐬊𐬋𐬥𐬨	Memory - Stim. vs. Foil - Set 1	3.00
9	efiru	𐬛𐬆𐬁𐬉𐬎	Memory - Stim. vs. Foil - Set 1	3.00
10	egpqt	𐬛𐬌𐬊𐬛𐬥	Memory - Stim. vs. Foil - Set 1	3.70

ID	Stimulus	BACS-2	Condition	$\log_{10}$ frequency
11	fmpsz	ᖁᖘᒪᓕᔆ	Memory - Stim. vs. Foil - Set 1	3.90
12	ghpqr	ᕃᑕᒪᖀᕮ	Memory - Stim. vs. Foil - Set 1	3.70
13	ginqs	ᕃᑎᔨᖀᓕ	Memory - Stim. vs. Foil - Set 1	3.00
14	imnqz	ᑎᖘᔨᖀᔆ	Memory - Stim. vs. Foil - Set 1	3.80
15	mpqtv	ᖘᒪᖀᓂᔪ	Memory - Stim. vs. Foil - Set 1	3.60
16	nptuz	ᔨᒪᓂᕀᔆ	Memory - Stim. vs. Foil - Set 1	3.50
1	aclrz	ᖇᑕᑐᓭᔆ	Memory - Stim. vs. Foil - Set 2	3.00
2	aeipz	ᖇᓀᑎᒪᔆ	Memory - Stim. vs. Foil - Set 2	3.05
3	chmnz	ᑕᑕᖘᔨᔆ	Memory - Stim. vs. Foil - Set 2	3.00
4	hipsu	ᑕᑎᒪᓕᕀ	Memory - Stim. vs. Foil - Set 2	2.95
5	cmnrv	ᑕᖘᔨᓭᔪ	Memory - Stim. vs. Foil - Set 2	3.05
6	bhpqt	ᖵᑕᒪᖀᓂ	Memory - Stim. vs. Foil - Set 2	3.85
7	afioz	ᖇᖁᑎᐅᔆ	Memory - Stim. vs. Foil - Set 2	3.00
8	efoqz	ᓀᖁᐅᖀᔆ	Memory - Stim. vs. Foil - Set 2	3.65
9	abefg	ᖇᖵᓀᖁᕃ	Memory - Stim. vs. Foil - Set 2	3.00
10	bersz	ᖵᓀᓭᓕᔆ	Memory - Stim. vs. Foil - Set 2	2.95
11	efimt	ᓀᖁᑎᖘᓂ	Memory - Stim. vs. Foil - Set 2	3.00
12	ghmuz	ᕃᑕᖘᔨᔆ	Memory - Stim. vs. Foil - Set 2	3.65
13	bgpsv	ᖵᕃᒪᓕᔪ	Memory - Stim. vs. Foil - Set 2	3.75
14	biquv	ᖵᑎᖀᔨᔪ	Memory - Stim. vs. Foil - Set 2	3.00
15	afhqu	ᖇᖁᑐᖀᔨ	Memory - Stim. vs. Foil - Set 2	3.85
16	mpqsz	ᖘᒪᖀᓕᔆ	Memory - Stim. vs. Foil - Set 2	3.90
1	bdfpt	ᖵᖠᖁᒪᓂ	Memory - Foil vs. Foil - Stimuli	3.50
2	bcfor	ᖵᑕᖁᐅᕮ	Memory - Foil vs. Foil - Stimuli	3.00
3	efstv	ᓀᖁᓕᓂᔪ	Memory - Foil vs. Foil - Stimuli	3.10
4	adnoq	ᖇᖠᔨᐅᖀ	Memory - Foil vs. Foil - Stimuli	3.00
5	abdlo	ᖇᖵᖠᑐᐅ	Memory - Foil vs. Foil - Stimuli	2.85
6	dhilz	ᖠᑕᑎᑐᓭ	Memory - Foil vs. Foil - Stimuli	3.00
7	achlo	ᖇᑕᑕᑐᐅ	Memory - Foil vs. Foil - Stimuli	2.85
8	abdfm	ᖇᖵᖠᖁᖘ	Memory - Foil vs. Foil - Stimuli	3.20
9	hlmsz	ᑕᑐᖘᓕᔆ	Memory - Foil vs. Foil - Stimuli	3.80
10	bhntz	ᖵᑕᔨᓂᔆ	Memory - Foil vs. Foil - Stimuli	3.00
11	fhnrz	ᖁᑕᔨᓭᔆ	Memory - Foil vs. Foil - Stimuli	3.50
12	cdehm	ᑕᖠᓀᑕᖘ	Memory - Foil vs. Foil - Stimuli	3.00
13	eglpz	ᓀᕃᑐᒪᔆ	Memory - Foil vs. Foil - Stimuli	3.60
14	bchrv	ᖵᑕᑕᓭᔪ	Memory - Foil vs. Foil - Stimuli	3.60
15	bclqv	ᖵᑕᑐᖀᔪ	Memory - Foil vs. Foil - Stimuli	3.75
16	iloqu	ᑎᑐᐅᖀᔨ	Memory - Foil vs. Foil - Stimuli	3.00
1	degoz	ᖠᓀᕃᐅᔆ	Memory - Foil vs. Foil - Foil	2.95
2	ahisz	ᖇᑕᑎᓭᔆ	Memory - Foil vs. Foil - Foil	3.00
3	ilstu	ᑎᑐᓭᓂᕀ	Memory - Foil vs. Foil - Foil	2.95
4	abefs	ᖇᖵᓀᖁᓭ	Memory - Foil vs. Foil - Foil	3.00
5	goprz	ᕃᐅᒪᖀᔆ	Memory - Foil vs. Foil - Foil	3.00
6	almoz	ᖇᑐᖘᐅᔆ	Memory - Foil vs. Foil - Foil	3.00
7	adguv	ᖇᖠᕃᔨᔪ	Memory - Foil vs. Foil - Foil	3.00

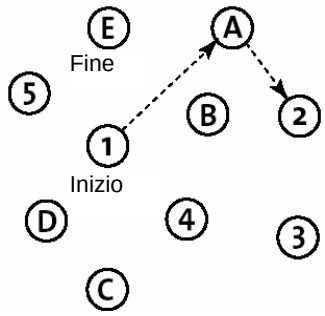
ID	Stimulus	BACS-2	Condition	$\log_{10}$ frequency
8	abmpu	ᚠᚡᚣᚳᚲ	Memory - Foil vs. Foil - Foil	3.00
9	bfhpz	ᚡᚨᚢᚳᚺ	Memory - Foil vs. Foil - Foil	3.90
10	demnz	ᚢᚽᚣᚱᚺ	Memory - Foil vs. Foil - Foil	3.00
11	alnuv	ᚠᚲᚱᚲᚴ	Memory - Foil vs. Foil - Foil	3.00
12	egmou	ᚽᚷᚣᚪᚲ	Memory - Foil vs. Foil - Foil	3.00
13	cfhrv	ᚩᚨᚢᚬᚴ	Memory - Foil vs. Foil - Foil	3.55
14	fimpu	ᚨᚿᚣᚳᚲ	Memory - Foil vs. Foil - Foil	3.00
15	ahnqz	ᚠᚢᚱᚹᚺ	Memory - Foil vs. Foil - Foil	3.85
16	fosuz	ᚨᚪᚲᚲᚺ	Memory - Foil vs. Foil - Foil	3.00
1	aclrz	ᚠᚩᚲᚬᚺ	Discrimination - Foil - Set 1	3.00
2	aeipz	ᚠᚽᚿᚷᚺ	Discrimination - Foil - Set 1	3.05
3	chmnz	ᚩᚢᚣᚱᚺ	Discrimination - Foil - Set 1	3.00
4	hipsu	ᚢᚿᚷᚲᚲ	Discrimination - Foil - Set 1	2.95
5	cmnrv	ᚩᚣᚱᚬᚴ	Discrimination - Foil - Set 1	3.05
6	bhpqt	ᚡᚢᚷᚹᚪ	Discrimination - Foil - Set 1	3.85
7	afioz	ᚠᚨᚿᚪᚺ	Discrimination - Foil - Set 1	3.00
8	efoqz	ᚽᚨᚪᚹᚺ	Discrimination - Foil - Set 1	3.65
9	abefg	ᚠᚡᚽᚨᚸ	Discrimination - Foil - Set 1	3.00
10	bersz	ᚡᚽᚬᚲᚺ	Discrimination - Foil - Set 1	2.95
11	efimt	ᚽᚨᚿᚣᚪ	Discrimination - Foil - Set 1	3.00
12	ghmuz	ᚸᚢᚣᚲᚺ	Discrimination - Foil - Set 1	3.65
13	bgpsv	ᚡᚸᚷᚲᚴ	Discrimination - Foil - Set 1	3.75
14	biquv	ᚡᚿᚹᚲᚴ	Discrimination - Foil - Set 1	3.00
15	afhqu	ᚠᚨᚢᚹᚲ	Discrimination - Foil - Set 1	3.85
16	mpqsz	ᚣᚷᚹᚲᚺ	Discrimination - Foil - Set 1	3.90
1	befgt	ᚡᚽᚨᚸᚪ	Discrimination - Foil - Set 2	3.00
2	adegv	ᚠᚹᚽᚸᚴ	Discrimination - Foil - Set 2	3.00
3	bcmnu	ᚡᚩᚣᚱᚲ	Discrimination - Foil - Set 2	3.45
4	bdert	ᚡᚹᚽᚬᚪ	Discrimination - Foil - Set 2	2.90
5	cfisv	ᚩᚨᚿᚲᚴ	Discrimination - Foil - Set 2	3.00
6	clmoz	ᚩᚸᚣᚪᚺ	Discrimination - Foil - Set 2	3.00
7	deqtz	ᚹᚽᚹᚪᚺ	Discrimination - Foil - Set 2	3.00
8	epstz	ᚽᚷᚲᚪᚺ	Discrimination - Foil - Set 2	3.00
9	efiru	ᚽᚨᚿᚬᚲ	Discrimination - Foil - Set 2	3.00
10	egpqt	ᚽᚸᚷᚹᚪ	Discrimination - Foil - Set 2	3.70
11	fmpsz	ᚨᚣᚷᚲᚺ	Discrimination - Foil - Set 2	3.90
12	ghpqr	ᚸᚢᚷᚹᚬ	Discrimination - Foil - Set 2	3.70
13	ginqs	ᚸᚿᚱᚹᚲ	Discrimination - Foil - Set 2	3.00
14	imnqz	ᚿᚣᚱᚹᚺ	Discrimination - Foil - Set 2	3.80
15	mpqtv	ᚣᚷᚹᚪᚴ	Discrimination - Foil - Set 2	3.60
16	nptuz	ᚱᚷᚪᚲᚺ	Discrimination - Foil - Set 2	3.50

Table S7. Montreal Cognitive Assessment (MoCA) Test

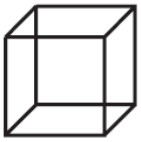
MONTREAL COGNITIVE ASSESSMENT (MOCA)
- ITALIA -

NOME: _____
Scolarità: _____ Data di nascita: _____
Sesso: _____ DATA: _____

VISUOSPAZIALE / ESECUTIVO



Copi Il cubo



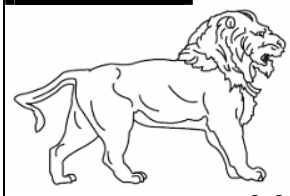
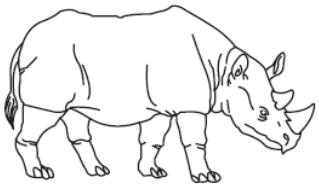
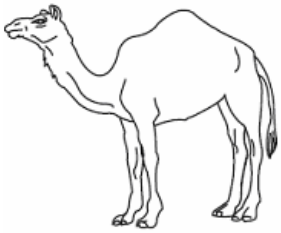
[ ] [ ]

Disegni un orologio (undici e dieci) (3 punti)

[ ] Contorno [ ] Numeri [ ] Lancette

PUNTI __/ 5

DENOMINAZIONE

[ ] [ ] [ ] __/ 3

MEMORIA Leggere la lista di parole: il soggetto deve ripeterle. Fare le prime 2 prove di seguito e **il " Richiamo " dopo 5 min.**

	Faccia	Velluto	Chiesa	Margherita	Rosso
1°prova					
2°prova					

0 punti

ATTENZIONE Leggere la serie di cifre (una cifra / sec.)
Il soggetto deve ripeterle **[ ] 2 1 8 5 4**
Il soggetto deve ripeterle in ordine inverso **[ ] 7 4 2** __/ 2

Leggere la serie di lettere. Il soggetto deve dare un colpetto con la mano sul tavolo ad ogni lettera "A". **0 punti se ≥ 2 errori**
[ ] F B A C M N A A G H L B A F A H D E A A A G A M O F A A B __/ 1

Sottrazione di 7 partendo da 100 per 5 volte **[ ] 93 [ ] 86 [ ] 79 [ ] 72 [ ] 65**
4 o 5 sottrazioni corrette: **3 pt,** 2 o 3 corrette: **2 pt,** 1 corretta: **1 pt,** 0 corretta: **0 pt** __/ 3

LINGUAGGIO Ripeta:
So solo che oggi dobbiamo aiutare Giovanni. [ ]
Il gatto si nascondeva sempre sotto il divano quando c'erano cani nella stanza. [ ] __/ 2

Fluenza / In 1 minuto, nomini il maggior numero possibile di parole che iniziano con la lettera "**F**". [ ] __ **(N ≥ 11 parole)** __/ 1

ASTRAZIONE Similitudini tra per es. banana / arancio = frutti; [ ] treno / bicicletta [ ] orologio / righello __/ 2

RICHIAMO DIFFERITO Deve ricordarsi le parole **SENZA AIUTO**

	Faccia	Velluto	Chiesa	Margherita	Rosso
	[ ]	[ ]	[ ]	[ ]	[ ]

Punti solo per ripetizione **SENZA AIUTO** __/ 5

Opzionale — AIUTO: Categoria Seman. / Scelta multipla

ORIENTAMENTO [ ] Data [ ] Mese [ ] Anno [ ] Giorno [ ] Luogo [ ] Città __/ 6

© Z. Nasreddine. Traduzione a cura di A. Pirani, C. Tulipani, M. Neri.
Versione 26 Luglio 2006
www.mocatest.org

Normale: ≥ 26 / 30

TOTALE __/30
Aggiungere 1 punto se ≤ 12 anni di istruzione

Table S8. Visual Statistical Learning (VSL) Task Stimuli

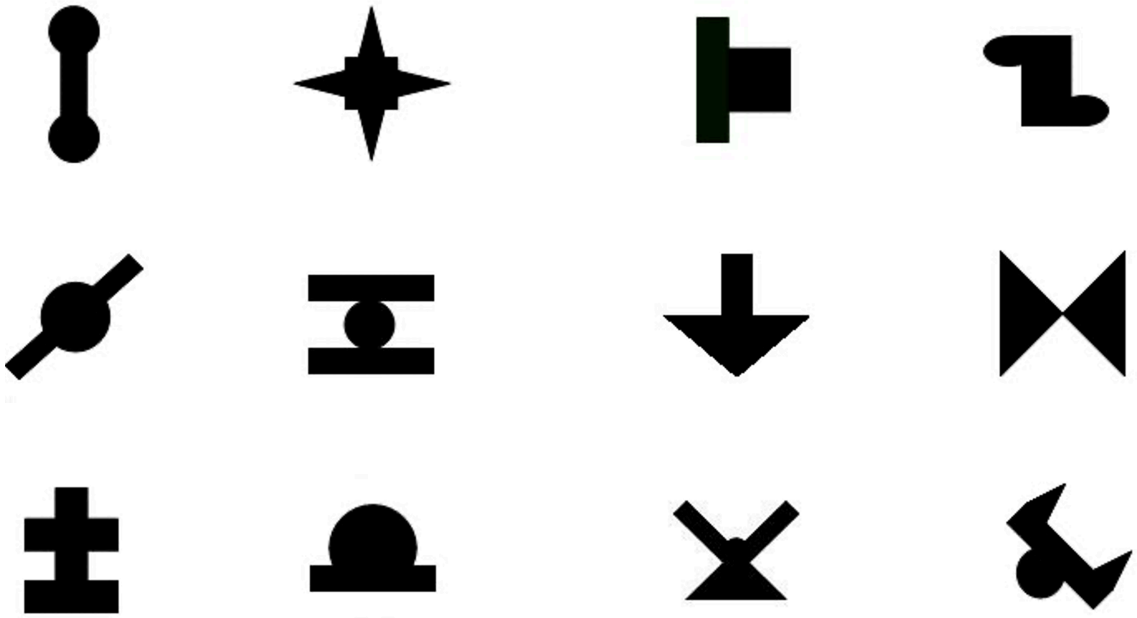


Table S9. Rating Practice Task Stimuli








Table S10. TOWRE-2 Word Practice Task Stimuli

on

my

bee

old

warm

bone

most

spell

Table S11. TOWRE-2 Nonword Practice Task Stimuli

ba

um

fos

gan

rup

nasp

luddy

dord

Table S12. TOWRE-2 Word Stimuli

is	work	inside	absentee
up	jump	plane	advertise
cat	part	pretty	pleasant
red	fast	famous	property
me	fine	children	distress
to	milk	without	information
no	back	finally	recession
we	lost	strange	understand
he	find	budget	emphasis
the	paper	repress	confident
and	open	contain	intuition
yes	kind	justice	boisterous
of	able	morning	plausible
him	shoes	resolve	courageous
as	money	describe	alienate
book	great	garment	extinguish
was	father	business	prairie
help	river	qualify	limousine
then	space	potent	valentine
time	short	collapse	detective
wood	left	elements	recently
let	people	pioneer	instruction
men	almost	remember	transient
baby	waves	dangerous	phenomenon
new	child	uniform	calculated
stop	strong	necessary	alternative
	crowd	problems	collective
	better		

Table S13. TOWRE-2 Nonword Stimuli

ip	plin	depate
ga	frip	glant
ko	poth	sploosh
ta	vasp	dreker
om	meest	ritlun
ig	shlee	hedfert
ni	guddy	bremick
pim	skree	nifpate
wum	felly	brinbert
lat	clirt	clabom
baf	sline	drepnort
din	dreff	shrattec
nup	prain	plofent
fet	bloot	smucrit
bave	trisk	pelnador
pate	kelm	fornalask
herm	strone	fermabalt
dess	lunaf	crenidmoke
chur	cratty	emulbatate
tive	trober	strotalanted
barp		prilingdorfent
stip		chundendilt