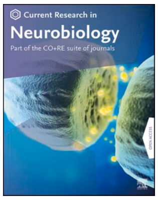
Current Research in Neurobiology 4 (2023) 100083

Contents lists available at ScienceDirect

Current Research in Neurobiology

journal homepage: www.sciencedirect.com/journal/current-research-in-neurobiology



Are musical activities associated with enhanced speech perception in noise in adults? A systematic review and meta-analysis

Elisabeth Maillard[a,b,c], Marilyne Joyal[a,b], Micah M. Murray[c,d,e], Pascale Tremblay[a,b,*]

[a] *CERVO Brain Research Center, Quebec City, G1J 2G3, Canada*
[b] *Université Laval, Faculté de Médecine, Département de Réadaptation, Quebec City, G1V 0A6, Canada*
[c] *The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland*
[d] *The Sense Innovation and Research Center, Lausanne, Sion, Switzerland*
[e] *Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA*

ARTICLE INFO

Keywords:
Speech processing
Speech-in-noise
Auditory masking
Cocktail party listening
Musician
Experience-dependent plasticity
Learning transfer

ABSTRACT

The ability to process speech in noise (SPiN) declines with age, with a detrimental impact on life quality. Music-making activities such as singing and playing a musical instrument have raised interest as potential prevention strategies for SPiN perception decline because of their positive impact on several brain system, especially the auditory system, which is critical for SPiN. However, the literature on the effect of musicianship on SPiN performance has yielded mixed results. By critically assessing the existing literature with a systematic review and a meta-analysis, we aim to provide a comprehensive portrait of the relationship between music-making activities and SPiN in different experimental conditions. 38/49 articles, most focusing on young adults, were included in the quantitative analysis. The results show a positive relationship between music-making activities and SPiN, with the strongest effects found in the most challenging listening conditions, and little to no effect in less challenging situations. This pattern of results supports the notion of a relative advantage for musicians on SPiN performance and clarify the scope of this effect. However, further studies, especially with older adults, using adequate randomization methods, are needed to extend the present conclusions and assess the potential for musical activities to be used to mitigate SPiN decline in seniors.

1. Introduction

The capacity to comprehend speech in noise (SPiN) is an essential tool for everyday life. From busy streets to supermarkets and restaurants, we are constantly faced with the challenge of untangling speech from competing sounds. SPiN capacity develops slowly, reaching adult-like performances in adolescence (e.g., for review see Leibold, 2017), and beginning to decline around the age of 50 (Moore et al., 2014) to become a common complaint among older adults (Working Group on Speech & Aging, 1988). Difficulties comprehending speech in noisy environments can lead to reduced social participation (Heine and Browning, 2002). Social disconnectedness and perceived loneliness have in turn been linked to depressive symptoms and an increased risk of dementia in elderly populations (Taylor et al., 2018; Sundström et al., 2020; Santini et al., 2020). Finding strategies to mitigate the decline of SPiN perception is therefore crucial to maintain a satisfying quality of life and wellbeing in older adults. However, the exact etiology of SPiN difficulties remains elusive. Factors such as hearing impairment, cognitive decline and functional and structural brain aging appear to contribute with varying degrees to these difficulties (Souza and Turner, 1994; Tun, 1998; Bilodeau et al., 2015; Presacco et al., 2016; Gordon-Salant and Cole, 2016; Vermeire et al., 2016; Humes, 2021).

The practice of a musical instrument is a known promoter of brain plasticity, affecting cortical as well as subcortical structures (Herholz and Zatorre, 2012; Kraus et al., 2009; Wan & Schlaug, 2010), due to its impact on auditory, cognitive, multisensory and motor processes as well as their integration (Zatorre et al., 2007). A wealth of research has examined the behavioural and neural impacts of music making activities (instrument playing, singing). Musicians exhibit better performance in auditory temporal processing (Kumar et al., 2014; Donai and Jennings,

---

Abbreviations: SPiN, speech perception in noise; M, musicians; NM, non-musicians.

[*] Corresponding author. Département de réadaptation, Université Laval, 1050 avenue de la Médecine, Quebec City, G1V 0A6, Canada.
E-mail address: pascale.tremblay@fmed.ulaval.ca (P. Tremblay).

https://doi.org/10.1016/j.crneur.2023.100083
Received 19 October 2022; Received in revised form 19 February 2023; Accepted 7 March 2023
Available online 24 March 2023
2665-945X/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

2016; Grassi et al., 2017), enhanced spectral processing with better frequency and pitch discrimination for both musical and speech stimuli (Schon et al., 2004; Micheyl et al., 2006; Deguchi et al., 2012; Kuhnis et al., 2013) and enhanced concurrent sound segregation (Zendel & Alain, 2009, 2013). Musicians also exhibit better performance on verbal working memory, especially in the auditory modality (Hanna-Pladdy and Gajewski, 2012; Zuk et al., 2014; Mansens et al., 2018; D'Souza et al., 2018; Groussard et al., 2020). Based on these findings, it has been hypothesized that music-making activities could induce cross-domain plasticity, with better SPiN performance in those practising musical activities compared to those not engaged in such activities (see the OPERA hypothesis (Patel, 2011, 2012, 2014) and Kraus & White-Schwoch for a recent review (2017)). Though the specific mechanism underlying beneficial effects associated with musical activities is not clear, cognitive reserve, either passive (structural) or active (more efficient processing) (see e.g. Stern, 2002 for a review of reserve theories) provides a general explanatory mechanism. While much of the research and debate has focused on the impact of music-making activities during development (Schellenberg, 2005, 2011, 2015; Schellenberg and Hallam, 2005; Schellenberg and Peretz, 2008; Swaminathan and Schellenberg, 2020), research suggests that the adult brain remains plastic in older ages (Anderson et al., 2013; Burke and Barnes, 2006); therefore, music-making activities could potentially have positive impacts on older adults too, including on their declining SPiN capacities.

The study of Parbery-Clark et al. (2009)—which showed that young adult musicians performed better than age-matched non-musicians on the HINT (Nilsson et al., 1994), a clinical SPiN task—opened the path for numerous studies that have explored the potential benefits of musicianship on SPiN perception under various experimental conditions. A recent review of 29 published articles (Coffey et al., 2017) on SPiN performance in musicians, while leaning towards a musician's advantage on SPiN perception, underlined the heterogeneity of the results. Multiple factors could explain this variability, including the characteristics of the participants, such as their age, their cognitive status and education levels. Characteristics related to musical training could also influence SPiN performance: the age of onset, the number of years and the intensity of musical practice could have an influence on SPiN skills, with more experience leading to better SPiN performance. In addition to participants' personal and music-related characteristics, heterogeneity could also be related to differences in SPiN tests, such as the type of task (e.g., repetition, discrimination, recognition), the response format (closed or open set), the nature of the speech targets (syllables, words, sentences), the use of spatial separation between the speech targets and the masking noise (referred to as "masker" in this article), the noise level (either positive signal-to-noise ratios (SNRs) or more challenging negative SNRs) and, importantly, the type of masking noise.

The masking properties of noises can be separated into two broad categories: energetic (noise) and informational (Freyman et al., 1999). Energetic masking competes with the target speech on a spectro-temporal level, reducing its audibility at the auditory periphery (Rennies et al., 2019). Typical energetic maskers (simply referred to as "noise maskers" in this article) used in SPiN tasks are speech-shaped, pink or white noises. In contrast, informational masking interferes with the target speech at and beyond the auditory periphery, due to perceptual similarities between the target and the masker. Informational masking modulates speech perception by affecting the segregation of the speech from the masker (object formation) and/or by drawing attention away from the target to the masker (object selection) (for a review, see Kidd et al., 2008; Shinn-Cunningham, 2013). Competing speech contains both energetic and informational masking components. The energetic masking is produced by the overlap of spectro-temporal energy between the target and the masker (Rennies et al., 2019). The informational masking of speech maskers derives from similarities between the target and masker talker's voice (e.g., same-sex speakers in both the target and the masker (Brungart, 2001)), as well as by the linguistic properties of the masker, such as the phonetic and semantic content (Hoen et al., 2007; Brouwer et al., 2012). Varying the number of talkers in a speech masker modulates both its informational and energetic masking properties. The informational power peaks with two-three talkers and then gradually decreases with the addition of talkers (assuming a constant SNR), whereas the energetic power increases proportionally with the number of talkers (Tun and Wingfield, 1999; Hoen et al., 2007; Rosen et al., 2013; Helfer and Freyman, 2014). It has been shown that task difficulty varies with the number of talkers (assuming a constant SNR): a one-talker masker represents the least challenging condition, whereas a 2–3 talker masker represents the most challenging condition (Freyman et al., 2004; Rosen et al., 2013; Helfer and Freyman, 2014). Studies on non-musicians have shown that both noise and speech maskers produce significant masking when presented collocated to the target speech (Taitelbaum-Swead and Fostick, 2016; Goossens et al., 2017). Spatially separating the target from the masker can alleviate the masking power of both noise and speech maskers in general, but the effect of spatial separation is particularly salient in speech maskers with high informational content (Freyman et al., 1999, 2001, 2004; Arbogast et al., 2005; Yost, 2017).

The effect of age on SPiN perception has been studied in non-musicians and has been shown to negatively impact performance under various experimental conditions. The negative impact of age is systematically shown in speech maskers, whereas SPiN performance in noise maskers appears to be less affected by age (Tun and Wingfield, 1999; Rajan and Cainer, 2008; Taitelbaum-Swead and Fostick, 2016; Goossens et al., 2017; Buss et al., 2019). Compared to younger adults, older adults require a more positive SNR than younger adults to reach the same performance (Desjardins and Doherty, 2013; Buss et al., 2019) and at the same SNR levels, older adults have worse performance than younger adults (Helfer and Freyman, 2014; Taitelbaum-Swead and Fostick, 2016).

The main objective of the present study was to provide a qualitative and quantitative analysis of the literature on SPiN performance in musicians compared to non-musicians through a systematic review of the literature and a meta-analysis. We aimed to evaluate the strength of the evidence for a musician advantage in SPiN performance as a function of different masker conditions, namely speech in noise maskers, speech in speech maskers, speech in spatially separated maskers and SPiN at different SNR levels (<0 dB; 0 dB; >0 dB). Assuming that the cognitive training that is inherent to music training leads to improved capacity to ignore auditory distractors, and that the sensory component of music training leads to improved auditory scene analysis capacity including the capacity to tease apart signal from noise, then one would expect to find a musician advantage with both noise and speech maskers. Further, we expected that a musician advantage would vary according to the difficulty of the SPiN task, with an advantage emerging in the most difficult conditions, given that performance in non-musician is good when noise level is low (i.e., ceiling effect). We expected a musician advantage to be stronger in collocated conditions compared to spatially separated conditions, especially with a two-talker masker (Freyman et al., 2004; Rosen et al., 2013; Helfer and Freyman, 2014) and under negative SNR levels, again because of a potential ceiling effect in easier conditions. The second objective was to evaluate whether the impact of musicianship changes as a function of the participants' age. We expected to find a musician advantage at all ages, but we expected the impact of musicianship to be more salient in older musicians especially in speech maskers, with the hypothesis of a potential cognitive reserve (Stern, 2002, 2009; Stern et al., 2020) mitigating SPiN perception decline.

2. Methods

2.1. Search strategy

A systematic literature review was conducted based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyzes (PRISMA) statement (Liberati et al., 2009). The electronic literature search was performed using PubMed and PsycNet on May 9' 2020, to identify studies comparing the speech-in-noise performance of musicians compared to non-musicians. The following keywords in these specific Boolean combinations were used: (music*[Title] OR sing [Title] OR singing [Title] OR singer [Title] OR instrument*[Title]) AND (noise OR intelligibility OR comprehensibility OR clear) AND (speech OR word OR signal OR sentence OR verbal) AND (perception OR comprehension OR intelligibility OR discrimination OR recognition OR understanding OR listening). The literature search was updated on the 15th of December 2021, filtering for the years 2020 and 2021.

2.1.1. Study selection

2.1.1.1. Selection process. A preselection on the title and the abstract of all retrieved articles was performed independently by two team members. The preselected articles were then independently assessed thoroughly on the full text by the same two reviewers. After full text assessment, the bibliographies of the selected articles were screened for additional articles, which were then independently assessed via the same preselection and selection processes by the two team members.

2.1.1.2. Selection criteria. To be included, studies had to be quantitative cross-sectional or longitudinal group studies published in French or English. There were no restrictions on publication dates. Studies had to include healthy male and female adults, with both musician and non-musician participants. In cross-sectional studies, musicianship was defined as the active solo/group practice of one/more musical instrument(s) and/or of singing—both practised at any level (amateur, professional). In longitudinal training studies, "musicians" were those enrolled in a musical instrument and/or singing training programs (solo and/or group practice). Finally, studies had to report the behavioural performance of musicians and non-musicians in at least one SPiN task. SPiN tasks had to present speech targets of any linguistic level (syllables, words, sentences) with a noise or a speech masker.

Single case, case series and multiple case studies were excluded. Studies focusing on participants affected by pathologies of any kind, including clinically diagnosed hearing impairments and/or wearing hearing aids were also excluded.

Specifically for the meta-analysis, we included only SPiN conditions with a binaural presentation of the stimuli (i.e., targets and maskers), either collocated or spatially separated. Furthermore, only SPiN conditions presenting natural speech targets were included, meaning that conditions with compressed, cochlear implant (CI) simulated, vocoded, monotonized, or whispered speech targets were excluded. With regards to noise maskers, only continuous energetic maskers were included, meaning that SPiN conditions with gated noise maskers were excluded.

2.1.2. Study designs and risk of bias assessment

The quality of each study was assessed with the QualSyst Tool (Kmet et al., 2004), which consists of 14 questions and a scoring system. When a study reported other (non-SPiN) tasks, those were not assessed. As shown in supplementary material 1a, the QualSyst Tool being a general tool, subcriteria were added to several questions to adapt the assessment to our specific question.

According to the QualSyst Tool scoring system, each question was answered by "No" (0 point), "Partial" (1 point), "Yes" (2 points) or "N/A" and a summative score was calculated (/14), which was then converted to a ratio score through a division of the sum by the maximum number of points. Based on previous systematic reviews (Lee et al., 2008; Maharaj and Harding, 2016), the summary score of each study was classified as either strong (>0.8), good (0.71-0.79), adequate (0.5-0.7) or limited (<0.5). Two of the authors (EMa, PT) independently assessed the quality of each study and their agreement was measured with an intraclass correlation coefficient (ICC) calculated with a two-way mixed model (Koo and Li, 2016). Discrepancies between authors were solved by consensus.

2.1.3. Data extraction

Data extraction was done independently by one author (EMa) and a research intern (JP) and verified by a second author (MJ). Discrepancies were solved by consensus. The extracted data are presented in supplementary materials 2 and 3. The cross-sectional and longitudinal studies are presented in separate tables. The following information was extracted for each study.

- group characteristics including sample size, mean age, sex distribution, education level and language (supplementary material 2, Tables 2a and 2b);
- musicianship, including type of musical activity [vocals, instruments], level of practice [amateur, professional], mean age of practice onset, mean number of years of practice, mean hours of weekly practice (supplementary material 2, Tables 2a and 2b);
- SPiN tasks, including type, characteristics of the target and the masker, target to masker localization, presentation mode.
- dependent variables, including accuracy (% correct), speech reception threshold (SRT), signal-to-noise ratio (SNR) loss score, reaction time (RT) and variables based on signal detection theory (sensitivity, criteria) (supplementary material 3, Table 3);
- group differences in speech performance and their direction (supplementary material 3, Table 3).

Finally, for the meta-analysis, we also extracted the mean and standard deviations for the performance of each group (musicians, non-musicians) in the selected SPiN tasks. For longitudinal studies, we extracted means and standard deviations for the post-training performance of each group.

2.1.4. Statistical analysis

The SPiN tasks of 38 studies were included in the meta-analysis. To test our hypotheses, separate meta-analyses were run on the following conditions.

a) speech in noise masker (19 studies);
b) speech in speech masker—one talker (7 studies);
c) speech in speech masker—two talkers (5 studies);
d) speech in speech masker—four talkers (16 studies);
e) speech in a spatially separated masker (6 studies);
f) speech in noise—SNR <0 dB (7 studies);
g) speech in noise—SNR = 0 dB (7 studies);
h) speech in noise—SNR> 0 dB (5 studies);

Task conditions with collocated stimuli were used in all meta-analyses except for (e). Due to the small number of studies, separate meta-analyses for speech and noise maskers for the conditions e to h could not be run (it would have led to fewer than 5 studies per analysis); hence both noise and speech maskers were included in these analyses.

In several studies, there were multiple (non-independent) outcomes, which were dealt with in the following way. First, for studies reporting SPiN performances in conditions eligible for different meta-analyses, for example one outcome with a noise masker and another one with a two-talker speech masker, the issue was automatically resolved by separating them in the different meta-analyses (Borenstein et al., 2009). Second, when studies presented various outcomes eligible for the same meta-analysis, the most represented experimental condition in the meta-analysis was always preferred. For example, when a study presented outcomes with two different noise maskers such as a white noise and a speech-shaped noise, the speech-shaped noise was always used in meta-analysis a); or when a study presented both the WIN and the QuickSIN (which only happened in 3 studies) —two tasks with a four-talker masker—the QuickSIN was always used in meta-analysis given that this was the most common test of SPiN, d). Third, when a

Table 1a
SPiN tasks and results—cross-sectional studies.

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
Studies on young adults													
Anaya et al. (2016)	M VS. NM	HINT	Oral REP	S	SPO (male)	American English	H	EM (speech-shaped noise)	COL	% correct	**SNR**: fixed at −3 dB	None	HINT & PRESTO composite score
	M VS. NM	PRESTO	Oral REP	S	SPO (different talkers)	–	H	6-TM	COL	% correct	**SNR**: fixed at 0 dB		
Baskent and Gaudrain (2016)	M VS. NM	Other	Oral REP	S	SPO (male)	Dutch	H	1-TM (S sequences; same male as the target)	COL	% correct	**Masker**: VTL shifted down by 0, 0.75, 1.5 semitones and F0 shifted up by 0, 4, 8 semitones; **SNR**: fixed at −6 dB	Every condition of the task	None
Bidelman and Yoo (2020)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	American English	H	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	QuickSIN	None
	M VS. NM	Other	Recall of callsigns (colour-number combination)	S	SPO (female (50%) or male (50%))	–	Circular array of 16 Ls	1- to 8-TM	SS (from ± 20° to +/−180°)	Accuracy; RT; Localization error	**Masker**: 1 to 8 talkers, with each talker in a different L; **SNR**: fixed at 0 dB; **Spatial configuration**: random spatial separation of targets and maskers from ± 20° to +/−180°.	S in TM with a rising number of talkers (accuracy, RT and localization error).	None
Boebinger et al. (2015)	M VS. NM	Other	Oral REP	S	SPO (female)	British English	H	1-TMs (1) clear speech; male talker, 2) spectrally rotated speech; male talker); EMs (1) speech amplitude modulated noise, 2) steady speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure; **Masker**: clear speech, spectrally rotated speech, speech amplitude modulated noise, speech spectrum steady noise.	None	Every condition of the task
Clayton et al. (2016)	M VS. NM	Other	Word IDE	S	SPO (female)	–	H	1-TM (S; different female as the target)	COL; SS (±15° azimuth)	SRT	**SNR**: adaptive procedure; **Spatial**	S in 1-TM, SS	S in 1-TM, COL

(*continued on next page*)

Table 1a (*continued*)

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
											configuration: 0°, ±15°.		
Coffey et al. (2017)	Continuum	Modified HINT	Oral REP	S	SPO (male)	American English	H	EM (speech-shaped noise)	COL	% correct	**SNR**: 2 dB, −2 dB, −6 dB.	Study with correlation/ regression	
Coffey et al. (2019)	M VS. NM	Modified HINT	Oral REP	S	SPO (male)	American English	H	EM (Speech-shaped noise)	COL	% correct	**SNR**: 2 dB to - 7 dB with 1 dB steps.	Modified HINT	None
Deroche et al. (2017)	M VS. NM	Other	Written REP	S	SPO (male)	–	H	2-TM (S; same male as the target); EM (Speech-modulated buzzes)	COL (diotic presentation); SS (dichotic presentation)	SRT	**SNR**: adaptive procedure; **F0**: difference between target and masker of 0, -2, -8 semitones; **Masker**: EM, 2-TM; **Spatial configuration**: diotic, dichotic.	None	Every condition of the task
		Other	Written REP	S	SPO (male)		H	2-TM (S; same male as the target); EM (Speech-modulated buzzes)	COL	SRT	**SNR**: adaptive procedure; **F0**: difference between target and masker of 0, -2, -8 semitones; **Masker**: EM, 2-TM; **Priming**: with priming, without priming	None	Every condition of the task
		Other	Written REP	S	SPO (male)		H	2-TM (S; same male as the target); EM (Speech-modulated buzzes)	COL	SRT	**SNR**: adaptive procedure; **F0**: difference between target and masker of −2, 0, +2, ±2 semitones; **Masker**: *masker type:* EM, 2-TM *masker roving:* F0 fixed at 125 Hz, variable over ten logarithmic steps between 100 and 150 Hz)	S in 2-TM with fixed and variable F0; S in EM variable F0.	S in EM with fixed F0.
		Other	Written REP	S	SPO (male)		H	2-TM (S; same male as the target); EM (Speech-modulated buzzes)	COL	SRT	**SNR**: adaptive procedure; **F0**: difference between target and masker of −8, 0, +8, ±8 semitones; **Masker**: *masker*	Every condition of the task	None

(*continued on next page*)

Table 1a (*continued*)

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
											type: EM, 2-TM; *masker roving*: F0 fixed at 125 Hz, variable over ten logarithmic steps between 100 and 150 Hz		
Du and Zatorre (2017)	M VS. NM	Other	IDE	SYL	SPO (female)	North American English	IE	EM (white noise)	COL	% correct; RT	**SNR**: 12 dB, −8 dB, −4 dB, 0 dB, 8 dB	SYL in EM (% correct)	SYL in EM (RT)
Escobar et al. (2020)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	IE	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	None	QuickSIN
	M VS. NM	HINT	Oral REP	S	SPO (male)	British English	L	EM (speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure	None	HINT
	M VS. NM	SPIN-R	Oral REP	S	SPO (male)	–	IE	12-TM	COL	% correct for final word	**Target**: high-predictability final word, low-predictability final words; **SNR**: fixed at +2 dB	None	Every condition of the task
Fuller et al. (2014)	M VS. NM	Other	Oral REP	W	SPO (female)	Dutch	L	EM (Speech-shaped noise)	COL	% correct (phonemes)	**SNR**: 0 dB, 5 dB, 10 dB; **Target**: natural speech, CI—simulated speech; No-noise condition	Every condition of the task	None
	M VS. NM	Other	Oral REP	S	SPO (female)	Dutch	L	EM (1) Steady speech-shaped noise; 2) fluctuating speech-shaped noise); 6-TM	COL	SRT	**SNR**: adaptive procedure; **Target**: natural speech, CI—simulated speech; **Masker**: EM1, EM2, 6-TM	None	Every condition of the task
Kaplan et al. (2021)	M VS. NM	Other	Oral REP	S	SPO (one female)	Dutch	L	2-TM (S; different female as the targets)	COL	% correct	**SNR**: 3, −5, −7, −9 dB	Every condition of the task	None
Jain and Nataraja (2019)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	Kannada	H	4-TM (one male and three females)	COL	SRT	**SNR**: +8 dB to −10 dB with 3 dB steps	None	QuickSIN
Madsen et al. (2017)	M VS. NM	Other	Written REP	S	SPO (male)	British English	H	2-TM (S); EM (Speech-shaped noise)	COL	% correct	**Target**: natural speech, monotonized speech; **Masker**: EM,2-TM	None	Every condition of the task

(*continued on next page*)

Table 1a (*continued*)

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
											(natural speech with F0 equal or 1/2/4/8 semitones lower than the target, monotonized speech); **SNR**: fixed at −3 dB		
Madsen et al. (2019)	M VS. NM	Other	Word IDE	S	SPO (3 different females, one at a time)	Danish	L	1-TM (S; same 3 females as the target, one at a time)	COL; SS (±15° azimuth)	SRT	**SNR**: adaptive procedure; **Spatial configuration**: 0°, ±15°	None	Every condition of the task
	M VS. NM	Other	Oral REP	S	SPO (male)		L	2-TM (two males); EM (Gaussian noise)	COL; SS (±15° azimuth)	SRT	**SNR**: adaptive procedure; **Masker**: 2TM, EM; **Spatial configuration**: 0°, ±15°; **Reverberation**: anechoic, reverberant	None	Every condition of the task
Mandikal Vasuki et al. (2016)	M VS. NM	LiSN–S test (2 conditions)	Oral REP	S	SPO (female)	Australian English	H	1-TM (continuous discourse; same voice as the target or a different female voice)	COL	SRT	**SNR**: adaptive procedure; **Masker**: same female, different female as the target	None	Every condition of the task
Mankel and Bidelman (2018)	M VS. NM (both high- and low-PROMS)	QuickSIN	Oral REP	S	SPO (female)	North American English	–	4-TM (one male and three females)	COL	SNR loss	**SNR**: from 25 dB to 0 dB with 5 dB steps	QuickSIN	None
Morse-Fortier et al. (2017)	M VS. NM	Other	Recall of target words	W	SPO (female)	North American English	L	2-TM (S; female)	COL; COL + extra masker at +60° azimuth	SNR threshold	**SNR**: adaptive procedure; **Target**: natural speech, vocoded speech; **Masker**: natural speech, vocoded speech; **Spatial configuration**: masker and speech COL at 0°, masker and speech COL at 0° + masker at +60°	Every condition of the task	None

(*continued on next page*)

Table 1a (*continued*)

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
Parbery-Clark et al. (2009)	M VS. NM	HINT	Oral REP	S	SPO (male)	American English	L	EM (speech-shaped noise)	COL; SS (+90° azimuth—90° azimuth)	SRT	**SNR**: adaptive procedure; **Spatial configuration**: 0°, +90°, −90°	HINTfront (S in EM, COL)	HINTright (S in EM, SS at +90° azimuth); HINTleft (S in EM, SS at −90° azimuth).
	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	IE	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	QuickSIN	None
Parbery-Clark et al. (2011)	M VS. NM	HINT	Oral REP	S	SPO (male)	American English	L	EM (speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure	HINT	None
Parbery-Clark et al. (2012)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	IE	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	QuickSIN	None
Parbery-Clark et al. (2013)	M VS. NM	HINT	Oral REP	S	SPO (male)	American English	L	EM (speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure	HINT	None
Puschmann et al. (2019)	Continuum	Other	Word detection	S	SPO (male)	American English	IE	1-TM (S; male)	COL	F1 score (= harmonic mean of recall (i.e., number of detected target W divided by the total number of targets W); Precision (number of detected target W divided by the total number of button presses).	**SNR**: fixed at 0 dB; No-noise condition	Study with correlation/regression	
		HINT	Oral REP	S	SPO (male)	American English	–	EM (speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure		
		Other	Oral REP	S	SPO (3 different males)	–	–	1-TM (S; 3 same males as the target)	COL	SRT	**SNR**: adaptive procedure		
Ruggles et al. (2014)	M VS. NM	Other	Written REP	S (nonsense)	SPO (female), WHI	–	H	EMs (1) continuous speech-	COL	% correct; masking release	**SNR**: 6 dB, −3 dB, 0 dB); **Target**: voiced	None	Every condition of the task

(*continued on next page*)

Table 1a (*continued*)

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
					(same female), modified WHI speech (same female)			shaped noise, 2) square-wave gated speech-shaped noise)		(between gated and continuous noise)	speech, WHI speech, modified WHI speech; **Masker**: continuous noise, gated noise		
	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	H	4-TM (one male and three females)	COL	SNR loss	**SNR**: from 25 dB to 0 dB with 5 dB steps	None	QuickSIN
	M VS. NM	HINT	Oral REP	S	SPO (male)	American English	H	EM (speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure	None	HINT
Skoe et al. (2019)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	IE	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	QuickSIN (with noise exposure as a covariate)	QuickSIN (without noise exposure as a covariate)
Slater and Kraus (2016)	Percussionists and Vocalists VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	IE	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	QuickSIN	None
	Percussionists and Vocalists VS. NM	WIN	Oral REP	W	SPO (female)	North American English	IE	4-TM	COL	SNR threshold	**SNR**: 24 dB–0 dB with 4 dB steps	None	WIN
Slater et al. (2018)	Percussionists VS. NM	QuickSIN	Oral REP	S	SPO (female)	American English	L	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	QuickSIN	None
Soncini and Costa (2006)	M VS. NM	Other	Oral REP	S	SPO	Brazilian Portuguese	H	EM (speech-shaped noise)	COL in right ear only (monaural presentation); COL in left ear only (monaural presentation)	SRT-quiet; SNR level	**SNR**: adaptive procedure; No-noise condition	S in EM a) left ear only, b) right ear only(SNR level).	S in quiet a) left ear only, b) right ear only(SRT-quiet).
Swaminathan et al. (2015)	M VS. NM	Other	Word IDE	S	SPO (female)	–	H	1-TM (S; female)EM (time reversed 1-TM)	COL; SS (±15° azimuth)	SRT	**SNR**: adaptive procedure. **Masker**: forward or time-reversed; **Spatial configuration:** 0°, −15°, +15°	S in 1-TM, SS; S in EM, COL.	S in 1-TM, COL; S in EM, SS.
Vanden Bosch der Nederlanden et al. (2020)	M VS. NM	Revised speech-in-noise (R-Spin)	Oral REP (last W in the carrier S)	S	SPO (male)	–	H	Multi-TM	COL in left ear only (monaural presentation)	% correct (at −1 dB SNR)	**SNR**: 1 dB–23 dB with 3 dB steps; **Target**: high predictability carrier S, low	Study with correlation/ regression	

(*continued on next page*)

Table 1a (*continued*)

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
											predictability carrier S		
Varnet et al. (2015)	M VS. NM	Other	SYL categorization (da or ga)	VCCV NW	SPO (male)	–	H	EM (white noise)	COL	% correct; signal detection theory [sensitivity (d'), RT, decision criteria (c)]; SNR threshold	**SNR**: adaptive procedure	SYL in EM (% correct, d', SNR)	SYL in EM (RT, c)
Yoo & Bidelman (2019)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	H	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	QuickSIN	None
	M VS. NM	HINT	Oral REP	S	SPO (male)	American English	H	EM (speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure	None	HINT
	M VS. NM	WIN	Oral REP	W	SPO (female)	North American English	H	4-TM	COL	SNR threshold	**SNR**: 24 dB–0 dB with 4 dB steps	None	WIN
Zendel et al. (2015)	M VS. NM	Other	Oral REP	W	SPO (male)	French (Québec)	IE	4-TM (two women and two men)	COL	% correct	**SNR**: 0 dB, 15 dB; No-noise condition	W in 4-TM at SNR = 0 dB	W in 4-TM at SNR = 15 dB; W in quiet
Zhang et al. (2019)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	H	4-TM (one male and three females)	COL in left ear only (monaural presentation), COL in the right ear only (monaural presentation), COL in both ears (binaural presentation)	SNR loss	**SNR**: from 25 dB to 0 dB with 5 dB steps; **Spatial configuration**: COL stimuli in the right ear only, the left ear only, in both ears	None	Every condition of the task
Studies on middle-aged adults													
Parbery-Clark et al. (2011)	M VS. NM	HINT	Oral REP	S	SPO (male)	American English	L	EM (speech-shaped noise)	COL	SRT	SNR: adaptive procedure	HINT	None
	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	IE	4-TM (one male and three females)	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	QuickSIN	None
	M VS. NM	WIN	Oral REP	W	SPO (female)	North American English	IE	4-TM	COL	SNR threshold	SNR: 24 dB–0 dB with 4 dB steps	WIN	None
Parbery-Clark et al. (2012)	M VS. NM	HINT	Oral REP	S	SPO (male)	American English	L	EM (speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure	HINT	None
Studies on older adults													

(*continued on next page*)

Table 1a (*continued*)

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
Fostick (2019)	M VS. NM (Card players) VS. NM (Controls)	AB word test	Oral REP	W	SPO	Hebrew	H	EMs (1) speech-shaped noise, 2) white noise)	COL	% correct	**Target**: natural speech; compressed speech; **Masker**: EM1, EM2; **SNR**: fixed at 0 dB; No noise condition	Every condition of the task	None
Mussoi (2021)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	North American English	IE	4-TM	COL	SNR loss	**SNR**: from 25 dB to 0 dB with 5 dB steps	None	QuickSIN
		HINT	Oral REP	S	SPO (male)	American English	L	EM (speech-shaped noise)	COL	SRT	**SNR**: adaptive procedure	None	HINT
		SPIN-R	Oral REP	S	SPO (male)	–	IE	12-TM	COL	% correct	**SNR**: fixed at +2 dB	None	SPIN-R
Studies on a broad age range													
Meha-Bettison et al. (2018)	M VS. NM	LiSN–S test	Oral REP	S	SPO (female)	Australian English	H	1-TM (continuous discourse; same voice as the target or a different female voice)	COL; SS (±90° azimuth)	SRT	**SNR**: adaptive procedure; **Masker**: same voice as the target, different voice; **Spatial configuration**: 0°, ±90°	Same voice for target and masker at 0° azimuth ("low-cue SRT"; no group effect, result of planned comparison)	Different voices for target and masker at ±90° azimuth ("high-cue SRT"); Same voice for target and masker at ±90°; Different voices for target and masker at 0° azimuth; Same voice for target and masker at 0° azimuth.
Perron et al. (2021)	M VS. NM	Other	Discrimination	SYL	SPO (male)	French (Québec)	H	4-TM (2 females and 2 males)	COL	Signal detection theory (Sensitivity (d'))	**SNR**: 3 dB, +3 dB	None	SYL in 4-TM (average d' between −3 dB and +3 dB)
Yeend et al. (2017)	Continuum	LiSN–S test (2 conditions)	Oral REP	S	SPO (female)	Australian English	H	1-TM (continuous discourse; female)	SS (±90° azimuth)	SRT	**SNR**: adaptive procedure	Study with correlation/ regression	
Zendel and Alain (2012)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	American English	IE	4-TM	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	Study with correlation/ regression	
Studies on young and older adults													

(*continued on next page*)

Table 1a (*continued*)

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	M > NM[b]	∅[c]
Zhang et al. (2021)	M VS. NM, Young VS. Older adults	Other	Oral REP	S	SPO (one female)	Chinese	H	2-TM (S; two different females); EM (speech spectrum noise)	COL (ITD of 0 ms); SS (ITD of 2 ms, −2 ms)	SRT	**SNR**: 12, −8, −4, 0, 4 dB); **Masker:** EM, 2-TM; **Spatial configuration:** ITD of 0 ms,ITD of 2 ms, −2 ms.	Older adults: Every condition of the task	Young adults:Every condition of the task

study presented various outcomes eligible for the same meta-analysis but with an additional experimental manipulation, such as results with the same noise masker but with variations in the stimuli's F(0) or the SNR levels, an average score of these different results was calculated and used in meta-analysis a) (Borenstein et al., 2009). Finally, when two studies eligible for the same meta-analysis had exactly the same participants in their sample, we included only one of them. Therefore, in the SNR> 0 dB meta-analysis, we included data from Zendel et al. (2019), but not from Fleming et al. (2019).

For each meta-analysis, a random-effect model was used to take into account the variability in SPiN tasks and participant characteristics (Borenstein et al., 2009). The random effect meta-analyses were conducted using the R *Metafor* package (R 4.2.1; R Development Core Team) running on Windows 11. Due to the small number of studies per analysis, Hedge's g (Borenstein et al., 2009) was used for each analysis as a power estimate (bias-corrected standardized mean differences). Hedge's g of <0.5, 0.5–0.8, and> 0.8 were considered as small, moderate, and high, respectively (Cohen, 1988). The Cochrane's Q test and I^2 statistics (Higgins et al., 2003) were used to assess the between-study heterogeneity. For the Cochrane's Q test, to indicate the presence of between-study heterogeneity, an alpha level of 0.05 was chosen for statistical significance. I^2values of 25%–50%, 50%–75%, and ≥75% were considered as low, moderate, and high between-study heterogeneity, respectively (Higgins et al., 2003). To test whether any study was overly influential on the effect size estimate, we conducted a leave-one-out cross-validation in each meta-analysis (Viechtbauer, 2010). Finally, meta-regressions (Huizenga et al., 2011) were also conducted to assess the potential moderating impact of different factors on SPiN performance. First, the mean of the musician and non-musicians' mean ages were computed and used as a continuous moderator, "mean age," in each analysis. Second, to evaluate the effect of study quality, the scores obtained on the QualSyst evaluation tool were used as a continuous moderator, "study quality," in each meta-analysis. The little number of available studies prevented the assessment of other potential moderators, such as the target's linguistic level (syllable, word, sentences) or the masker type (noise masker, speech masker) in meta-analyses c) and d).

The extracted data and analysis scripts are available on Borealis, the Canadian Dataverse Repository (DOI: https://doi.org/10.5683/SP3/TQAUAS).

3. Results

3.1. Study selection

Electronic literature searches conducted in May 2020 identified a total of 579 studies matching the search terms: 322 in Pubmed and 257 in PsycNet. Five additional records were obtained: 3 from the bibliography screening of all selected articles identified through database searching and 2 articles were obtained directly from their authors. 460 records remained after the removal of duplicates. References not meeting the inclusion criteria on the screening of titles and abstracts were discarded (N = 400 records). All disagreements were solved by either consensus (23/28) or by involving the senior author (PT) (5/28). After evaluation of the remaining 60 articles on their full text, 17 articles that did not meet the inclusion criteria for the systematic review were discarded, and 43 articles were retained. All disagreements (3/3) were resolved by involving the senior author (PT), who evaluated the articles independently. A record of both the articles and the reason for the exclusion were kept, as well as the number of disagreements. 6 new records were selected following the update conducted in December 2021 with the same search criteria and procedure. In total, the systematic review included 49 studies, including 43 cross-sectional and 6 longitudinal studies, conducted between 2009 and 2021. The selection process is illustrated in Fig. 1.

17 of the included studies reported the mean and standard deviation

Table 1b
SPiN tasks—longitudinal studies.

Article	Groups	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation	Masker	Spatial configuration[a]	DV	Experimental manipulation	Pre/post training improvement
Studies on middle-aged adults												
Hennessy et al. (2021)	M VS. NM	BKB-SIN	Oral REP	S	SPO (male)	British English	L	4-TM	COL	SRT	**SNR**: 21, 18, 15, 12, 9, 6, 3, 0, −3, −6 dB	None (No Group × Time interaction)
Studies on older adults												
Dubinsky et al. (2019)	M VS. NM	QuickSIN	Oral REP	S	SPO (female)	American English	H	4-TM	COL	SNR loss	**SNR**: 25 dB–0 dB with 5 dB steps	Improvement for the musical training group only
Fleming et al. (2019)	M VS. NM (video game players) VS. NM (No contact group)	Other	S-picture matching	S	SPO (male)	French (Québec)	IE	Multi-TM	COL	% correct	**SNR**: 5 dB, 20 dB; No noise condition	None (No Group × Time interaction)
Merten et al. (2021)	M VS. NM	Other	Oral REP	W	SPO (female)	North American English	–	1-TM (S; male talker)	COL to better ear (monaural presentation)	% correct	**SNR**: fixed at +8 dB.	Less decline over time in M vs. NM
Worschech (2021)	M VS. NM	International Matrix Test	Oral REP	S	SPO (Fr: one female; Ge: one male)	French/ German	H	EM (speech-shaped noise)	COL in left ear only (monaural presentation), COL in the right ear only (monaural presentation), COL in both ears (binaural presentation)	SRT	**SNR**: adaptive procedure; **Spatial configuration**: COL stimuli in the right ear only, the left ear only, in both ears	Greater improvement in M vs. NM for the monaural (left ear) condition only
Zendel (2019)	M VS. NM (video game players) VS. NM (No contact group)	Other	Oral REP	Words	SPO (male)	-	IE	4-TM	COL	% correct	**SNR**: 0 dB, 15 dB; No-noise condition	M improved over time in the 0 dB condition (Group X Time × Condition interaction)

Note. CI = cochlear implant; COL = collocated; DV = dependent variable; EM = noise masker; F0 = fundamental frequency; H = headphones; IDE = identification (multiple choice); IE = insert earphones; L = loudspeaker; M = musicians; NM = non-musicians; NW = non-words; REP = repetition; RT = reaction time; S = sentences; SNR = signal-to-noise ratio; SRT = speech reception threshold (SNR threshold with 50% correct answer); SS = spatially separated; SPO = spoken; SYL = syllables; TM = talker masker; VTL = vocal tract length; W = words; WHI = whispered.

[b] M > NM: conditions with an M advantage.

[c] ∅: condition without an M advantage.

[a] Spatial configuration = target-masker localization, in ° azimuth or ITD (interaural difference).

Table 2
Tasks and descriptive data used in the “speech in noise masker” meta-analysis.

Article	Tasks	Score	Musicians M (SD)	Non-Musicians M (SD)
Anaya et al. (2016)	HINT at −3 dB	% correct	50.74 (8.59)	47.22 (8.11)
Boebinger et al. (2015)	Sentences in steady speech shaped noise	SRT	−3 (1.00)	−2.7 (0.90)
Coffey et al. (2019)	modified HINT	% correct	−0.66 (0.06)	−0.58 (0.08)
Deroche et al. (2017) (exp.1)	Sentences in buzz—mean score for buzz and diotic conditions	SRT	−8.67 (0.50)	−7.96 (1.44)
Deroche et al. (2017) (exp.4)	Sentences in buzz—mean score fur buzz and fixed masker F0 conditions	SRT	−8.95 (1.57)	−7.52 (1.70)
Du and Zatorre (2017)	Syllables in white noise—mean score of SNR conditions (−8 dB, −4 dB, 0 dB, 8 dB)	% correct	76.73 (3.57)	69 (6.03)
Escobar et al., 2019	HINT	SRT	−0.57 (0.95)	−0.69 (0.93)
Fostick (2019)	AB word test: words in speech shaped noise	% correct	72 (15)	55 (15)
Fuller et al. (2014)	Sentences in steady speech shaped noise	SRT	−6.25 (1.20)	−5.77 (1.38)
Madsen et al. (2017)	Sentences in Gaussian noise	% correct	73.53 (9.59)	72.91 (6.84)
Madsen et al. (2019)	Sentences in Gaussian noise (anechoic condition)	SRT	−0.52 (1.02)	−0.58 (0.82)
Mussoi (2021)	HINT	SRT	0.13 (2.81)	−1.18 (1.64)
Parbery-Clark et al., 2011b	HINT	SRT	−3.37 (0.52)	−2.24 (0.87)
Parbery-Clark et al., 2012b	HINT	SRT	−3.16 (0.61)	−2.34 (0.63)
Puschmann et al. (2019)	HINT	SRT	−3.02 (1.06)	−2.24 (0.95)
Ruggles et al. (2014)	HINT	SRT	−2.34 (0.59)	−2.21 (0.58)
Varnet et al. (2015)	Syllables in white noise	Mean SNR threshold	−13.37 (1.19)	−11.91 (1.01)
Worschech et al., 2021	International Matrix Test	SRT	−6.61 (2.21)	−6.90 (2.21)
Yoo & Bidelman, 2019	HINT	SRT	−6.5 (2.00)	−6.13 (2.07)
Zhang et al. (2021) (older adults)	Sentences in sentences	SRT	−2.95 (0.94)	−2.07 (1.17)
Zhang et al. (2021) (young adults)	Sentences in sentences	SRT	−3.34 (0.91)	−3.24 (0.60)

Note. M = mean, SD = standard deviation, SRT = speech reception threshold, F0 = fundamental frequency.

for each group on the SPiN task(s). The authors of the remaining 32 studies were contacted. The descriptive data were obtained for 24 additional articles. A total of 38 articles were included in the quantitative analyses based on our selection criteria.

3.1.1. *Study quality*

Most studies were cross-sectional (N = 43) and only 6 reported an intervention (longitudinal). The mean summary score for study quality was 0.58 ± 0.09, corresponding to an adequate quality score (0.5-0.7). Specifically, 39 studies exhibited an adequate quality score (0.5-0.7), 7 studies exhibited a limited (<0.5) quality score, 3 studies exhibited a good quality score (0.71-0.79), and 1 study exhibited a strong (>0.8) quality score. The detailed results are presented in supplementary material 1, Table 1b. The single measure ICC between both judges was high (0.980, with a 95% confidence interval from 0.965 to 0.989, $F_{(48,\,48)} = 100.498$, $p < .001$).

Table 3
Moderator analysis for the “speech in noise masker” condition.

Moderator	ß	ß's	β's $CI_{95\%}$		SE	Q's	I^2 (%)
		p-value	LL	UL		*p-value*	
Mean age	0.01	0.84	−0.01	0.01	0.01	0.00	71.93
Study quality	−1.51	0.38	−4.89	1.87	1.72	0.00	72.27

Note. ß = regression coefficient, $CI_{95\%}$ = 95% confidence interval, LL = lower limit of the $CI_{95\%}$, UL = upper limit of the $CI_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

The most frequent issues were the following. Few studies specified the age group of interest (Qual Sys tool question 1). Most studies were not randomized, even the longitudinal (training) studies (Qual Sys tool question 2). The sample size for the groups in most studies was small, with 20 or fewer participants, and not justified using power calculation (Qual Sys tool question 9). Regarding the recruitment of participants, the selection criteria were generally detailed but the recruitment settings (e.g., place of recruitment such as a university or a music school or the general community), and especially the recruitment methods (e.g., emails, flyers) were lacking in most studies (Qual Sys tool question 3). While participants demographic information was generally reported, 8 of the studies included in the review did not provide any information about the musicians' musical practices, such as a mean number of years of practice and/or a mean age of onset. Musicians were generally matched on age with the control group, but were matched on hearing in only 14 studies, introducing potentially important confounds (Qual Sys tool question 12). The statistical analyses were in general well described and appropriate (Qual Sys tool question 10). The reporting of results was, however, less strong (Qual Sys tool question 13), due to several articles reporting the descriptive data of SPiN performances either graphically only (19 studies) or not at all (9 studies).

3.1.2. *Focus(es) of the selected studies*

In 46/49 of the articles, one of the main goals was to explore the effect of musical training on SPiN tasks. In 3/43 cross-sectional studies, performance in the SPiN tasks was used as a predictor or comparator and was not the focus of the analyses. 35/43 cross-sectional studies reported the performance of musicians and non-musicians, and 5/49 studies consisted of correlation/regression studies. Among the longitudinal studies, 5/6 were experimental (i.e., studies with pre-post designs, with or without randomization) and one was a cohort study.

3.1.3. *Participant description*

Across all studies, there were 5434 participants in total (57% female). The language of the participants was specified in 30 studies. The majority (15) of the studies involved North American English speakers. The other languages that were reported are British English (1 study), Australian English (2 studies), French from Québec (3 studies), French from Switzerland (1 study), Dutch (3 studies), Kannada (1 study), Brazilian Portuguese (1 study), Danish (1 study), Chinese (1 study) and Hebrew (1 study).

Among cross-sectional studies, 34 focused on young adults (18–40 years old), 2 on middle-aged adults (45–65 years) and 2 on older adults (≥60 years old). 4 studies employed groups with a broad age range (18–91, 30–60, 22–59, 20–87 years old) and one study recruited both younger adults and older adults. Among the longitudinal studies, 5 focused on older adults (≥60 years old) and one study on middle-aged adults (50–65 years old).

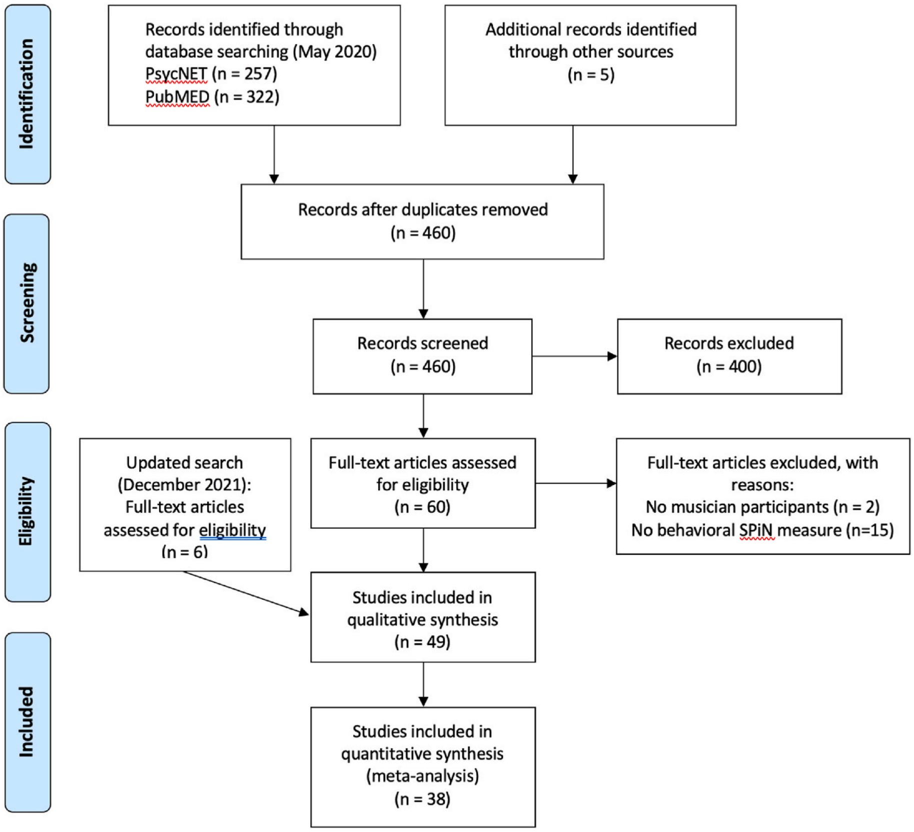


Fig. 1. PRISMA figure.

Information on the participants' education was inconsistently available in 23 studies, with musicians and non-musicians being matched on education in only 10/23 studies. Furthermore, the available information revealed that in most studies, a majority of highly educated participants (college students) were recruited.

The pure tone thresholds of participants were measured in most studies (46/49). All but 8 studies specified a pure tone threshold level below which participants hearing fell; this threshold ranged from 15 to 50 dB HL. Only 13/49 studies reported the actual hearing thresholds for each group. In 16/49 studies, it was reported that M and NM were matched on hearing; in the other studies, the groups were not matched.

The cognitive functioning of participants was verified in only 25/49 studies; and matched on at least one measure in 20/49. The scores at the Montreal Cognitive Assessment were measured in 3 studies; participants were matched in two studies. Specific executive functions (working memory, attention, speed of processing) were measured in 9 other studies, with participants being matched in 6/9 studies. In two other studies, the cognitive scores were used as a factor in linear regression analyses. Participants' IQ was measured in 13 studies; participants were matched across groups in 12 of those studies. A significant difference in IQ was found in one study, which was then controlled for in the analyses. Finally, the Cognitive Telephone Screening Instrument and the Mini Mental State Examination were measured in one study each. Participant information is detailed in Supplementary material 2.

3.1.4. *Musicianship*

Among the cross-sectional studies, two studies involved a single group of participants, with musicianship measured in terms of the number of years of musical training (continuous quantitative variable). For the other 41 cross-sectional studies, the musicianship factor was defined by specific inclusion criteria to differentiate a group of musicians (M) from a group of non-musicians (NM). Across all cross-sectional studies, there were 1149 M and 1041 NM. Among the longitudinal studies, there were 152 M and 176 NM across all experimental studies and 630 M and 2308 NM in the only cohort study. In the experimental studies, 23 NM were enrolled in a music-listening group, 82 in a musical culture group, 16 in a non-musical activity group and 55 were enrolled in a no-activity group. For the 2308 NM in the cohort study, there was unspecific information characterizing their inclusion in the NM group. The average number of participants per cross-sectional study was relatively small with 21.22 participants (SD = 11.13; range: 10–74) in the M group and 21.69 (SD = 13.22; range: 4–89) in the NM groups. The average number of participants in the longitudinal experimental studies was slightly higher, with 30.40 participants (SD = 25.85; range: 13–74) in the M group and 25.14 (SD = 26.25; range: 8–82) in the NM groups.

The inclusion criteria reported for the musician groups in the cross-sectional studies were either a maximum onset age for musical practice (26 studies, range: 7–16 years old), a minimum number of years of practice (29 studies, range: 2–10) and a minimum weekly practice (10 studies), either in terms of the number of practice sessions per week (3x/week) or in terms of the number of hours of practice per week (range: 5–15 h/week). In 17 studies, both onset age and minimum number of years of practice were used as inclusion criteria. In 3 studies, a maximum onset age and a minimal amount of training per week were used as inclusion criteria. In 2 studies, a minimum number of years of practice and a minimal amount of training per week were used as inclusion criteria. Finally, only 3 studies combined all three inclusion criteria (i.e., age of onset, number of sessions per week and session duration).

Among the cross-sectional studies that provided information about the participants' type of musical activities (32/43), 12 involved instrumentalists (any instrument) and 17 involved both instrumentalists (any instrument) and vocalists. In four studies, the M group was homogeneous: there was one study on choir singers, one study on singers and percussionists, one study on keyboard players (piano or organ), one study on percussionists and one study involving a group of pianists, one group of percussionists, one group of vocalists and one group of string players. Four studies focused on amateur musicians, 6 on professionals and 8 on both (participants enrolled in a university music program were considered as professionals). Most studies did not provide this information (30/43). The average number of musical training years was available for the musician group in 29 studies, with an average number of 15.20 ± 2.62 years of musical training for younger adults, 49.68 ± 0.45 years for middle-aged adults, 27.75 ± 17.70 for one study on older adults and 39.7 ± 12.03 years for one study with a broad age group. The age of onset of musical training was available in 28/43 studies with an average starting age of 6.26 ± 1.23 years old for young adults, 6.06 ± 0.65 years old for middle-aged adults, 7.86 ± 3.74 for one study on older adults and 7.2 ± 1.9 years old for one study with a broad age group. The average hours of weekly practice for the musician group were available in only 4 studies: 19.85 ± 5.01 h for studies on young adults, 11.47 ± 2.98 for older adults and 24.65 ± 7.12 h for one study on a broad age group.

All longitudinal experimental studies compared a musical training against another kind of training, a do-nothing control group or both. The training lasted from 10 weeks to 6 months, with a total weekly training duration of 2.5 h up to 4.25 h.

3.1.5. SPiN tasks

70 SPiN tasks (Tables 1a and 1b) were used across the 49 studies; the QuickSIN (17 studies) (Killion et al., 2004), the Hearing in Noise Test (HINT, 11 studies) (Nilsson et al., 1994), a modified (computerized) version of the HINT (2 studies) (Coffey et al., 2017), the Listening in spatialized noise sentence test (LiSN–S T, 3 studies) (Cameron and Dillon, 2007), the PRESTO (1 study) (Gilbert et al., 2013), the Words-in-noise test (WIN, 3 studies) (Wilson, 2003; Wilson et al., 2007), the Revised Speech-in-noise (R–SPIN, 3 studies) and the AB-word test (Hebrew version, n = 1) (Boothroyd, 1968; Fostick et al., 2014), the BKB-SIN, the International Matrix Test and 29 experimental tasks.

3.1.5.1. Targets: characteristics and experimental manipulations. Sentences were the most common targets in the selected studies, being used in 58/70 tasks. Words were used in 9/70 tasks, non-words in 1/70 tasks and syllables in 2/70 tasks. Targets were acoustically modified in 9/70 experimental tasks (5 studies), including variation of the fundamental frequency (F0), whispered speech, monotone speech, CI-simulated speech and vocoded speech.

3.1.5.2. Maskers: characteristics and manipulations. The most frequent masker type was speech, which was used in 49/70 tasks (43 studies): four-talker babble was the most represented speech masker (24 tasks), followed by one-talker babble (12 tasks), two-talker babble (10 tasks), six-talker babble (3 tasks), twelve-talker babble (3 tasks), eight-talker babble (1 task), multi-talker babbles not otherwise specified (1 task), and time-reversed speech (1 task).

Noise maskers were used in 28/70 tasks (24 studies). Speech shaped noises were the most frequent (21 tasks) and the other noise maskers found were white noise (1 task), buzzes (4 tasks), Gaussian noise (1 task) and time-reversed speech (1 task).

The masker's properties were modified in 12/70 experimental tasks. Of these, 9/12 tasks used both speech and non-speech maskers. In 2/12 tasks the speech masker's amount of informational content was varied by increasing the number of talkers or by playing the one-talker babble in reverse. In 2/12 tasks, the vocal properties of the masker's talker were modified by varying the fundamental frequency (F0) of the talker(s) (1 task) or by varying both the F0 and the vocal tract length (VTL) of the talker (1 task). In 1/12 task, the masking power of the non-speech masker was varied by using a continuous and a gated noise.

3.1.5.3. Stimuli presentation. The most common task presentation was via headphones (33/70 tasks). Loudspeakers were used in 16/70 tasks and insert earphones in 16/70 tasks. The presentation was not specified for 4 tasks.

3.1.5.4. SNR manipulation. A SNR manipulation was present in 59/70 tasks. In 32 tasks, the SNR was varied along predefined SNR levels and in 28 tasks, the SNR varied in an adaptive procedure, according to the participants' performance. In 10 tasks, the stimuli were presented at a single SNR level.

3.1.5.5. Spatial separation between stimuli. A spatial separation between targets and maskers (° azimuth) was present in only 9/70 tasks. Of these, there were 5 bilateral masker presentations (±90°, ±15° azimuth), 3 unilateral masker presentations (+90°, −90°, ITD of 2 ms, ITD of −2 ms) and one presentation via a circular array of loudspeakers with azimuths between 20° and 180°.

3.1.6. SPiN performance (Table 1a)

3.1.6.1. Speech in noise masker. 20 cross-sectional studies comprising 26 tasks included a speech in a noise masker condition. In the 16 studies involving young adults, M outperformed NM in 10/21 tasks (10/16 studies). In the 2 studies involving middle-aged adults, M outperformed NM in 2/2 tasks. In the one study involving older adults, M outperformed NM in 1/1 task. Finally, in the one study on young and older adults, M outperformed NM only in the older adult group (1 task). Overall, M outperformed NM in 14/20 studies (14/26 tasks).

One longitudinal study conducted on older adults (1 task) included a speech in a non-speech masker condition. This study found a greater improvement in M as compared to NM.

Meta-analysis: There were 21 data sets (19 studies) in this meta-analysis (Table 2). Note that the study of Deroche et al. (2017) reported the outcome for two different (independent) sets of participants, labelled as "exp.1" and "exp.4" and that in the study of Zhang et al. (2021), both young and older participants were recruited. As shown in Fig. 2, musicianship had a moderate but significant effect on SPiN perception (Hedge's $g = .52$; $p < .0001$). The confidence interval was narrow and did not include 0. The between-study heterogeneity across studies was moderate (Q's *p-value* = <.0001; $I^2 = 71.40\%$). Cross-validation indicated that the effect of musicianship was significant in each iteration of the analysis (all $p < .0002$; Hedge's g between 0.47 and 0.56). The moderator analysis indicated that the effect of both mean age and study quality were nonsignificant (all $p > .38$; see Table 3).

3.1.6.2. Speech in one-talker masker. There were 7 cross-sectional studies comprising 7 experimental tasks in this analysis. Of those, 6 focused on younger adults and 1 involved a broad age group. In the broad age group (1 task), no musician advantage was found. In the younger adult group, M outperformed NM in only 1/6 studies (1 task). Overall, M outperformed NM in only 1/7 studies (1/7 task). One longitudinal study on older adults (1 task) included a speech in one-talker masker condition. In this study, M showed less decline over time as compared to NM.

Meta-analysis: There were 7 studies in this analysis (Table 4). As shown in Fig. 3, the effect of musicianship was small and not significant (*Hedge's* $g = 0.08$; $p = .799$; *Q's p-value* = .004; I^2=83.99%). Cross-validation indicated that the effect of musicianship remained nonsignificant in each iteration of the analysis (all $p > .116$; Hedge's g between −0.09 and 0.09). The moderator analysis indicated that the effect of both mean age and study quality were nonsignificant (all $p > .26$) (Table 5).

3.1.6.3. Speech in two-talker masker. There were 6 cross-sectional studies comprising 10 experimental tasks in this analysis, including five studies on younger adults and one study on both young and older adults. M outperformed NM in 3 studies on young adults (4 tasks) and only older M outperformed older NM in the study on young and older

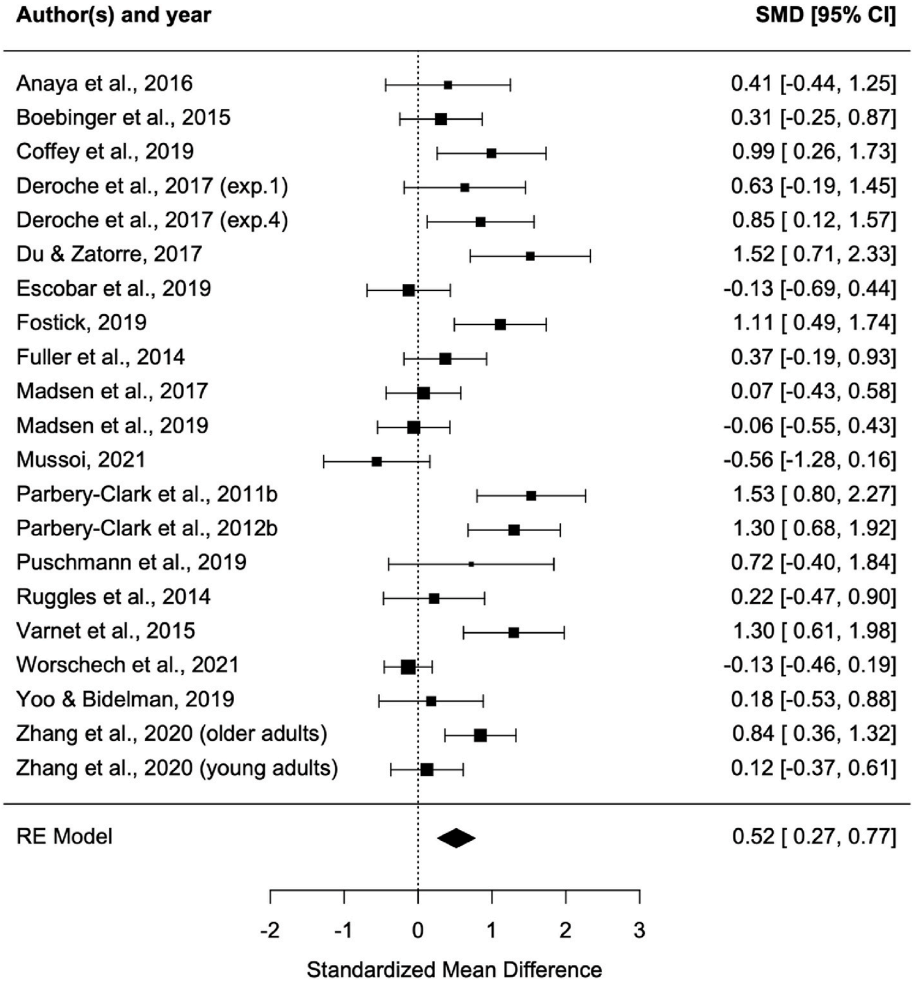


Fig. 2. Effect sizes for the speech in noise masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge's G), 95% CI = 95% confidence interval.

adult. Overall, M outperformed NM in 5/10 tasks. None of the longitudinal studies measured speech perception in a two-talker masker condition.

Meta-analysis: There were 7 data sets (5 studies) in this meta-analysis (Table 6). Note that the study of Deroche et al. (2017) reported the outcome for two different (independent) sets of participants, labelled as "exp.1" and "exp.4" and that in the study of Zhang et al. (2021), both young and older participants were recruited. As shown in Fig. 4, the effect of musicianship on this analysis was moderate and significant (*Hedge'g* = .63; *p* = .0009). The confidence interval was broad but did not include 0. Between study heterogeneity was significant and moderate (*Q's p-value* = .01; $I^2 = 61.49\%$). Cross-validation indicated that the effect of musicianship was significant in every iteration (*Hedge's g* between .50 and .76; all $p < .008$). The moderating effect of mean age and study quality were not significant (all $p > .058$) (Table 7).

3.1.6.4. Speech in four-talker masker. There were 19 tasks in this analysis from 16 cross-sectional studies; 13 studies on younger adults (15 tasks), 1 study on middle-aged adults (2 tasks), 1 study on older adults (1 task) and 1 study on a broad age group (1 task). In the studies on older adults and the broad age group, M did not outperform NM. In studies on younger adults, M outperformed NM in 8/15 tasks (8 studies). Overall, M outperformed NM in 8/16 studies (8/20 tasks). Three longitudinal studies (3 tasks) on older adults were included in this analysis and 2/3 found an improvement in the musical training group only.

Meta-analysis: There were 16 data sets (16 studies) in this meta-analysis (Table 8). As shown in Fig. 5, the effect of musicianship was moderate and significant (*Hedge's g* = 0.53; *p* = <.0001). The confidence interval was broad but did not include 0. Between study heterogeneity was significant (Q's *p-value* = .02; $I^2 = 57.30\%$). Cross-validation indicated that that the effect of musicianship was significant in each iteration (Hedge's *g* between .43 and .57; all $p < .0001$). Moderator analysis indicated that the effect of mean age and study quality were nonsignificant (all $p > .79$) (Table 9).

3.1.6.5. Speech in a spatially separated masker. There were 8 cross-sectional studies in this analysis comprising 10 tasks, 6 studies on younger adults (7 tasks), 1 study on both young and older adults and 1 study on a sample with a broad age range (1 task). In the study with the broad age sample, M did not outperform NM. M outperformed NM in the study on older adults. For the younger adults, M outperformed NM in 4/7 tasks (5 studies). Overall, M outperformed NM in 5/8 studies (5/10 tasks). None of the longitudinal study measured speech perception in spatially separated noise.

Meta-analysis: There were 7 data sets (6 studies) in this analysis (Table 10). Note that in the study of Zhang et al. (2021), both young and older participants were recruited. As seen on Fig. 6, the effect of musicianship was small and significant (*Hedge's g* = 0.35; *p* = .004). The confidence interval did not include 0. Between study heterogeneity was nonsignificant (Q's *p-value* = .133; $I^2 = 24.87\%$). Cross-validation indicated that the effect of musicianship remained significant (all $p > .013$), except when removing the data from older adults in the study

Table 4
Tasks and descriptive data used in the "speech in one-talker masker" meta-analysis.

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Baskent and Gaudrain (2016)	Sentences in sentences—composite score of all F0 and VTL modification	% correct	0.81 (0.05)	0.70 (0.12)
Boebinger et al. (2015)	Sentences in clear speech	SRT	−11.9 (2.40)	−11.5 (3.20)
Clayton et al. (2016)	Sentences in sentences	SRT	3.8(2.24)	4.4 (2.05)
Madsen et al. (2019)	Sentences in sentences	SRT	2.03 (1.49)	2.22 (1.08)
Mandikal Vasuki et al. (2016)	LiSN–S test: natural speech (talker1) in natural speech (talker 2)	SRT	−7.8 (3.50)	−7.9 (3.10)
Mankel and Bidelman (2018)	LiSN–S test: natural speech (talker1) in natural speech (talker 2)	SRT	11.9 (3.47)	11.64 (2.81)
Puschmann et al. (2019)	Sentences in sentences	SRT	6.77 (0.95)	4.96 (1.25)

Note. M = mean; SD = standard deviation, SRT = speech reception threshold, F0 = fundamental frequency, VTL = vocal tract length.

of Zhang (2021) (*Hedge's g* = 0.23; *p* = .053). Moderator analysis indicated that the effect of mean age and study quality were nonsignificant (all *p* > .455) (Table 11).

3.1.6.6. Speech in noise, SNR <0 dB. There were 6 cross-sectional studies comprising 6 experimental tasks in this analysis, 1 study on older adults and 5 studies on younger adults. Overall, M outperformed NM in 2 studies on younger adults (2/6 tasks). None of the longitudinal studies was included in this analysis.

Meta-analysis. There were 7 data sets (7 studies) in this meta-analysis (Table 12). As shown in Fig. 7, the effect of musicianship was moderate and significant (*Hedge's g* = 0.59; *p* = .01). The confidence interval was broad but did not include 0. Between study heterogeneity was moderate (Q's *p-value* = .005; I^2 = 70.57%). Cross-validation indicated that the effect size remained significant in every iteration (*Hedge's g* between 0.46 and 0.69; all *p* < *.045*). Moderators indicated that the effect of mean age and study quality were nonsignificant (all *p* > *.432)* (see Table 13).

3.1.6.7. Speech in noise, SNR = 0 dB. There were 5 cross-sectional studies comprising 5 tasks in this condition, 4 studies on younger adults and 1 study on older adults. M outperformed NM in 2/4 tasks in 2/4 studies on younger adults. M outperformed NM in 1/1 tasks (1 study) on older adults. Overall, M outperformed NM in 3/5 studies (3/5 tasks). One longitudinal study (1 task) conducted with older adults was included in this analysis, where only the M group improved over time.

Meta-analysis: There were 7 studies in this analysis (Table 14). As shown in Fig. 8, the effect of musicianship was moderate but significant (*Hedge's g* = 0.61; *p* < .0001). The confidence interval was narrow and did not include 0. Between study heterogeneity was nonsignificant (Q's *p-value* = .315; I^2= 0.00%). Cross-validation indicated that the effect was significant in each iteration of the analysis (*Hedge's g* between 0.50 and 0.73; all *p* < .0008). Moderator analysis indicated that the effect of mean age and study quality was nonsignificant (all p > .103) (Table 15).

3.1.6.8. Speech in noise, SNR> 0 dB. There were 3 cross-sectional studies comprising 3 experimental tasks in this analysis, all focusing on younger adults. M outperformed NM in 2/3 tasks, namely one task of syllables in white noise and one task of words in speech-shaped noise. Three longitudinal studies (3 tasks) in older adults were included in this analysis. 1/3 reported a benefit of musical practice, with less decline over time in M as compared to NM.

Meta-analysis: There were 5 studies in this meta-analysis (Table 16). As seen in Fig. 9, the effect of musicianship was not significant (*Hedge's*

Table 5
Moderator analysis for the "speech in one-talker masker" condition.

Moderator	ß	ß's *p-value*	ß's $CI_{95\%}$ LL	ß's $CI_{95\%}$ UL	SE	Q's *p-value*	I^2 (%)
Mean age	0.00	0.93	−0.09	0.10	0.05	0.00	89.74
Study quality	4.82	0.26	−3.64	13.29	4.32	0.00	81.78

Note. ß = regression coefficients, $CI_{95\%}$ = 95% confidence interval, LL = lower limit of the $CI_{95\%}$, UL = upper limit of the $CI_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

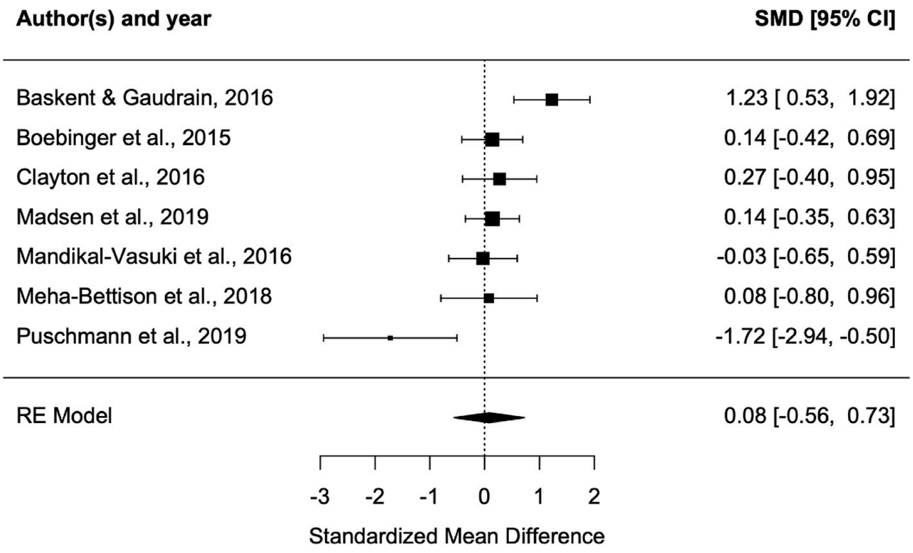


Fig. 3. Effect sizes for the speech in one-talker masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge's G), 95% CI = 95% confidence interval.

Table 6
Tasks and descriptive data used in the “speech in two-talker masker” analysis.

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Deroche et al. (2017)	Exp.1: sentences in sentences—mean score for diotic presentation	SRT	−2.76 (1.14)	−1.81 (1.20)
Deroche et al. (2017)	Exp.4: sentences in sentences—mean score for fixed masker presentation	SRT	−2.26 (1.06)	0.02 (1.79)
Kaplan et al. (2021)	Sentences in sentences—mean score of SNR conditions (−9 dB, −7 dB, −5 dB, −3 dB)	% correct	64.86 (11.57)	55.26 (15.53)
Madsen et al. (2019)	Sentences in sentences	SRT	2.41 (0.94)	2.83 (1.27)
Morse-Fortier et al. (2017)	CVC words in nonsense sentences	SNR threshold	−7.75 (5.19)	−4.56 (3.04)
Zhang et al. (2021) (older adults)	Sentences in sentences	SRT	−1.78 (1.65)	−0.31 (1.83)
Zhang et al. (2021) (young adults)	Sentences in sentences	SRT	−3.24 (1.45)	−3.51 (1.72)

Note. M = mean, SD = standard deviation, SRT = speech reception threshold, CVC = consonant-vowel-consonant.

$g = 0.28$; $p = .157$). The confidence interval included 0. Between study heterogeneity was low (Q’s *p-value* = .215; $I^2 = 47.40\%$). Cross-validation indicated that the effect of musicianship was nonsignificant in every iteration (all $p > .162$). Moderator analysis indicated that the effect of mean age and study quality were nonsignificant (all $p > .50$) (Table 17).

A summary of all analyses is presented in Table 18.

4. Discussion

The present study is the first systematic review and meta-analysis on the effect of musicianship on SPiN performance, extending the results of a detailed narrative review of 29 articles (Coffey et al., 2017). Our systematic review of 49 studies documents the state of knowledge on this topic, describing the participants and experimental approaches that have been used, evaluating their methodological quality and examining their results. With the meta-analyses (38 studies), we quantitatively assessed the effect of musicianship on SPiN in different experimental conditions, namely speech in noise masker, speech in different speech maskers, speech in a spatially separated masker and speech in noise at different SNR levels. We hypothesized that there would be a musician advantage in both noise and speech maskers, and that this advantage would increase with increasing task difficulty. Furthermore, we expected to observe a greater musician advantage in older adult musicians, especially in speech maskers. Our results show that musicianship was associated with better SPiN performance in both noise and speech masker conditions, as well as in SNR <0 dB and SNR = 0 dB conditions. However, since very few studies included participants aged ≥55 years, it was not possible to assess the effect of musicianship in an aging population. Overall, our results argue in favour of a musicianship effect on SPiN perception, consistent with the qualitative assessment conducted by Coffey et al. (2017).

4.1. Speech in noise and in speech maskers

Our main hypothesis was that there would be a musician advantage in both noise and speech maskers and that the effect would be stronger with two-talker maskers. The meta-analyses for noise maskers revealed a moderate effect size of .52. The meta-analyses for speech in different speech maskers revealed the largest effect size in the two-talker condition, with a moderate though larger effect of 0.63 (with one of the largest confidence intervals), followed by a moderate effect in the four-talker condition (0.53) and a small non-significant effect in the 1-talker condition (0.08). These observations suggest that musicianship impacts both noise and speech masking conditions with an impact of the number of talkers in the masker. Speech maskers display both noise and informational properties and their masking power varies according to the

Table 7
Moderator analysis for the “speech in two-talker masker” condition.

Moderator	ß	ß’s *p-value*	β’s $CI_{95\%}$ LL	β’s $CI_{95\%}$ UL	SE	Q’s *p-value*	I^2 (%)
Mean age	0.01	0.62	−0.02	0.03	0.01	0.01	64.23
Study quality	−2.77	0.06	−5.63	0.09	1.46	0.05	42.30

Note. ß = regression coefficients, $CI_{95\%}$ = 95% confidence interval, LL = lower limit of the $CI_{95\%}$, UL = upper limit of the $CI_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

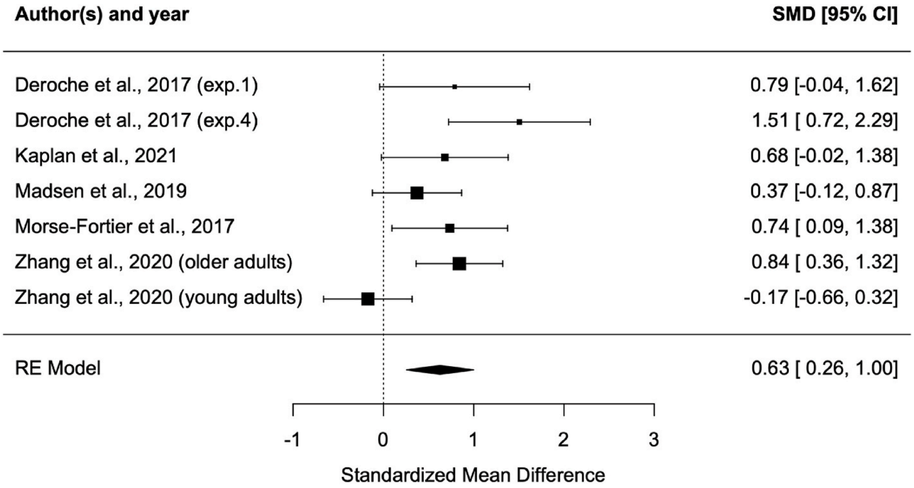


Fig. 4. Effect sizes for the speech in two-talker masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study’s weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge’s G), 95% CI = 95% confidence interval.

Table 8

Tasks and descriptive data used in the “speech in 4-talker masker” meta-analysis.

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Bidelman and Yoo (2020)	QuickSIN	SNR loss	−1 (1.29)	−0.21 (1.14)
Dubinsky et al. (2019)	QuickSIN	SNR loss	2.24 (2.10)	3.40 (2.27)
Escobar et al., 2019	QuickSIN	SNR loss	0.19 (0.85)	0.15 (0.90)
Hennessy et al. (2021)	BKB-SIN	Score	25.03 (1.04)	24.83 (1.24)
Jain and Nataraja (2019)	QuickSIN	SNR loss	−7.55 (1.75)	−6.58 (1.90)
Mussoi (2021)	QuickSIN	SNR loss	0.64 (1.20)	1.41 (1.32)
Parbery-Clark et al., 2011b	QuickSIN	SNR loss	0.22 (0.39)	−0.51 (0.38)
Perron et al., 2022a	CVC syllables in four talker babble composite score for −3 and +3 dB	Sensitivity	2(0.58)	1.92 (0.50)
Ruggles et al. (2014)	QuickSIN	SNR loss	0.8(1.03)	1.19 (0.72)
Skoe et al. (2019)	QuickSIN	SNR loss	0.61 (0.75)	0.86 (0.58)
Slater et al. (2018)	QuickSIN	SNR loss	−1.04 (0.70)	−0.36 (1.00)
Yoo & Bidelman, 2019	QuickSIN	SNR loss	−1.56 (1.81)	0.7 (1.32)
Zendel and Alain (2012)	QuickSIN	SNR loss	0.84 (1.57)	1.65 (2.50)
Zendel et al. (2015)	Words in four talker babble—composite score for 0 and 15 dB	% correct	91.67 (3.34)	88.72 (3.42)
Zendel et al., 2019	Words in four talker babble—composite score for 0 and + 15 dB	% correct	86.69 (7.57)	82.12 (6.15)
Zhang et al. (2019)	QuickSIN	SNR loss	1.48 (1.42)	1.35 (1.01)

Note. M = mean, SD = standard deviation, SRT = speech reception threshold, CVC = consonant-vowel-consonant.

characteristics and the number of talkers in the masker (Hoen et al., 2007; Rosen et al., 2013; Helfer and Freyman, 2014). Perceiving speech in a one-talker masker has been shown to be less challenging than in maskers with additional talkers (Rosen et al., 2013). Despite the informational load present in a one-talker masker, its lower masking power could be explained by the energy modulation present in the speech of a single talker, allowing more access to the target speech due to glimpses, defined as “spectrotemporal regions where the energy of the target speech exceeds that of the masker by at least 3 dB” (Cooke, 2006). The addition of a second talker in the masker reduces glimpsing opportunities while raising its informational load (as well as, presumably, its energetic masking too), leading to a dramatic decrease in SPiN performance (Tun and Wingfield, 1999; Rosen et al., 2013). However, increasing further the number of talkers in the masker to ≥4 has been shown to gradually improves SPiN performance by decreasing the informational masking power, due to a reduced capacity for the listener to decipher the content of the multiple simultaneous talkers (Hoen et al., 2007).

To summarize, our results show that the most challenging condition in terms of informational masking—the two-talker masker—is associated with the strongest musicianship SPiN advantage, with a significant but slightly lower advantage in the four-talker maskers and in the noise masker conditions, and not associated with performance on the less challenging condition, the one-talker masker. These findings show that musicianship is associated with resilience to both noise and speech maskers, which may prove helpful in many daily situations, where both types of maskers are present.

4.2. *Speech in spatially separated masker*

We hypothesized to find a lesser musician advantage in the spatially separated conditions compared to the collocated conditions. Consistent with this hypothesis, the results of the meta-analysis showed that the effect of musicianship for speech in spatially separated masker was lesser than in collocated noise, two- and four-talker maskers.

Previous studies on non-musicians have shown that spatially separating the target speech from the masker facilitates SPiN perception, a phenomenon referred to as “spatial release from masking” (SRM) (for a review, see Wang and Xu (2021)). Spatially separating stimuli improves perception by allowing binaural processes of interaural time and level differences (Glyde et al., 2011; Yost, 2017). SRM has been shown to be lesser in noise maskers than in speech maskers (Jones and Freyman, 2012). In speech maskers, SRM is inversely proportional to the number of talkers (Freyman et al., 2004; Helfer and Freyman, 2014). The result of our meta-analysis on speech in spatially separated speech maskers suggests that musicianship provides a lesser advantage for the perception of spatially separated stimuli. It is conceivable that the advantages of musicianship are mitigated in this easier condition.

4.3. *SPiN at various SNR levels*

We hypothesized to find a greater musician advantage in the most challenging listening condition—the SNR <0 dB condition. The meta-analysis revealed a moderate effect of musicianship in the SNR <0 dB condition and in the 0 dB SNR condition, and no advantage in the easier condition SNR >0 dB, in partial support of our hypothesis.

Lowering the SNR level raises the loudness of the background noise compared to the target speech, rendering it less audible. Performance in SPiN tasks drops with a decrease in the SNR in young adults; and the effect of SNR is even more detrimental in older adults, especially at SNRs ≤0 dB (Taitelbaum-Swead and Fostick, 2016; Heidari et al., 2018). This suggests that the perception of musicians is more resistant to the disruptive effect of low SNR compared to non-musicians, that is when the masker is louder (SNR <0 dB) or equal to the target (SNR = 0 dB) but not when the target is louder than the masker (SNR> 0 dB), consistent with prior experimental work (Fuller et al., 2014; Parbery-Clark et al., 2009; Parbery-Clark et al., 2009; Zendel et al., 2015). Interestingly, Parbery-Clark and colleagues showed that musical experience resulted in more robust subcortical representation of speech in the presence of noise while electrophysiological responses in quiet were similar in musicians and non-musicians, which could contribute to the musician advantage for SPiN at low SNR (Parbery-Clark et al., 2009).

A question one might ask is whether this pattern of advantage has an impact on real-life situations. Studies reporting real life SNRs calculated from recordings made from a listener’s position have shown that SNRs ≤0 dB are scarce in real life situations (Smeds et al., 2015; Wu et al., 2018; Brungart et al., 2020). However, SNRs calculated from real life situations may be overestimated, and corresponding experimental SNRs would be lower (possibly negative), especially for bars and restaurants (Brungart et al., 2020). Additional data is needed to determine whether experimental SPiN tasks mimic real life SPiN situations, and to understand the extent to which a musician advantage may facilitate communication in day-to-day situations.

4.4. *Effect of age*

We hypothesized that there would be a greater effect of musicianship with advancing age, especially in speech maskers, due to a potential cognitive reserve resulting from years of musical training, which would mitigate the normal age-related decline of speech perception. There has been much interest in music training for older adults, and music education more generally, from a rehabilitation perspective (e.g., Alain

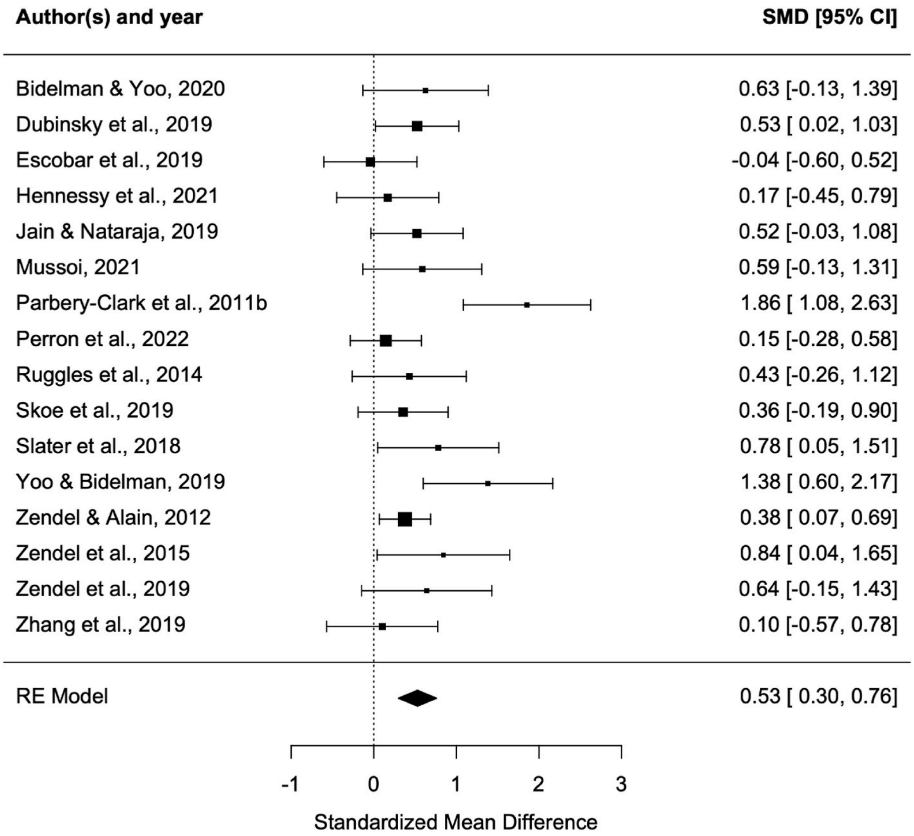


Fig. 5. Effect sizes for the speech in four-talker masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge's G), 95% CI = 95% confidence interval.

Table 9
Moderator analysis for the "speech in four-talker masker" condition.

Moderator	ß	ß's	β's $CI_{95\%}$		SE	Q's	I^2 (%)
		p-value	LL	UL		*p-value*	
Mean age	0.00	0.87	−0.01	0.02	0.01	0.01	66.51
Study quality	−0.37	0.79	−3.04	2.31	1.36	0.02	60.32

Note. ß = regression coefficients, $CI_{95\%}$ = 95% confidence interval, LL = lower limit of the $CI_{95\%}$, UL = upper limit of the $CI_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

et al., 2014; Anderson et al., 2013; Grenier et al., 2021; Kraus and Anderson, 2013a; Kraus and Anderson, 2013b; Parbery-Clark et al., 2012; Parbery-Clark et al., 2011; Perron et al., 2021; Perron et al., 2022; Zendel and Alain, 2012) but the level of evidence is currently too low to guide policies. In the present systematic review, 2/2 of the cross-sectional studies on middle-aged and 2/3 studies in older adults showed a musician advantage in noise and in speech maskers. Longitudinal studies conducted in older adults (4/6) also reported benefits of musical practice on speech perception, mostly using speech maskers. In the meta-analysis, no positive association between mean age and the effect of musicianship was observed. However, the limited number of datasets including middle-aged (55–65 years old, n = 2) and older adults (>60 years old, n = 7, accounting for an average of only 17.9% of the data included in the meta-analyses) does not warrant the interpretation of any age effect, or lack thereof, in the different experimental conditions.

Difficulties perceiving speech in noisy situations appear gradually in adulthood, beginning to emerge in midlife, in the absence of abnormal audiograms (Demeester et al., 2012). These difficulties appear predominantly in conditions of speech maskers, rather than in noise masking (for a review, see Helfer and Jesse (2021)). Echoing the literature on middle-aged adults, studies comparing young and older adults have also shown that the effect of increasing age systematically decreases the performance of speech perception in speech maskers as opposed to noise maskers for which there is little effect (Tun and Wingfield, 1999; Rajan and Cainer, 2008; Taitelbaum-Swead and Fostick, 2016; Goossens et al., 2017). It is plausible that this discrepancy in the age effect between noise and speech maskers could be partially explained by the age-related decline in cognitive faculties (Harada et al., 2013), on which speech maskers place higher demands than noise maskers (Brungart, 2001; Freyman et al., 2004, 2007; Meister et al., 2013; Fitzhugh et al., 2021). Hence, to properly assess whether musicianship mitigates the decline in speech perception, and whether this mitigation impacts the perception in noise and/or informational masker, additional studies on middle-aged and older adult comparing the performance of musicians and non-musicians in both masking conditions are required.

4.5. *Critical appraisal of the literature and future considerations*

The present review, through the evaluation of the methodological quality and the extraction of studies' characteristics, highlights some important limitations of this literature.

A first and most important limitation relates to the lack of randomized training studies, which are key to determining whether the practice of musical activities can have a beneficial impact, in a causal sense, on speech perception in noise. Only seven training studies were included, and of these, the use of a randomization method was either not reported (n = 2) or only partially described (n = 5). As pointed out by McKay (2021), without proper randomization, it is impossible to control for the influence of innate abilities that may differentiate musicians from

Table 10

Tasks and descriptive data used in the “speech in spatially separated noise” meta-analysis.

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Clayton et al. (2016)	Sentences in sentences (1-talker-babble), ±15° azimuth	SRT	−11.7 (5.14)	−6.6 (6.34)
Madsen et al. (2019)	Sentences in sentences (1-talker-babble), ±15° azimuth	SRT	−10.27 (3.55)	−9.33 (3.89)
Meha-Bettison et al. (2018)	LiSN–S test: sentences (talker1) in natural speech (talker 2), ±90° azimuth	SRT	−18.6 (2.6)	−18.1 (2.93)
Morse-Fortier et al. (2017)	CVC words in sentences (2-talker babble), +60° azimuth	SNR threshold	−25.18 (1.65)	−24.56 (1.70)
Yeend et al. (2017)	LiSN–S test: sentences (talker1) in natural speech (talker 2), ±90° azimuth	SRT	−19.59 (2.31)	−19.65 (2.74)
Zhang et al. (2021) (older adults)	Sentences in sentences	SRT	−7.29 (1.55)	−5.99 (1.51)
Zhang et al. (2021) (young adults)	Sentences in sentences	SRT	−7.23 (1.14)	−7.06 (1.81)

Note. M = mean, SD = standard deviation, SRT = speech reception threshold, SNR = signal-to-noise ratio, CVC = consonant-vowel-consonant.

non-musicians, as in cross-sectional studies. High-quality randomized controlled studies, currently lacking in the literature, are thus needed to investigate a causal relationship between the practice of musical activities and SPiN abilities. This suggestion was made six years ago by Coffey et al. (2017), and remain relevant today, as the number of such studies still remain low. Our results only attest to an *association* between musicianship and SPiN ability.

A second main limitation relates to the description of participants. The characterization of musician participants was homogeneous in terms of reporting the age of onset and the years of musical practice. However, the frequency, intensity and lifelong duration of musical practice and musical training as well as the proficiency level were largely undocumented. Additionally, the type of music played by the musicians was not always reported. Hence, it is difficult to determine which aspect of musical practice is associated with a musicianship advantage and to outline a profile of successful musical practice for SPiN perception. From the perspective of prevention and rehabilitation, the lack of information on the musicians’ practice complicates the elaboration of guidelines and recommendations. If benefits emerge only in those with an intense musical practice or those that are highly trained (e.g., those who studied music at the conservatory or university level), or perhaps in those who practise certain musical activities but not others (e.g., instrument playing vs. singing) then SPiN benefits may be out of reach to large segments of the general population. It would therefore be helpful in further studies to provide more information about the musicians (e.g., the frequency and intensity of their musical practice, the total amount of practice over their lifespan, their ability to read music, their progression over time, the proficiency level). Further, this information could be integrated into statistical analyses to identify the parameters that most influence SPiN performance. A recent study has done so and showed a benefit for those with a more frequent practice compared with those with less regular practice (Perron et al., 2022ab). Studying various types of musicians and describing these thoroughly will help understand the parameters that best promote plasticity, and in turn, SPiN enhancement. But currently, the lack of information about the musicians complicates interpretation of the findings and limit their generalizability. It could also have contributed to the high heterogeneity found in the results, with several studies not actually reporting benefits despite globally positive analyses, and with I^2 values of over 60% in 4/8 meta-analyses. The between-study heterogeneity is illustrated in different funnel plots, which show the presence of studies out of the

Table 11

Moderator analysis for the “speech in spatially separated noise” meta-analysis.

Moderator	ß	ß’s	β’s $CI_{95\%}$		SE	Q’s	I^2 (%)
		p-value	LL	UL		*p-value*	
Mean age	0.01	0.45	−0.01	0.02	0.01	0.13	34.40
Study quality	0.80	0.62	−2.32	3.91	1.59	0.10	35.43

Note. ß = regression coefficients, $CI_{95\%}$ = 95% confidence interval, LL = lower limit of the $CI_{95\%}$, UL = upper limit of the $CI_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

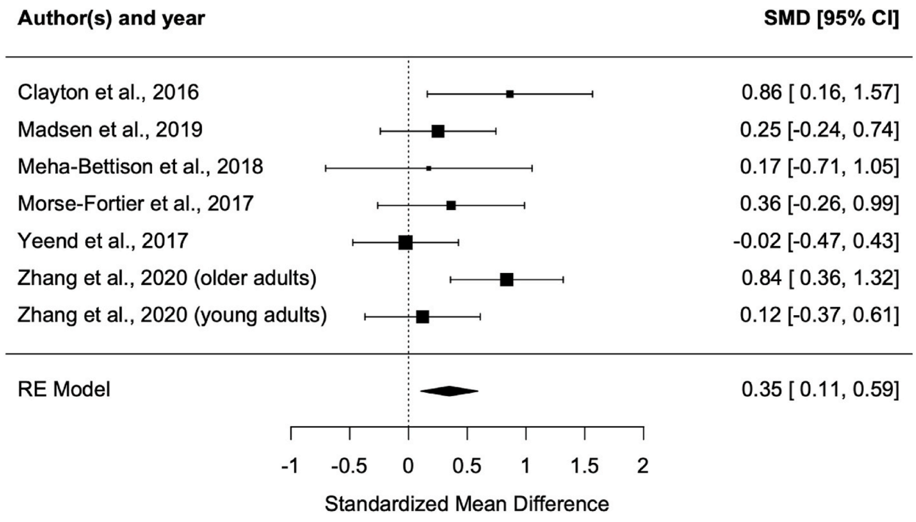


Fig. 6. Effect sizes for the speech in spatially separated masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study’s weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge’s G), 95% CI = 95% confidence interval.

Table 12
Tasks and descriptive data used in the “speech in noise—SNR <0 dB” meta-analysis.

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Anaya et al. (2016)	HINT, SNR = −3 dB	% correct	50.74 (8.59)	47.22 (8.11)
Baskent and Gaudrain (2016)	Sentences in sentences, SNR = −6 dB	% correct	0.82 (0.05)	0.70 (0.12)
Du and Zatorre (2017)	CV syllables in white noise—mean score of SNR conditions −4, −8 and −12 dB	% correct	75.21 (6.35)	65.04 (7.37)
Kaplan et al. (2021)	Sentences in sentences—mean score of SNR conditions −9 dB, −7 dB, −5 dB and −3 dB	% correct	64.86 (11.57)	55.26 (15.53)
Madsen et al. (2017)	Sentences in Gaussian noise, SNR = −3 dB	SRT	73.53 (9.59)	72.91 (6.84)
Perron et al., 2022a	Syllables in four-talker babble, SNR = −3 dB	Sensitivity	1.26 (0.55)	1.17 (0.56)
Ruggles et al. (2014)	Sentences in continuous noise—mean score of SNR conditions −6 dB and −3 dB	% correct	64.84 (5.82)	63.30 (10.20)

Note. M = mean, SD = standard deviation, SRT = speech reception threshold, SNR = signal-to-noise ratio, CV = consonant-vowel.

funnel, for instance in the funnel plot of the “speech in noise masker” meta-analysis (see Supplementary material 3).

A third major limitation is that the control for hearing was uneven between studies. While the vast majority of studies measured pure-tone thresholds (~94%), with several even taking additional hearing measures (i.e., otoscopy, ABR, tympanometry and distortion products), the issues is that musicians and non-musicians were matched on their hearing status in a surprisingly low number of studies (16/49). Further, in 36/49 studies, the specific hearing levels were not reported, that is, it was reported that participants had normal hearing, but the actual thresholds were not provided. Additionally, the definition of normal hearing varied from a study to the other (15–50 dB HL) and it was not always clear whether a given threshold for normality was tested at each frequency, across all frequency or across a subset of frequencies (i.e., pure tone average). While in studies with stricter definitions of normal (e.g., 15 or 20 dB HL), not directly comparing groups is less of an issue, in studies with more liberal definitions of normal hearing thresholds (e.g., 30, 35, 40, 45, 50 dB HL), significant group differences could be present and meaningful. The practice of a musical instrument, especially in a band/orchestra can cause hearing loss (Pouryaghoub et al., 2017; Ramrattan and Gurevich, 2020; Teie, 1998). Further, hearing is known to decline significantly with age in most adults, often beginning in middle-aged adults and becoming evident around 60 years (Gates and

Table 13
Moderator analysis for the “speech in noise—SNR <0 dB” meta-analysis.

Moderator	ß	ß's	β's CI$_{95\%}$		SE	Q's	I^2 (%)
		p-value	LL	UL		*p-value*	
Mean age	−0.02	0.43	−0.05	0.02	0.02	0.01	68.16
Study quality	−1.79	0.61	−8.73	5.14	3.54	0.01	73.74

Note. ß = regression coefficients, CI$_{95\%}$ = 95% confidence interval, LL = lower limit of the CI$_{95\%}$, UL = upper limit of the CI$_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

Table 14
Tasks and descriptive data used for the “speech in noise—SNR = 0 dB” meta-analysis.

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Anaya et al. (2016)	PRESTO	% correct	68.13 (6.11)	63.5 (7.48)
Du and Zatorre (2017)	CV syllables in white noise	% correct	92.33 (2.91)	88 (7.56)
Fostick (2019)	Words in speech shaped noise	% correct	0.72 (0.15)	0.55 (0.15)
Fuller et al. (2014)	Words in speech shaped noise	% correct	89.55 (6.72)	88 (9.14)
Ruggles et al. (2014)	Sentences in continuous noise	% correct	86.97 (6.01)	85.17 (11.73)
Zendel et al. (2015)	Words in multitalker babble noise	% correct	85.28 (5.51)	79.54 (6.10)
Zendel et al., 2019	Words in four talker babble	% correct	76.23 (12.65)	67.23 (12.15)

Note. M = mean, SD = standard deviation, SRT = speech reception threshold, CV = consonant-vowel.

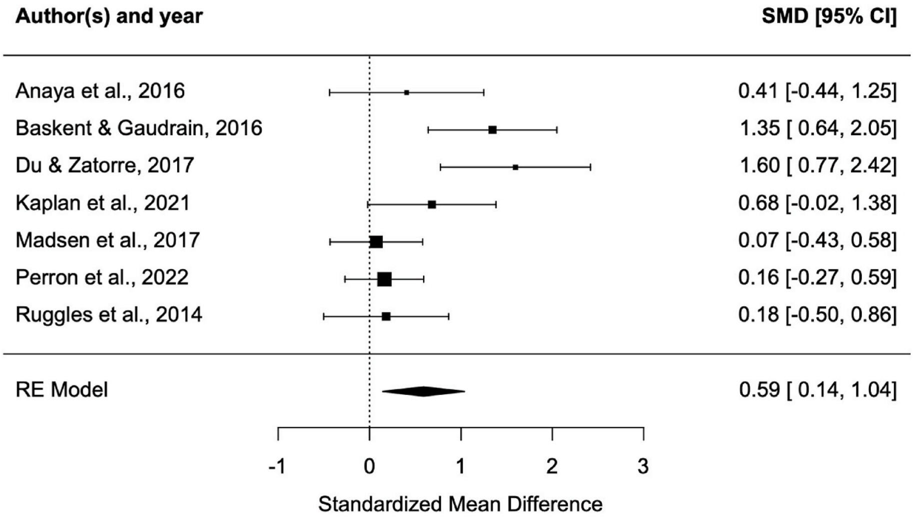


Fig. 7. Effect sizes for the speech in SNR <0 dB condition. Each box represents the effect estimate of an individual study and its size is proportional to that study’s weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge’s G), 95% CI = 95% confidence interval.

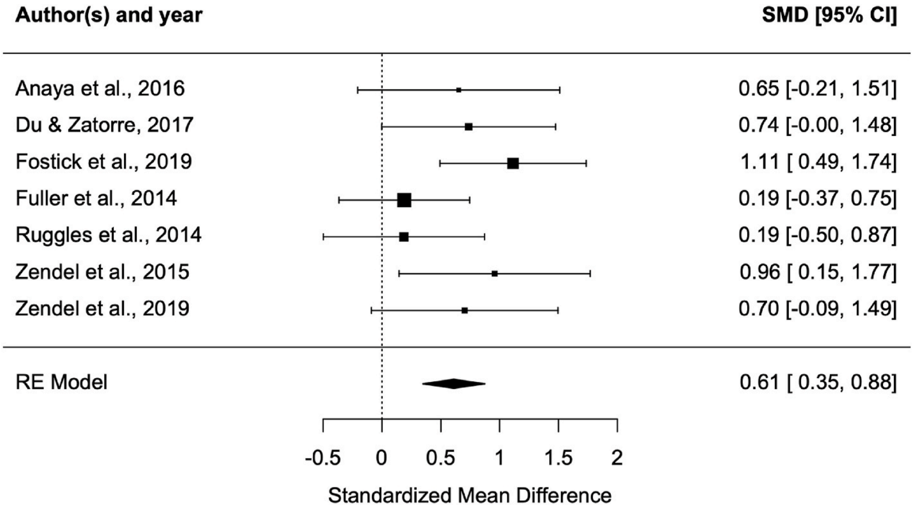


Fig. 8. Effect sizes for the speech in SNR = 0 dB condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge's G), 95% CI = 95% confidence interval.

Table 15
Moderator analysis for the "speech in noise—SNR = 0 dB" analysis.

Moderator	ß	ß's *p-value*	β's CI$_{95\%}$ LL	UL	SE	Q's *p-value*	I^2 (%)
Mean age	0.01	0.10	−0.00	0.02	0.01	0.49	0.00
Study quality	0.45	0.77	−2.60	3.50	1.55	0.22	9.54

Note. ß = regression coefficients, CI$_{95\%}$ = 95% confidence interval, LL = lower limit of the CI$_{95\%}$, UL = upper limit of the CI$_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

Table 16
Tasks and descriptive data used for the "speech in noise—SNR> 0 dB" meta-analysis.

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Du and Zatorre (2017)	CV syllables in white noise, SNR = +8 dB	% correct	97.33 (2.31)	91.5 (6.72)
Fuller et al. (2014)	Words in speech shaped noise, composite score for +5 and + 10 dB SNR	% correct	96.10 (2.54)	95.88 (2.67)
Perron et al., 2022a	CVC syllables in four-talker babble, SNR = +3 dB	Sensitivity	2.75 (0.69)	2.63 (0.58)
Zendel et al. (2015)	Words in multitalker babble, SNR = +15 dB	% correct	98.05 (1.35)	97.90 (1.21)
Zendel et al., 2019	Words in four talker babble	% correct	97.15 (3.60)	97.00 (1.41)

Note. M = mean, SD = standard deviation.

Mills, 2005), with surprisingly high number of undiagnosed and uncorrected hearing loss (Ramage-Morin et al., 2019). Finally, considering that hearing levels can influence SPiN performance, measuring, reporting and controlling for hearing is a necessity. Perhaps one attenuating factor is that, typically, the sound level during a SPiN task is adjusted to participants hearing, which limits the impact of not controlling for hearing level. Further, older adults with normal hearing also commonly complain of difficulty with SPiN, which suggests that SPiN difficulties are not entirely attributable to peripheral hearing decline (Frisina and Frisina, 1997; Gordon-Salant and Fitzgibbons, 1993; Hopkins and Moore, 2011; Snell and Frisina, 2000). Nevertheless, it is still possible that hearing level differences between groups could have attenuated or masked group differences in SPiN, which could contribute to explaining the rather large proportion of studies that did not report a robust musician's advantage, despite the overall analysis being significant in most conditions. Echoing a recent narrative review (Grenier et al., 2021), our recommendation is therefore to measure and report pure tone thresholds for each group (audiograms and the calculation of PTA). Considering that a recent literature review showed that the classical PTA frequencies (500, 1000 and 2000 Hz) are not affected in musicians, but that higher frequencies (3000–8000 Hz) are (Di Stadio et al., 2018), measurements should include frequencies up to 8 KHz. Given the high prevalence of hearing loss among musicians (estimated to be around ∼40% among professional musicians (Di Stadio et al., 2018; Pouryaghoub et al., 2017), whether abnormal hearing levels should be an exclusion criterion if one aims to maintain representativeness, is a debate that is beyond the scope of this review. However, a more thorough reporting of hearing status, and the use of a measure of hearing level (e.g. an extended PTA) in the statistical analyses would provide some control over this potentially confounding factor, as well as reveal its impact on SPiN.

A fourth major limitation is that the control for cognitive level was largely lacking. In the 20 studies that matched the groups on cognition, most used IQ. Considering that high-level cognitive and executive functions have been shown to contribute to about 9% (r ∼ 0.3) of SPiN performance (Dryden et al., 2017) and given that there could be baseline cognitive differences between musicians and non-musicians (e.g., predisposition (Schellenberg and Peretz, 2008)), measuring cognition/executive functions and matching groups on this is essential. A general screening measure could be the basis of such a matching process (e.g., the MoCA test, which is a multilingual validated tool that only takes a few minutes to administer), as was also suggested by Kraus and Anderson (Kraus and Anderson, 2013c).

This would also ensure that the musicianship advantage that we report here is truly related to music-making activities and not to genetic or other predispositions. More in depth measures of cognitive functions that have been associated with SPiN perception such as auditory working memory, auditory attention, and speed of processing (e.g., Bidelman and Yoo, 2020; Dryden et al., 2017; Parbery-Clark et al., 2011; Strait and Kraus, 2011) and an evaluation of their impact on SPiN performance would also contribute to reveal the extent to which these domains are impacted by musicianship, if at all, and their contribution to SPiN perception. In their narrative review, Coffey et al. suggested that

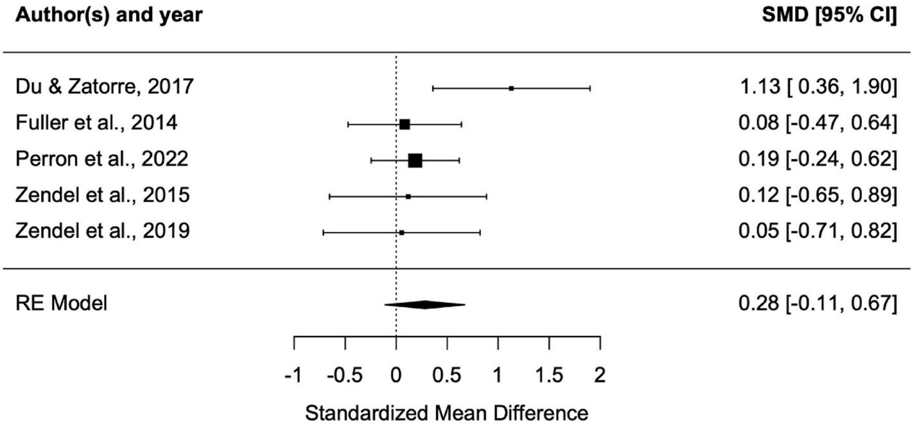


Fig. 9. Effect sizes for the speech in SNR> 0 dB condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge's G), 95% CI = 95% confidence interval.

Table 17
Moderator analysis for the "speech in noise—SNR> 0 dB" analysis.

Moderator	ß	ß's	β's CI$_{95\%}$		SE	Q's	I^2 (%)
		p-value	LL	UL		*p-value*	
Mean age	−0.01	0.50	−0.03	0.01	0.01	0.16	52.49
Study quality	0.46	0.83	−3.74	4.66	2.14	0.13	60.33

Note. ß = regression coefficients, CI$_{95\%}$ = 95% confidence interval, LL = lower limit of the CI$_{95\%}$, UL = upper limit of the CI$_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

cognitive performance enhance SPiN performance but that even in SPiN tasks with low cognitive load, a musician advantage has been reported, which suggest that, though important, a musician advantage in SPiN may not be fully explained by higher-level processes (Coffey et al., 2017). Together with the current observation that controls for cognitive level are scarce in the SPiN literature, these findings show that additional studies are needed to determine the part of the musician advantage on SPiN that is related to cognitive and executive functions.

4.6. *Limitations*

There are a few limitations to our study. Our analyses and conclusions are limited by the available literature: (1) a restricted number of studies, (2) frequently missing descriptive statistics (means and standard deviations), (3) variable quality of the studies included, (4) several studies reported no reliable group difference, and (5) a potential publication bias. (1) The limited number of studies (5–21 per analysis) impacts the reliability of the effect sizes, which should be interpreted with some caution. Because of the paucity of studies on middle-aged and older adults, the moderating impact of age could not be consistently assessed in every analysis. In addition, the effect of masker type (noise VS. speech) in analyses of different SNR levels ((f) to h)) could not be analyzed because of the small number of studies. Furthermore, the assessment of other potentially moderating variables, such as the targets' linguistic level (e.g., words vs. sentences) were also impossible, limiting our ability to explore the between-study heterogeneity. (2) We were not able to obtain the descriptive data for all studies. Specifically, despite our efforts in reaching out to authors, 7 studies that were eligible for our meta-analyses could not be included. As a result, the effect sizes do not reflect the entirety of the literature selected for the systematic review. We strongly encourage authors to report means, standard deviations, and total sample for each primary and secondary outcome measure, following most guidelines for reporting quantitative research such as those of the Publication Manual of the APA, and, whenever ethically possible, to share anonymized data on dedicated platforms to facilitate reanalyses including meta-analyses. (3) We chose to include all studies that provided data in our meta-analyses, even those with lower quality ratings, for the sake of completeness, and because of the overall low number of studies. It should be noted that, due to the issues discussed in the previous section, the average score across all studies was not very high (0.58). When further studies are available, it will be important to validate the current findings with hopefully better controlled studies including more thoroughly described participants. (4) A relatively large number of studies reported no reliable group difference despite the meta-analysis being significant. In several cases, the confidence interval of these studies barely passed through zero suggesting that they were near significance, which could explain in part this finding. A lack of statistical power might have occurred in some studies due to relatively small sample sizes (i.e., 20 or fewer participants in most studies), as pointed out in the results section. Some tools (e.g., https://shiny.york.ac.uk/powercontours/) may help researchers determine sample sizes as well as the number of trials in experimental SPiN tasks to

Table 18
Summary statistics for all meta-analyses.

Experimental condition	N	G	G's	SE	G's CI$_{95\%}$		Q's	I^2 (%)	I^2's CI$_{95\%}$	
			p-value		LL	UL	*p-value*		LL	UL
Speech in noise masker	21	**0.52**	**<0.01**	0.13	0.27	0.77	<0.01	71.40	51.10	86.94
Speech in one-talker masker	7	0.08	0.80	0.33	−0.56	0.73	<0.01	83.99	33.54	96.70
Speech in two-talker masker	7	**0.63**	**<0.01**	0.19	0.26	1.00	0.01	61.49	12.08	92.46
Speech in four-talker masker	16	**0.53**	**<0.01**	0.12	0.30	0.76	0.02	57.30	0.30	82.56
Speech in spatially separated noise	7	**0.35**	**<0.01**	0.12	0.11	0.59	0.13	24.87	0.00	86.39
Speech in noise—SNR <0 db	7	**0.59**	**<0.01**	0.23	0.14	1.04	0.01	70.57	23.19	94.14
Speech in noise—SNR = 0 db	7	**0.61**	**<0.01**	0.14	0.35	0.88	0.31	0.00	0.00	79.17
Speech in noise—SNR> 0 db	5	0.28	0.16	0.20	−0.11	0.67	0.21	47.40	0.00	93.96

Note. N = number of articles, G = Hedge's *g*, SE = standard error, CI$_{95\%}$ = 95% confidence interval, LL = lower limit, UL = upper limit, Q = Cochran's Q.

maximize statistical power (Baker et al., 2021). (5) Finally, we cannot rule out a publication bias, specifically for small studies, as can be noted by the visual inspection of the funnel plot for the "speech in four-talker masker" meta-analysis (see Supplementary material 3). Such a publication bias could have inflated the effect size in favour of musicians in this condition. The consequences of publication bias are profound and detrimental. We strongly encourage authors to publish their results whether consistent or not with the notion of a musician's advantage. There are several avenues nowadays where solid but nonsignificant results can be published. Pre-registering studies may be a way to protect oneself against the hardship that publishing null findings sometimes represent.

5. Conclusion

The study of a potential musician advantage on SPiN performance has garnered a wide interest for the prevention/rehabilitation of SPiN perception in the aging population. Despite the relatively small number of studies on the topic, and the quasi absence of randomized studies, our results show a robust association between musicianship and SPiN performance in challenging conditions, while this association is reduced or not significant in easier conditions. This pattern of results is compelling and supports the notion of a musician advantage. However, the current data does not warrant conclusion regarding older adults. Additional studies on older adults are needed to confirm and refine the present results. Further, randomized studies are critically needed to assess the causality of this relationship in older adults, but also, more generally, in adults of all ages. Finally, a more careful characterization of musicians, in terms of their hearing level, their cognitive level and their musical practice, would hone our understanding of the practices associated with benefits in SPiN performance and contribute to the elaboration of prevention/rehabilitation strategies.

CRediT authorship contribution statement

Elisabeth Maillard: Conceptualization, Methodology, Investigation, Validation, Project administration, Formal analysis, Visualization, Writing – original draft. **Marilyne Joyal:** Methodology, Validation, Investigation, Formal analysis, Visualization, Writing – original draft. **Micah M. Murray:** Writing – review & editing. **Pascale Tremblay:** Conceptualization, Methodology, Investigation, Validation, Supervision, Resources, Project administration, Writing – original draft, Data curation.

Declaration of competing interest

Pascale Tremblay reports financial support was provided in the form of a research grant by the Natural Sciences and Engineering Research Council of Canada.

Data availability

The data and R scripts are available on on Borealis, the *Canadian Dataverse Repository* (DOI: https://doi.org/10.5683/SP3/TQAUAS).

Acknowledgments

The project was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC grant # RGPIN-2019-06534) to PT. EM was funded by a graduate scholarship from the Fonds Theodor and Gabriela Kummer (Université de Lausanne, Switzerland).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crneur.2023.100083.

References

Alain, C., Zendel, B.R., Hutka, S., Bidelman, G.M., 2014. Turning down the noise: the benefit of musical training on the aging auditory brain. Hear. Res. 308, 162–173. https://doi.org/10.1016/j.heares.2013.06.008.

Anaya, E.M., Pisoni, D.B., Kronenberger, W.G., 2016. Long-term musical experience and auditory and visual perceptual abilities under adverse conditions. J Acoust Soc Am 140 (3), 2074. https://doi.org/10.1121/1.4962628.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., Kraus, N., 2013. Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. J. Speech Lang. Hear. Res. 56 (1), 31–43. https://doi.org/10.1044/1092-4388(2012/12-0043.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., Kraus, N., 2013. Reversal of age-related neural timing delays with training. Proc. Natl. Acad. Sci. U. S. A. 110 (11), 4357–4362. https://doi.org/10.1073/pnas.1213555110.

Arbogast, T.L., Mason, C.R., Kidd Jr., G., 2005. The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 117 (4 Pt 1), 2169–2180. https://doi.org/10.1121/1.1861598.

Baker, D.H., Vilidaite, G., Lygo, F.A., Smith, A.K., Flack, T.R., Gouws, A.D., Andrews, T. J., 2021. Power contours: optimising sample size and precision in experimental psychology and human neuroscience. Psychol. Methods 26 (3), 295–314. https://doi:10.1037/met0000337.

Baskent, D., Gaudrain, E., 2016. Musician advantage for speech-on-speech perception. J Acoust Soc Am 139 (3), EL51–56. https://doi.org/10.1121/1.4942628.

Bidelman, G.M., Yoo, J., 2020. Musicians show improved speech segregation in competitive, multi-talker cocktail party scenarios. Front. Psychol. 11 (1927) https://doi.org/10.3389/fpsyg.2020.01927.

Bilodeau-Mercure, M., Lortie, C.L., Sato, M., Guitton, M.J., Tremblay, P., 2015. The neurobiology of speech perception decline in aging. Brain Struct. Funct. 220 (2), 979–997. https://doi.org/10.1007/s00429-013-0695-3.

Boebinger, D., Evans, S., Rosen, S., Lima, C.F., Manly, T., Scott, S.K., 2015. Musicians and non-musicians are equally adept at perceiving masked speech. J Acoust Soc Am 137 (1), 378–387. https://doi.org/10.1121/1.4904537.

Boothroyd, A., 1968. Statistical theory of the speech discrimination score. J. Acoust. Soc. Am. 43 (2), 362–367. https://doi.org/10.1121/1.1910787.

03/01 Borenstein, M., Hedges, L., Higgins, J., Rothstein, H., 2009. An introduction to meta-analysis. Introduction to Meta-Analysis 19. https://doi.org/10.1002/9780470743386.

Brouwer, S., Van Engen, K.J., Calandruccio, L., Bradlow, A.R., 2012. Linguistic contributions to speech-on-speech masking for native and non-native listeners: language familiarity and semantic content. J. Acoust. Soc. Am. 131 (2), 1449–1464. https://doi.org/10.1121/1.3675943.

Brungart, D.S., 2001. Informational and energetic masking effects in the perception of two simultaneous talkers. J. Acoust. Soc. Am. 109 (3), 1101–1109. https://doi.org/10.1121/1.1345696.

Brungart, D.S., Barrett, M.E., Cohen, J.I., Fodor, C., Yancey, C.M., Gordon-Salant, S., 2020. Objective assessment of speech intelligibility in crowded public spaces. Ear Hear. 41 (Suppl. 1), 68s–78s. https://doi.org/10.1097/aud.0000000000000943. Suppl 1.

Burke, S.N., Barnes, C.A., 2006. Neural plasticity in the ageing brain. Nat. Rev. Neurosci. 7 (1), 30–40. https://doi.org/10.1038/nrn1809.

Buss, E., Hodge, S.E., Calandruccio, L., Leibold, L.J., Grose, J.H., 2019. Masked sentence recognition in children, young adults, and older adults: age-dependent effects of semantic context and masker type. Ear Hear. 40 (5), 1117–1126. https://doi.org/10.1097/AUD.0000000000000692.

Cameron, S., Dillon, H., 2007. Development of the listening in spatialized noise-sentences test (LISN-S). Ear Hear. 28 (2), 196–211. https://doi.org/10.1097/AUD.0b013e318031267f.

Clayton, K.K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A.D., Kidd Jr., G., 2016. Executive function, visual attention and the cocktail party problem in musicians and non-musicians. PLoS One 11 (7), e0157638. https://doi.org/10.1371/journal.pone.0157638.

Coffey, E.B.J., Arseneau-Bruneau, I., Zhang, X., Zatorre, R.J., 2019. The music-in-noise task (MINT): a tool for dissecting complex auditory perception. Front Neurosci 13, 199. https://doi.org/10.3389/fnins.2019.00199.

Coffey, E.B.J., Chepesiuk, A.M.P., Herholz, S.C., Baillet, S., Zatorre, R.J., 2017. Neural correlates of early sound encoding and their relationship to speech-in-noise perception. Front. Neurosci. 11, 479. https://doi.org/10.3389/fnins.2017.00479.

Coffey, E.B.J., Mogilever, N.B., Zatorre, R.J., 2017. Speech-in-noise perception in musicians: a review. Hear. Res. 352, 49–69. https://doi.org/10.1016/j.heares.2017.02.006.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, second ed. Routledge. https://doi.org/10.4324/9780203771587.

Cooke, M., 2006. A glimpsing model of speech perception in noise. J. Acoust. Soc. Am. 119 (3), 1562–1573. https://doi.org/10.1121/1.2166600.

D'Souza, A.A., Moradzadeh, L., Wiseheart, M., 2018. Musical training, bilingualism, and executive function: working memory and inhibitory control. Cogn Res Princ Implic 3 (1), 11. https://doi.org/10.1186/s41235-018-0095-6.

May 21 Deguchi, C., Boureux, M., Sarlo, M., Besson, M., Grassi, M., Schon, D., Colombo, L., 2012. Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: behavioral, electrophysiological and psychoacoustic study. Brain Res. 1455, 75–89. https://doi.org/10.1016/j.brainres.2012.03.034.

Demeester, K., Topsakal, V., Hendrickx, J.-J., Fransen, E., van Laer, L., Van Camp, G., Van de Heyning, P., van Wieringen, A., 2012. Hearing disability measured by the speech, spatial, and qualities of hearing scale in clinically normal-hearing and hearing-impaired middle-aged persons, and disability screening by means of a reduced SSQ (the SSQ5). Ear Hear. 33 (5), 615–616. https://doi.org/10.1097/AUD.0b013e31824e0ba7.

Deroche, M.L.D., Limb, C.J., Chatterjee, M., Gracco, V.L., 2017. Similar abilities of musicians and non-musicians to segregate voices by fundamental frequency. J. Acoust. Soc. Am. 142 (4), 1739–1755. https://doi.org/10.1121/1.5005496.

Desjardins, J.L., Doherty, K.A., 2013. Age-related changes in listening effort for various types of masker noises. Ear Hear. 34 (3), 261–272. https://doi.org/10.1097/AUD.0b013e31826d0ba4.

Sep. 26 Di Stadio, A., Dipietro, L., Ricci, G., Della Volpe, A., Minni, A., Greco, A., de Vincentiis, M., Ralli, M., 2018. hearing loss, tinnitus, hyperacusis, and diplacusis in professional musicians: a systematic review. Int. J. Environ. Res. Publ. Health 15 (10). https://doi.org/10.3390/ijerph15102120.

Donai, J.J., Jennings, M.B., 2016. Gaps-in-noise detection and gender identification from noise-vocoded vowel segments: comparing performance of active musicians to non-musicians. J. Acoust. Soc. Am. 139 (5), EL128–EL134. https://doi.org/10.1121/1.4947070.

Dryden, A., Allen, H.A., Henshaw, H., Heinrich, A., 2017. The association between cognitive performance and speech-in-noise perception for adult listeners: a systematic literature review and meta-analysis. Trends Hear 21. https://doi.org/10.1177/2331216517744675, 2331216517744675.

Du, Y., Zatorre, R.J., 2017. Musical training sharpens and bonds ears and tongue to hear speech better. Proc Natl Acad Sci U S A 114 (51), 13579–13584. https://doi.org/10.1073/pnas.1712223114.

Dubinsky, E., Wood, E.A., Nespoli, G., Russo, F.A., 2019. Short-term choir singing supports speech-in-noise perception and neural pitch strength in older adults With age-related hearing loss. Front Neurosci 13, 1153. https://doi.org/10.3389/fnins.2019.01153.

Escobar, J., Mussoi, B.S., Silberer, A.B., 2020. The effect of musical training and working memory in adverse listening situations. Ear Hear 41 (2), 278–288. https://doi.org/10.1097/AUD.0000000000000754.

Fitzhugh, M.C., Schaefer, S.Y., Baxter, L.C., Rogalsky, C., 2021. Cognitive and neural predictors of speech comprehension in noisy backgrounds in older adults. Lang Cogn Neurosci 36 (3), 269–287. https://doi.org/10.1080/23273798.2020.1828946.

Fleming, D., Belleville, S., Peretz, I., West, G., Zendel, B.R., 2019. The effects of short-term musical training on the neural processing of speech-in-noise in older adults. Brain Cogn 136, 103592. https://doi.org/10.1016/j.bandc.2019.103592.

Fostick, L., 2019. Card playing enhances speech perception among aging adults: comparison with aging musicians. Eur J Ageing 16 (4), 481–489. https://doi.org/10.1007/s10433-019-00512-2.

Jun 1) Fostick, L., Babkoff, H., Zukerman, G., 2014. Effect of 24 hours of sleep deprivation on auditory and linguistic perception: a comparison among young controls, sleep-deprived participants, dyslexic readers, and aging adults. J. Speech Lang. Hear. Res. 57 (3), 1078–1088. https://doi.org/10.1044/1092-4388(2013/13–0031.

Freyman, R.L., Helfer, K.S., McCall, D.D., Clifton, R.K., 1999. The role of perceived spatial separation in the unmasking of speech. J. Acoust. Soc. Am. 106 (6), 3578–3588. https://doi.org/10.1121/1.428211.

Freyman, R.L., Balakrishnan, U., Helfer, K.S., 2001. Spatial release from informational masking in speech recognition. J. Acoust. Soc. Am. 109 (5 Pt 1), 2112–2122. https://doi.org/10.1121/1.1354984.

Freyman, R.L., Balakrishnan, U., Helfer, K.S., 2004. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. J. Acoust. Soc. Am. 115 (5), 2246–2256. https://doi.org/10.1121/1.1689343.

Freyman, R.L., Helfer, K.S., Balakrishnan, U., 2007. Variability and uncertainty in masking by competing speech. J. Acoust. Soc. Am. 121 (2), 1040–1046. https://doi.org/10.1121/1.2427117.

Frisina, D.R., Frisina, R.D., 1997. Speech recognition in noise and presbycusis: relations to possible neural mechanisms [clinical trial comparative study controlled clinical trial research support, U.S. Gov't, P.H.S.]. Hear. Res. 106 (1–2), 95–104. http://www.ncbi.nlm.nih.gov/pubmed/9112109.

Fuller, C.D., Galvin 3rd, J.J., Maat, B., Free, R.H., Baskent, D., 2014. The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? Front. Neurosci. 8, 179. https://doi.org/10.3389/fnins.2014.00179.

Sep 24-30) Gates, G.A., Mills, J.H., 2005. Presbycusis, submitted for publication Lancet 366 (9491), 1111–1120. https://doi.org/10.1016/S0140-6736(05)67423-5.

Gilbert, J.L., Tamati, T.N., Pisoni, D.B., 2013. Development, reliability, and validity of PRESTO: a new high-variability sentence recognition test. J. Am. Acad. Audiol. 24 (1), 26–36. https://doi.org/10.3766/jaaa.24.1.4.

Glyde, H., Hickson, L., Cameron, S., Dillon, H., 2011. Problems hearing in noise in older adults: a review of spatial processing disorder. Trends Amplif. 15 (3), 116–126. https://doi.org/10.1177/1084713811424885.

Goossens, T., Vercammen, C., Wouters, J., van Wieringen, A., 2017. Masked speech perception across the adult lifespan: impact of age and hearing impairment. Hear. Res. 344, 109–124. https://doi.org/10.1016/j.heares.2016.11.004.

Gordon-Salant, S., Cole, S.S., 2016. Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. Ear Hear. 37 (5), 593–602. https://doi.org/10.1097/AUD.0000000000000316.

Gordon-Salant, S., Fitzgibbons, P.J., 1993. Temporal factors and speech recognition performance in young and elderly listeners. J. Speech Lang. Hear. Res. 36 (6), 1276–1285. https://doi.org/10.1044/jshr.3606.1276.

Grassi, M., Meneghetti, C., Toffalini, E., Borella, E., 2017. Auditory and cognitive performance in elderly musicians and nonmusicians. PLoS One 12 (11), e0187881. https://doi.org/10.1371/journal.pone.0187881.

Grenier, A.S., Lafontaine, L., Sharp, A., 2021. Use of music therapy as an audiological rehabilitation tool in the elderly population: a mini-review. Front. Neurosci. 15, 662087 https://doi.org/10.3389/fnins.2021.662087.

Groussard, M., Coppalle, R., Hinault, T., Platel, H., 2020. Do musicians have better mnemonic and executive performance than actors? Influence of regular musical or theater practice in adults and in the elderly. Front. Hum. Neurosci. 14, 557642 https://doi.org/10.3389/fnhum.2020.557642.

Hanna-Pladdy, B., Gajewski, B., 2012. Recent and past musical activity predicts cognitive aging variability: direct comparison with general lifestyle activities. Front. Hum. Neurosci. 6, 198. https://doi.org/10.3389/fnhum.2012.00198.

Harada, C.N., Natelson Love, M.C., Triebel, K.L., 2013. Normal cognitive aging. Clin. Geriatr. Med. 29 (4), 737–752. https://doi.org/10.1016/j.cger.2013.07.002.

Heidari, A., Moossavi, A., Yadegari, F., Bakhshi, E., Ahadi, M., 2018. Effects of age on speech-in-noise identification: subjective ratings of hearing difficulties and encoding of fundamental frequency in older adults. J Audiol Otol 22 (3), 134–139. https://doi.org/10.7874/jao.2017.00304.

Heine, C., Browning, C.J., 2002. Communication and psychosocial consequences of sensory loss in older adults: overview and rehabilitation directions. Disabil. Rehabil. 24 (15), 763–773. https://doi.org/10.1080/09638280210129162.

Helfer, K.S., Freyman, R.L., 2014. Stimulus and listener factors affecting age-related changes in competing speech perception. J. Acoust. Soc. Am. 136 (2), 748–759. https://doi.org/10.1121/1.4887463.

Helfer, K.S., Jesse, A., 2021. Hearing and speech processing in midlife. Hear. Res. 402 https://doi.org/10.1016/j.heares.2020.108097.

Hennessy, S., Wood, A., Wilcox, R., Habibi, A., 2021. Neurophysiological improvements in speech-in-noise task after short-term choir training in older adults. Aging (Albany NY) 13 (7), 9468–9495. https://doi.org/10.18632/aging.202931.

Herholz, Sibylle C., Zatorre, Robert J., 2012. Musical training as a framework for brain plasticity: behavior, function, and structure. Neuron 76 (3), 486–502. https://doi.org/10.1016/j.neuron.2012.10.011.

Sep. 6 Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. BMJ 327 (7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.

Hoen, M., Meunier, F., Grataloup, C.-L., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X., Collet, L., 2007. Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. Speech Commun. 49 (12), 905–916. https://doi.org/10.1016/j.specom.2007.05.008.

Hopkins, K., Moore, B.C., 2011. The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. J. Acoust. Soc. Am. 130 (1), 334–349. https://doi.org/10.1121/1.3585848.

Huizenga, H.M., Visser, I., Dolan, C.V., 2011. Testing overall and moderator effects in random effects meta-regression. Br. J. Math. Stat. Psychol. 64 (Pt 1), 1–19. https://doi.org/10.1348/000711010X522687.

Humes, L.E., 2021. Factors underlying individual differences in speech-recognition threshold (SRT) in noise among older adults. Front. Aging Neurosci. 13, 702739 https://doi.org/10.3389/fnagi.2021.702739.

Jain, S., Nataraja, N.P., 2019. The Effect of Fatigue on Working Memory and Auditory Perceptual Abilities in Trained Musicians. Am J Audiol 28 (2S), 483–494. https://doi.org/10.1044/2019_AJA-IND50-18-0102.

Jones, J.A., Freyman, R.L., 2012. Effect of priming on energetic and informational masking in a same-different task. Ear Hear. 33 (1), 124–133. https://doi.org/10.1097/AUD.0b013e31822b5bee.

Kaplan, E.C., Wagner, A.E., Toffanin, P., Baskent, D., 2021. Do musicians and non-musicians differ in speech-on-speech processing? Front Psychol 12, 623787. https://doi.org/10.3389/fpsyg.2021.623787.

Kidd, G., Mason, C.R., Richards, V.M., Gallun, F.J., Durlach, N.I., 2008. Informational masking. In: Yost, W.A., Popper, A.N., Fay, R.R. (Eds.), Auditory Perception of Sound Sources. Springer US, pp. 143–189. https://doi.org/10.1007/978-0-387-71305-2_6.

Killion, M.C., Niquette, P.A., Gudmundsen, G.I., Revit, L.J., Banerjee, S., 2004. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 116 (4 Pt 1), 2395–2405. https://doi.org/10.1121/1.1784440.

Kmet, L., Lee, R., Cook, L., 2004. Standard Quality Assessment Criteria for Evaluating Primary Research Papers for a Variety of Fields. AHFMR—HTA Initiative #13). http://www.ahfmr.ab.ca/download.php/.

Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15 (2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Kraus, N., Anderson, S., 2013a. Hearing matters: in older adults, the brain can still Be trained to hear in noise. Hear. J. 66 (5), 32.

Kraus, N., Anderson, S., 2013b. Hearing matters: music training: an antidote for aging? Hear. J. 66 (3), 52. https://doi.org/10.1097/01.HJ.0000427538.01582.c4.

Kraus, N., Anderson, S., 2013c. Hearing matters: the auditory–cognitive system: to screen, or not to screen. Hear. J. 66 (7), 36. https://doi.org/10.1097/01.HJ.0000432405.85455.91.

Kraus, N., White-Schwoch, T., 2017. Neurobiology of everyday communication: what have we learned from music? Neuroscientist 23 (3), 287–298. https://doi.org/10.1177/1073858416653593.

Kraus, N., Skoe, E., Parbery-Clark, A., Ashley, R., 2009. Experience-induced malleability in neural encoding of pitch, timbre, and timing. Ann. N. Y. Acad. Sci. 1169, 543–557. https://doi.org/10.1111/j.1749-6632.2009.04549.x.

Kuhnis, J., Elmer, S., Meyer, M., Jancke, L., 2013. The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: an EEG study. Neuropsychologia 51 (8), 1608–1618. https://doi.org/10.1016/j.neuropsychologia.2013.04.007.

Kumar, P.V., Rana, B., Krishna, R., 2014. Temporal processing in musicians and non-musicians [journal article]. J. Hear. Sci. 4 (3), 35–42. In: https://www.journalofhearingscience.com/TEMPORAL-PROCESSING-IN-MUSICIANS-nAND-NON-MUSICIANS, 120632,0,2.html.

Lee, L., Packer, T.L., Tang, S.H., Girdler, S., 2008. Self-management education programs for age-related macular degeneration: a systematic review. Australas. J. Ageing 27 (4), 170–176. https://doi.org/10.1111/j.1741-6612.2008.00298.x.

Oct 17) Leibold, L.J., 2017. Speech perception in complex acoustic environments: developmental effects. J. Speech Lang. Hear. Res. 60 (10), 3001–3008. https://doi.org/10.1044/2017_JSLHR-H-17-0070.

Jul 21 Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gotzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 6 (7), e1000100. https://doi.org/10.1371/journal.pmed.1000100.

Madsen, S.M.K., Marschall, M., Dau, T., Oxenham, A.J., 2019. Speech perception is similar for musicians and non-musicians across a wide range of conditions. Sci Rep 9 (1), 10404. https://doi.org/10.1038/s41598-019-46728-1.

Madsen, S.M.K., Whiteford, K.L., Oxenham, A.J., 2017. Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds. Sci Rep 7 (1), 12624. https://doi.org/10.1038/s41598-017-12937-9.

Maharaj, S., Harding, R., 2016. The needs, models of care, interventions and outcomes of palliative care in the Caribbean: a systematic review of the evidence. BMC Palliat. Care 15, 9. https://doi.org/10.1186/s12904-016-0079-6.

Mandikal Vasuki, Sharma, M., Demuth, K., Arciuli, J., 2016. Musicians' edge: A comparison of auditory processing, cognitive abilities and statistical learning. Hear Res 342, 112–123. https://doi.org/10.1016/j.heares.2016.10.008.

Mankel, K., Bidelman, G.M., 2018. Inherent auditory skills rather than formal music training shape the neural encoding of speech. Proc Natl Acad Sci U S A 115 (51), 13129–13134. https://doi.org/10.1073/pnas.1811793115.

Mansens, D., Deeg, D.J.H., Comijs, H.C., 2018. The association between singing and/or playing a musical instrument and cognitive functions in older adults. Aging Ment. Health 22 (8), 964–971. https://doi.org/10.1080/13607863.2017.1328481.

McKay, C.M., 2021. No evidence that music training benefits speech perception in hearing-impaired listeners: a systematic review. Trends Hear 25, 1–16. https://doi.org/10.1177/2331216520985678.

Meha-Bettison, K., Sharma, M., Ibrahim, R.K., Mandikal Vasuki, P.R., 2018. Enhanced speech perception in noise and cortical auditory evoked potentials in professional musicians. Int J Audiol 57 (1), 40–52. https://doi.org/10.1080/14992027.2017.1380850.

Meister, H., Schreitmuller, S., Grugel, L., Ortmann, M., Beutner, D., Walger, M., Meister, I.G., 2013. Cognitive resources related to speech recognition with a competing talker in young and older listeners. Neuroscience 232, 74–82. https://doi.org/10.1016/j.neuroscience.2012.12.006.

Merten, N., Fischer, M.E., Dillard, L.K., Klein, B.E.K., Tweed, T.S., Cruickshanks, K.J., 2021. Benefit of musical training for speech perception and cognition later in life. J Speech Lang Hear Res 64 (7), 2885–2896. https://doi.org/10.1044/2021_JSLHR-20-00588.

Micheyl, C., Delhommeau, K., Perrot, X., Oxenham, A.J., 2006. Influence of musical and psychoacoustical training on pitch discrimination. Hear. Res. 219 (1–2), 36–47. https://doi.org/10.1016/j.heares.2006.05.004.

Moore, D.R., Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R.H., Munro, K.J., 2014. Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. PLoS One 9 (9), e107720. https://doi.org/10.1371/journal.pone.0107720.

Morse-Fortier, C., Parrish, M.M., Baran, J.A., Freyman, R.L., 2017. The effects of musical training on speech detection in the presence of informational and energetic masking. Trends Hear 21. https://doi.org/10.1177/2331216517739427, 2331216517739427.

Mussoi, B.S., 2021. The impact of music training and working memory on speech recognition in older age. J Speech Lang Hear Res 64 (11), 4524–4534. https://doi.org/10.1044/2021_JSLHR-20-00426.

Nilsson, M., Soli, S.D., Sullivan, J.A., 1994. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J. Acoust. Soc. Am. 95 (2), 1085–1099. https://doi.org/10.1121/1.408469.

Parbery-Clark, A., Skoe, E., Lam, C., Kraus, N., 2009. Musician enhancement for speech-in-noise. Ear Hear. 30 (6), 653–661. https://doi.org/10.1097/AUD.0b013e3181b412e9.

Parbery-Clark, A., Skoe, E., Lam, C., Kraus, N., 2009. Musician enhancement for speech-in-noise. Ear Hear. 30 (6), 653–661. https://doi.org/10.1097/AUD.0b013e3181b412e9.

May 11) Parbery-Clark, A., Strait, D.L., Anderson, S., Hittner, E., Kraus, N., 2011. Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. PLoS One 6 (5), e18082. https://doi.org/10.1371/journal.pone.0018082.

Parbery-Clark, A., Anderson, S., Hittner, E., Kraus, N., 2012. Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. Front. Aging Neurosci. 4, 30. https://doi.org/10.3389/fnagi.2012.00030.

Parbery-Clark, A., Strait, D.L., Hittner, E., Kraus, N., 2013. Musical training enhances neural processing of binaural sounds. J Neurosci 33 (42), 16741–16747. https://doi.org/10.1523/JNEUROSCI.5700-12.2013.

Patel, A.D., 2011. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. Front. Psychol. 2, 142. https://doi.org/10.3389/fpsyg.2011.00142.

Patel, A.D., 2012. The OPERA hypothesis: assumptions and clarifications. Ann. N. Y. Acad. Sci. 1252, 124–128. https://doi.org/10.1111/j.1749-6632.2011.06426.x.

Patel, A.D., 2014. Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. Hear. Res. 308, 98–108. https://doi.org/10.1016/j.heares.2013.08.011.

Perron, M., Theaud, G., Descoteaux, M., Tremblay, P., 2021. The frontotemporal organization of the arcuate fasciculus and its relationship with speech perception in young and older amateur singers and non-singers. Hum. Brain Mapp. 42 (10), 3058–3076. https://doi.org/10.1002/hbm.25416.

Jan 11) Perron, M., Vaillancourt, J., Tremblay, P., 2022. Amateur singing benefits speech perception in aging under certain conditions of practice: behavioural and neurobiological mechanisms. Brain Struct. Funct.. https://doi.org/10.1007/s00429-021-02433-2.

Pouryaghoub, G., Mehrdad, R., Pourhosein, S., 2017. Jan 24). Noise-Induced hearing loss among professional musicians. J. Occup. Health 59 (1), 33–37. https://doi.org/10.1539/joh.16-0217-OA.

Nov 1) Presacco, A., Simon, J.Z., Anderson, S., 2016. Effect of informational content of noise on speech representation in the aging midbrain and cortex. J. Neurophysiol. 116 (5), 2356–2367. https://doi.org/10.1152/jn.00373.2016.

Jul 22 Puschmann, S., Baillet, S., Zatorre, R.J., 2019. Musicians at the cocktail party: neural substrates of musical training during selective listening in multispeaker situations. Cerebr. Cortex 29 (8), 3253–3265. https://doi.org/10.1093/cercor/bhy193.

Jun 23) Rajan, R., Cainer, K.E., 2008. Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. Neuroscience 154 (2), 784–795. https://doi.org/10.1016/j.neuroscience.2008.03.067.

Ramage-Morin, P.L., Banks, R., Pineault, D., Atrach, M., 2019. Perte auditive non perçue chez les Canadiens de 40 à 79 ans. Rapports sur la santé, Issue.

Ramrattan, H., Gurevich, N., 2020. Prevalence of noise-induced hearing loss in middle and high school band members: a preliminary study. Folia Phoniatrica Logop. 72 (4), 302–308. https://doi.org/10.1159/000501154.

Rennies, J., Best, V., Roverud, E., Kidd Jr., G., 2019. Energetic and informational components of speech-on-speech masking in binaural speech intelligibility and perceived listening effort. Trends Hear 23, 2331216519854597. https://doi.org/10.1177/2331216519854597.

Rosen, S., Souza, P., Ekelund, C., Majeed, A.A., 2013. Listening to speech in a background of other talkers: effects of talker number and noise vocoding. J. Acoust. Soc. Am. 133 (4), 2431–2443. https://doi.org/10.1121/1.4794379.

Ruggles, D.R., Freyman, R.L., Oxenham, A.J., 2014. Influence of musical training on understanding voiced and whispered speech in noise. PLoS One 9 (1), e86980. https://doi.org/10.1371/journal.pone.0086980.

Santini, Z.I., Jose, P.E., York Cornwell, E., Koyanagi, A., Nielsen, L., Hinrichsen, C., Meilstrup, C., Madsen, K.R., Koushede, V., 2020. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. Lancet Public Health 5 (1), e62–e70. https://doi.org/10.1016/s2468-2667(19)30230-0.

Schellenberg, E.G., 2005. Music and cognitive abilities. Curr. Dir. Psychol. Sci. 14 (6), 317–320. https://doi.org/10.1111/j.0963-7214.2005.00389.x.

Schellenberg, E.G., 2011. Examining the association between music lessons and intelligence. Br. J. Psychol. 102 (3), 283–302. https://doi.org/10.1111/j.2044-8295.2010.02000.x.

Schellenberg, E.G., 2015. Music training and speech perception: a gene-environment interaction. Ann. N. Y. Acad. Sci. 1337, 170–177. https://doi.org/10.1111/nyas.12627.

Schellenberg, E.G., Hallam, S., 2005. Music listening and cognitive abilities in 10- and 11-year-olds: the blur effect. Ann. N. Y. Acad. Sci. 1060, 202–209. https://doi.org/10.1196/annals.1360.013.

Schellenberg, E.G., Peretz, I., 2008. Music, language and cognition: unresolved issues. Trends Cognit. Sci. 12 (2), 45–46. https://doi.org/10.1016/j.tics.2007.11.005.

Schon, D., Magne, C., Besson, M., 2004. The music of speech: music training facilitates pitch processing in both music and language. Psychophysiology 41 (3), 341–349. https://doi.org/10.1111/1469-8986.00172.x.

Shinn-Cunningham, B., 2013. Understanding Informational Masking from a Neural Perspective.

Skoe, E., Camera, S., Tufts, J., 2019. Noise exposure may diminish the musician advantage for perceiving speech in noise. Ear Hear 40 (4), 782–793. https://doi.org/10.1097/AUD.0000000000000665.

Slater, J., Kraus, N., 2016. The role of rhythm in perceiving speech in noise: a comparison of percussionists, vocalists and non-musicians. Cogn Process 17 (1), 79–87. https://doi.org/10.1007/s10339-015-0740-7.

Slater, J., Kraus, N., Carr, K.W., Tierney, A., Azem, A., Ashley, R., 2018. Speech-in-noise perception is linked to rhythm production skills in adult percussionists and non-musicians. Lang Cogn Neurosci 33 (6), 710–717. https://doi.org/10.1080/23273798.2017.1411960.

Smeds, K., Wolters, F., Rung, M., 2015. Estimation of signal-to-noise ratios in realistic sound scenarios. J. Am. Acad. Audiol. 26 (2), 183–196. https://doi.org/10.3766/jaaa.26.2.7.

Snell, K.B., Frisina, D.R., 2000. Relationships among age-related differences in gap detection and word recognition [Comparative Study Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. J. Acoust. Soc. Am. 107 (3), 1615–1626. http://www.ncbi.nlm.nih.gov/pubmed/10738815.

Soncini, F., Costa, M.J., 2006. [The effect of musical practice on speech recognition in quiet and noisy situations]. Pro Fono 18 (2), 161–170. https://doi.org/10.1590/s0104-56872006000200005 (Efeito da pratica musical no reconhecimento da fala no silencio e no ruido.).

Souza, P.E., Turner, C.W., 1994. Masking of speech in young and elderly listeners with hearing loss. J. Speech Hear. Res. 37 (3), 655–661. https://doi.org/10.1044/jshr.3703.655.

Stern, Y., 2002. What is cognitive reserve? Theory and research application of the reserve concept. J. Int. Neuropsychol. Soc. 8 (3), 448–460.

Stern, Y., 2009. Cognitive reserve. Neuropsychologia 47 (10), 2015–2028. https://doi.org/10.1016/j.neuropsychologia.2009.03.004.

Stern, Y., Arenaza-Urquijo, E.M., Bartres-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W.S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Valenzuela, M., Vemuri, P., Vuoksimaa, E., Reserve, R., Protective Factors, P. I. A. E. D., & Conceptual Frameworks, W, 2020. Whitepaper: defining and investigating cognitive reserve, brain reserve, and brain maintenance. Alzheimers Dement 16 (9), 1305–1311. https://doi.org/10.1016/j.jalz.2018.07.219.

Strait, D.L., Kraus, N., 2011. Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. Front. Psychol. 2, 113. https://doi.org/10.3389/fpsyg.2011.00113.

Apr 16) Sundström, A., Adolfsson, A.N., Nordin, M., Adolfsson, R., 2020. Loneliness increases the risk of all-cause dementia and alzheimer's disease. J. Gerontol. B Psychol. Sci. Soc. Sci. 75 (5), 919–926. https://doi.org/10.1093/geronb/gbz139.

Swaminathan, J., Mason, C.R., Streeter, T.M., Best, V., Kidd Jr., G., Patel, A.D., 2015. Musical training, individual differences and the cocktail party problem. Sci Rep 5, 11628. https://doi.org/10.1038/srep11628.

Swaminathan, S., Schellenberg, E.G., 2020. Musical ability, music training, and language ability in childhood. J. Exp. Psychol. Learn. Mem. Cogn. 46 (12), 2340–2348. https://doi.org/10.1037/xlm0000798.

Taitelbaum-Swead, R., Fostick, L., 2016. The effect of age and type of noise on speech perception under conditions of changing context and noise levels. Folia Phoniatrica Logop. 68 (1), 16–21. https://doi.org/10.1159/000444749.

Taylor, H.O., Taylor, R.J., Nguyen, A.W., Chatters, L., 2018. Social isolation, depression, and psychological distress among older adults. J. Aging Health 30 (2), 229–246. https://doi.org/10.1177/0898264316673511.

Teie, P.U., 1998. Noise-induced hearing loss and symphony orchestra musicians: risk factors, effects, and management. Md. Med. J. 47 (1), 13–18.

Tun, P.A., 1998. Fast noisy speech: age differences in processing rapid speech with background noise. Psychol. Aging 13 (3), 424–434. https://doi.org/10.1037//0882-7974.13.3.424.

Tun, P.A., Wingfield, A., 1999. One voice too many: adult age differences in language processing with different types of distracting sounds. J. Gerontol. B Psychol. Sci. Soc. Sci. 54 (5), P317–P327.

Vanden Bosch der Nederlanden, Zaragoza, C., Rubio-Garcia, A., Clarkson, E., Snyder, J. S., 2020. Change detection in complex auditory scenes is predicted by auditory memory, pitch perception, and years of musical training. Psychol Res 84 (3), 585–601. https://doi.org/10.1007/s00426-018-1072-x.

Varnet, L., Wang, T., Peter, C., Meunier, F., Hoen, M., 2015. How musical expertise shapes speech perception: evidence from auditory classification images. Sci Rep 5, 14489. https://doi.org/10.1038/srep14489.

Vermeire, K., Knoop, A., Boel, C., Auwers, S., Schenus, L., Talaveron-Rodriguez, M., De Boom, C., De Sloovere, M., 2016. Speech recognition in noise by younger and older adults: effects of age, hearing loss, and temporal resolution. Ann. Otol. Rhinol. Laryngol. 125 (4), 297–302. https://doi.org/10.1177/0003489415611424.

(2010, 2010-08-05) Viechtbauer, W., 2010. Conducting Meta-Analyses in R with the metafor Package 36 (3), 48. https://doi.org/10.18637/jss.v036.i03.

Wan, C.Y., Schlaug, G., 2010. Music making as a tool for promoting brain plasticity across the life span. Neuroscientist 16 (5), 566–577. https://doi.org/10.1177/1073858410377805.

Wang, X., Xu, L., 2021. Speech perception in noise: masking and unmasking. J. Otolaryngol. 16 (2), 109–119. https://doi.org/10.1016/j.joto.2020.12.001.

Wilson, R.H., 2003. Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. J. Am. Acad. Audiol. 14 (9), 453–470.

Wilson, R.H., McArdle, R.A., Smith, S.L., 2007. An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. J. Speech Lang. Hear. Res. 50 (4), 844–856. https://doi.org/10.1044/1092-4388(2007/059).

1988/03/01 Working Group on Speech, U., & Aging, 1988. Speech understanding and aging. J. Acoust. Soc. Am. 83 (3), 859–895. https://doi.org/10.1121/1.395965.

Wu, Y.H., Stangl, E., Chipara, O., Hasan, S.S., Welhaven, A., Oleson, J., 2018. Characteristics of real-world signal to noise ratios and speech listening situations of older adults with mild to moderate hearing loss. Ear Hear. 39 (2), 293–304. https://doi.org/10.1097/aud.0000000000000486.

Yeend, I., Beach, E.F., Sharma, M., Dillon, H., 2017. The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. Hear. Res. 353, 224–236. https://doi.org/10.1016/j.heares.2017.07.006.

Yost, W.A., 2017. Spatial release from masking based on binaural processing for up to six maskers. J. Acoust. Soc. Am. 141 (3), 2093. https://doi.org/10.1121/1.4978614.

2007/07/01 Zatorre, R.J., Chen, J.L., Penhune, V.B., 2007. When the brain plays music: auditory—motor interactions in music perception and production. Nat. Rev. Neurosci. 8 (7), 547–558. https://doi.org/10.1038/nrn2152.

Zendel, B.R., Alain, C., 2009. Concurrent sound segregation is enhanced in musicians. J. Cognit. Neurosci. 21 (8), 1488–1498. https://doi.org/10.1162/jocn.2009.21140.

Zendel, B.R., Alain, C., 2012. Musicians experience less age-related decline in central auditory processing. Psychol. Aging 27 (2), 410–417. https://doi.org/10.1037/a0024816.

Zendel, B.R., Alain, C., 2013. The influence of lifelong musicianship on neurophysiological measures of concurrent sound segregation. J. Cognit. Neurosci. 25 (4), 503–516. https://doi.org/10.1162/jocn_a_00329.

Zendel, B.R., Tremblay, C.D., Belleville, S., Peretz, I., 2015. The impact of musicianship on the cortical mechanisms related to separating speech from background noise. J. Cognit. Neurosci. 27 (5), 1044–1059. https://doi.org/10.1162/jocn_a_00758.

Zhang, L., Fu, X., Luo, D., Xing, L., Du, Y., 2021. Musical experience offsets age-related decline in understanding speech-in-noise: type of training does not matter, working memory Is the key. Ear Hear 42 (2), 258–270. https://doi.org/10.1097/AUD.0000000000000921.

Zhang, F., Roland, C., Rasul, D., Cahn, S., Liang, C., Valencia, G., 2019. Comparing musicians and non-musicians in signal-in-noise perception. Int J Audiol 58 (11), 717–723. https://doi.org/10.1080/14992027.2019.1623424.

Zuk, J., Benjamin, C., Kenyon, A., Gaab, N., 2014. Behavioral and neural correlates of executive functioning in musicians and non-musicians. PLoS One 9 (6), e99868. https://doi.org/10.1371/journal.pone.0099868.