

BMJ Open Hearing, speech reception, vocabulary and language: population epidemiology and concordance in Australian children aged 11 to 12 years and their parents

Julia Smith,^{1,2} Jing Wang,^{1,2} Anneke C Grobler,^{1,2} Katherine Lange,^{1,2} Susan A Clifford,^{1,2} Melissa Wake^{1,2,3}

To cite: Smith J, Wang J, Grobler AC, *et al.* Hearing, speech reception, vocabulary and language: population epidemiology and concordance in Australian children aged 11 to 12 years and their parents. *BMJ Open* 2019;9:85–94. doi:10.1136/bmjopen-2018-023196

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-023196>).

Received 26 March 2018
Revised 18 October 2018
Accepted 4 April 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Murdoch Children's Research Institute, Parkville, Victoria, Australia

²Department of Paediatrics, The University of Melbourne, Parkville, Victoria, Australia

³Department of Paediatrics and The Liggins Institute, The University of Auckland, Auckland, New Zealand

Correspondence to

Professor Melissa Wake;
melissa.wake@mcri.edu.au

ABSTRACT

Objectives To describe the epidemiology and parent-child concordance of hearing, speech reception, vocabulary and language in Australian parent-child dyads at child age 11 to 12 years.

Design Population-based cross-sectional study (Child Health CheckPoint) nested within the Longitudinal Study of Australian Children.

Setting Assessment centres in seven Australian cities and eight regional towns or home visits around Australia, February 2015 to March 2016.

Participants Of all participating CheckPoint families (n=1874), 1516 children (50% female) and 1520 parents (87% mothers, mean age 43.8 years) undertook at least one of four measurements of hearing and language.

Outcome measures Hearing threshold (better ear mean of 1, 2 and 4 kHz) from pure-tone audiometry, speech reception threshold, receptive vocabulary, expressive and receptive languages using a sentence repetition task. Parent-child concordance was examined using Pearson's correlation coefficients and adjusted linear regression models. Survey weights and methods accounted for Longitudinal Study of Australian Children's complex sampling and stratification.

Results Children had a similar speech reception threshold to parents (children mean –14.3, SD 2.4; parents –14.9, SD 3.2 dB) but better hearing acuity (children 8.3, SD 6.3; parents 13.4, SD 7.0 decibels hearing level). Standardised sentence repetition scores were similar (children 9.8, SD 2.9; parents 9.1, SD 3.3) but, as expected, parents had superior receptive vocabularies. Parent-child correlations were higher for the cognitively-based language measures (vocabulary 0.31, 95% CI 0.26 to 0.36; sentence repetition 0.29, 95% CI 0.24 to 0.34) than the auditory measures (hearing 0.18, 95% CI 0.13 to 0.23; speech reception threshold 0.18, 95% CI 0.13 to 0.22). Mother-child and father-child concordances were similar for all measures.

Conclusions We provide population reference values for multiple measures spanning auditory and verbal communication systems in children and mid-life adults. Concordance values aligned with previous twin studies and offspring studies in adults, in keeping with polygenic heritability that is modest for audition but around 60% for language by late childhood.

Strengths and limitations of this study

- This is the first study to our knowledge to measure multiple elements of the auditory and verbal communication systems across two generations at the population level.
- Hearing, speech reception threshold, vocabulary and sentence repetition were all measured with objective standardised tools.
- This was also the first time to our knowledge that a sentence repetition measure was administered via an iPad application, with high inter- and intra-rater reliability.
- Our parent sample comprised mostly mothers, resulting in less precise descriptive and concordance estimates for fathers.
- Results may not generalise to very disadvantaged families and adults who are not parents.

INTRODUCTION

Effective oral communication usually requires that we hear what is said, process it accurately against background noise and have adequate grammar and vocabulary to formulate our responses. Dysfunction anywhere along this auditory-verbal communication system can affect physical health and psychosocial outcomes,^{1–4} culminating in the profound global burden related to hearing loss and declining cognitive language abilities in the elderly. Brief, objective, adaptive measurement tools now allow researchers to measure multiple elements of hearing and language across the life course within large population studies, including the relatively understudied stages of adolescence and mid-life. Objectively describing the epidemiology of these abilities is a necessary first step towards a deeper exploration of their inter-relationships.

Different components of this system follow their own developmental trajectories.^{5–7} Hearing acuity (for which pure-tone

audiometry is the gold-standard assessment⁸) peaks in infancy then declines throughout life. Hearing is well characterised in young children and the elderly^{9 10}; yet relatively fewer population data are available during adolescence¹¹ and mid-life, despite the high prevalence of noise exposure and risk factors for non-communicable diseases in these age groups. Speech perception in noise, which relies on central auditory processing functions coupled with hearing acuity, may be more representative than pure-tone audiometry of listening situations encountered in everyday life. Speech perception (measured as speech reception threshold) develops rapidly to adolescence and then more slowly until it plateaus in mid-adulthood, before declining from approximately 60 years of age.¹² However, this life course information comes from population studies with small sample sizes and heterogeneous assessment tools,¹³ limiting generalisability.

Language skills are also important determinants of daily functioning and health and are closely linked to academical and employment outcomes.¹⁴ While language acquisition largely occurs in early childhood, linguistic sophistication develops through mid-childhood and adolescence, and vocabulary grows throughout much of the life course.⁷ Measuring its epidemiology is challenging, with the very complexity of language resulting in a diversity of multi-construct measures that themselves evolve over time. Lack of short gold-standard measures in turn means that language often does not appear in broad-based population health studies.^{15 16} Existing population studies are skewed towards early childhood, which is unsurprising given an emphasis on prevention and early treatment to improve learning and other outcomes.¹⁷ Nonetheless, language difficulties represent a significant burden beyond early childhood, with impairment reported to affect 12% to 13% of 5 to 18 year olds.¹⁸ Furthermore, language and communication deficits often persist into adulthood,¹⁴ affecting up to 10% of adults.¹⁹

Genetic factors contribute to hearing,^{20–22} central auditory processing²³ and language skills.²⁴ The US Framingham study²⁵ estimated correlations in hearing thresholds for parents and offspring at comparable midlife ages to be 0.10 for middle-frequency (0.5, 1, 2 kHz), 0.19 for high-frequency average (4, 6, 8 kHz) and 0.13 for low-frequency average (0.25, 0.5, 1.0 kHz); estimated heritability ranged from 0.26 to 0.32. Language heritability appears to be even stronger and increases throughout childhood, with a Dutch twin study reporting a heritability of 0.48 for language at age 9 rising to 0.84 by age 18 years.²⁶ However, to our knowledge no studies have directly measured either concordance or heritability of *both* hearing and language simultaneously. Furthermore, most hearing concordance studies have been conducted when offspring were themselves mid-life adults, whereas language studies have mainly occurred when the offspring were children. Thus, different study methods, constructs, measures and offspring ages may all preclude a population-level understanding of differences

in cross-generational concordance across the hearing and language continuum.

In summary, the descriptive epidemiology and inter-generational concordance of diverse elements of the auditory and verbal communication systems are poorly documented in population samples of older children and adults. The Child Health CheckPoint,²⁷ a cross-sectional biophysical wave nested within the Longitudinal Study of Australian Children (LSAC), provided an opportunity to redress this within a broadly-focused population-based health assessment of Australian parent-child dyads at child age 11 to 12 years. In this paper, we aimed to (1) describe the distribution of standardised measures of hearing acuity, speech reception threshold, receptive vocabulary and a marker of receptive and expressive language skills in both age groups in an Australian population, and (2) investigate parent-child concordance in these measures. In addition, we examined inter- and intra-rater reliability for the language measure as this was the first time to our knowledge that sentence repetition was administered via an iPad application.

METHODS

Study design and participants

Details of the initial study design and recruitment are outlined elsewhere.^{28 29} Briefly, LSAC recruited a nationally representative cohort of 5107 infants using a two-stage clustered design, and followed them up in biennial 'waves' of data collection since. The initial recruitment rate in 2004 was 57.2%, of whom 73.7% (n=3764) participated in LSAC wave 6 in 2014. During the wave 6 visit, families were invited to consent to their contact details being shared with the Child Health CheckPoint team. In 2015, families that consented at wave 6 to CheckPoint contact (n=3513) were sent an information pack via post and received an information and recruitment phone call. The Child Health CheckPoint was a one-off detailed cross-sectional biophysical assessment for the study child and one of their parents.²⁷ The CheckPoint was nested between LSAC waves 6 and 7, taking place between February 2015 and March 2016 at child age 11 to 12 years (see detailed description of CheckPoint methods³⁰). A total of 1874 families participated.

Consent

A parent or guardian provided written informed consent for their own and their child's participation.

Patient and public involvement

Because LSAC is a population-based longitudinal study, no patient groups were involved in its design or conduct. To our knowledge, the public was not involved in the study design, recruitment or conduct of the LSAC study or its CheckPoint module. Parents received a summary health report for their child and themselves at or soon after the assessment visit. They consented to take part

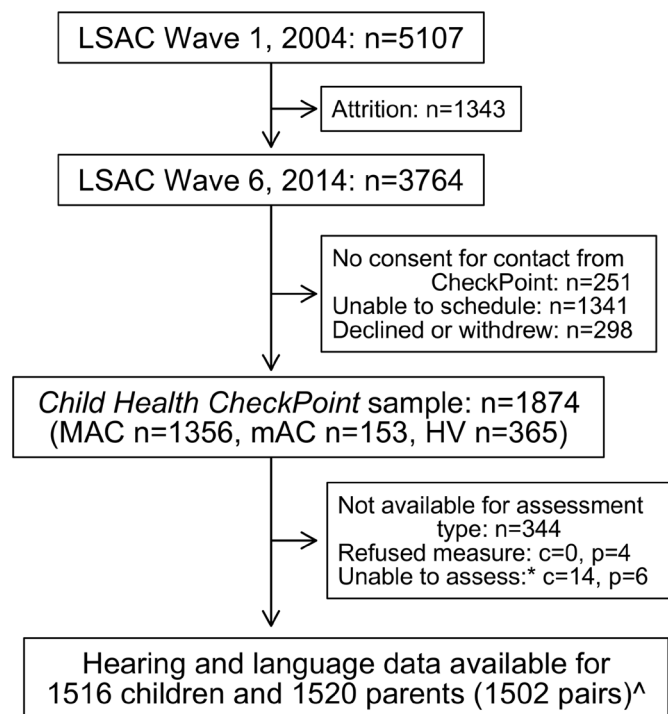


Figure 1 Participant flow chart. *Unable to assess due to equipment failure, poor quality data or time constraints. ^Data from thirteen non-biological child-parent pairs excluded from concordance analyses. n, number of families; c, number of children; p, number of attending adults; MAC, main assessment centre; mAC, mini assessment centre; HV, home visit assessment; LSAC, Longitudinal Study of Australian Children.

knowing that they would not otherwise receive individual results about themselves or their child.

Procedure

All four auditory and language measures were collected at a specialised 3.5 hour (six capital cities and one large regional city) or 2.75 hour (eight smaller regional centres) CheckPoint assessment centre visit. At the assessment centre, participants completed numerous measurements of multiple body systems in a standard sequence that differed slightly for children and parents. Hearing and language testing occurred early in the visit because these tasks required concentration on the part of the participant, unlike most of the other physical measurements. Each child and parent separately visited the 15 min 'Listen Up' station (30 or 45 min into the visit, respectively), where they underwent pure-tone audiometry, tympanometry (not reported here), speech reception threshold and sentence repetition tasks. Shortly afterwards, they visited the 15 min 'Bone Zone' station together, where each completed the receptive vocabulary task while the other underwent bone imaging. Families who could not attend a centre received a 1.5 hour home visit (figure 1), which included only the vocabulary measure because of time constraints and practical constraints. Participants were included in the current analyses if they had data for at least one auditory or language measure (figure 1). Dyads

were excluded from concordance analyses if the attending caregiver was not a biological parent to the study child.

Audiometry

Trained examiners conducted air-conduction pure-tone audiometry using an Oscilla USB-330 (V.3.3.4) computer-based audiometer with Oscilla headphone and a standardised modified Hughson-Westlake audiometric technique. Testing of the first frequency began at 30 decibels hearing level (dB HL); if within normal limits, testing for each frequency thereafter began at 20 dB HL. If participants wore hearing aids and/or cochlear implant speech processors, they were asked to remove them where possible, and the test started at the 1 kHz frequency at 60 dB HL. Those who could not remove their hearing devices were not tested. Responses were recorded in the database. For the first 143 dyads, only three frequencies were tested (1, 2, 4 kHz) for each ear across an intensity range of -10 to 120 dB HL. As CheckPoint systems became faster and additional funding was sourced, testing at 8 kHz (1342 dyads), and soundproof booths (813 children, 812 parents) were successively added. If participants' hearing thresholds at two or more frequencies in at least one ear were >20 dB HL, parents' written feedback stated that hearing was outside the usual range and they should consider a clinical audiology assessment. In our analyses below, hearing thresholds are defined as the mean hearing threshold (dB HL) at 1, 2 and 4 kHz in the better ear, known as the high Fletcher index, which maps closely to the range of speech sound frequencies so is functionally relevant to oral communication.³¹

Speech reception threshold

Trained examiners determined the speech reception threshold of the participants in a simulated three-dimensional listening environment, using Phonak's computerised adaptive Listening in Spatialized Noise-Sentences test (LiSN-S) (V.1.104) with Sennheiser HD215 circumaural headphones. Participants listened to 30 simple target sentences (average of five words), which are localised to 0 degrees azimuth and with varying loudness. The high-cue LiSN-S condition was used (different voices±90° azimuth condition). Competing speech using voices different to that of the target speech were presented binaurally, with loudness held constant at 55 dB sound pressure levels. After each sentence the examiner scored the number of words repeated correctly, resulting in an automatic adjustment to the sound-to-noise ratio (ratio of target speech to competing noise measured in dB) for the next target sentence to rapidly achieve a very precise estimate. The speech reception threshold was determined as the lowest intensity sound (in dB) of target words that can be understood against competing speech at 55 dB for at least 50% of the time. At between 22 and 30 sentences, the speech reception threshold was automatically calculated. If there was insufficient time to finish the 30 sentences, the speech reception threshold was estimated from the first 22 sentences (by which time the speech reception

threshold closely approximates that from the full test). Further details regarding administration are reported by Cameron.³²

Receptive vocabulary

The computerised adaptive National Institutes of Health Toolbox Picture Vocabulary Test (NPVT)³³ was used to assess receptive vocabulary. Wearing headphones, participants were presented with a series of auditory recordings of single words each with four photographic images on an iPad screen, from which they were required to select the image that most closely represented the meaning of the word. NPVT included two practice items (with feedback from the examiner) followed by 25 test items selected by the computer from its bank of words of a wide range of difficulty. The starting word was determined according to the participant's age and educational level, and subsequent difficulty adjusted according to the performance on all preceding words presented. More details regarding administration are reported by Gershon *et al.*³⁴ A 'theta score' was calculated for each participant, which represents an algorithm-derived raw-score that typically ranges between -4 and 4 for 3 to 17 year olds³⁵ with a full range between -6.15 and 6.76.³⁶

Sentence repetition

Participants completed the Recalling Sentences task from the Clinical Evaluation of Language Fundamentals fourth edition (CELF-4, Australian version),³⁷ administered by an iPad. Recalling Sentences task is a subscale of the CELF-4, which is a normed measure that is widely used in educational and research settings to assess expressive and receptive language. We selected this following analysis within our large population-based language oriented cohort, the Early Language in Victoria Study. Of all the CELF subscales, Recalling Sentences had the highest area under the curve (0.96) in receiver operating characteristics curves predicting the total CELF Core Language score (data not shown). It is also the subscale that best corresponds to both receptive and expressive skills.

Wearing a headset with microphone plugged into the iPad, participants were presented with an audio recording of each sentence via audio file, and asked to repeat it back verbatim, without any visual cues. After two trial sentences, sentences of progressive length and difficulty were presented. The trained examiner scored each sentence in real time at the assessment, with an audio recording of the participant's repeated sentences also recorded on the iPad for later rescoring if necessary. As the participant repeated each sentence, the research assistant scored it on the iPad as 'correct', 'intermediate/uncertain' or 'incorrect' (four or more errors). The test ended at sentence 32 or after three consecutive 'incorrect' scores. Total raw score could range from 0 to 96. This was converted to an age-related scaled score (ranging from 1 to 18) based on normative data with a mean of 10 and SD of 3, using the oldest age-bracket norms (17 years

0 month to 21 years 11 months) for the parent scores as the CELF is not normed beyond 21 years.³⁷

At a later date, one of four trained scorers listened to the recordings of the sentences with 'intermediate/uncertain' scores. These sentences were scored as correct (no errors, score of 3), one error (score of 2), two to three errors (score of 1) or incorrect (four or more errors, score of 0). To check the accuracy of the on-the-spot scoring, when a scorer listened to 'intermediate/uncertain' sentences, they also rescored all other recorded sentences for that participant (including those scored as correct or incorrect) for 1377 participants. Full rescoring of all sentences, regardless of the indication of sentence correctness given on the spot, resulted in mean differences in raw total scores of <0.1% (ie, 71.28 vs 71.24 for children). Therefore, only 'intermediate/uncertain' sentences were rescored for the remaining participants.

Other sample characteristics

Measurement of baseline characteristics are described in the cohort profile in this Special Issue.³⁰ Additional characteristics pertinent to this study include the presence of pre-existing hearing conditions or hearing aid use (reported by the parent and also self-reported by the child). Sex and date of birth were exported from Medicare Australia's database at the time of LSAC enrolment (for the child) or self-reported (parent). Age was calculated to the nearest week using date of birth and date of assessment. Neighbourhood socioeconomic position was measured using the Socio-Economic Indexes for Areas Index of Relative Socioeconomic Disadvantage (disadvantage index), derived from the 2011 Australian Census. This numerically summarises the social and economical conditions of Australian neighbourhoods by geographical area (postcode of family domicile), and has a national mean of 1000 and SD of 100, where higher values represent less disadvantage.³⁸

Statistical analysis

All analyses were performed using Stata V.14.2 (StataCorp, College Station, Texas). Sample demographical characteristics were summarised using weighted means and SD or number and weighted per cent for children and adults separately and by sex. Parent and child hearing and language characteristics were summarised using weighted means, SDs and 95% CI, for the whole sample and by sex. Survey weights were calculated taking into account the selection probability of each child, and were adjusted for non-response, loss to follow-up and benchmarked to population numbers in major (post stratification) categories of the population of children born in 2004. Standard errors were calculated taking into account the complex design and weights.³⁹ More detail on the calculation of weights is provided elsewhere.⁴⁰

Concordance between parents and children was assessed by (1) Pearson's correlation coefficients with 95% CI and (2) linear regression with the child variable as the dependent variable and the parent variable as the independent

Table 1 Sample characteristics; values are weighted mean (SD), except where specified as (%)

Characteristics	All	Male	Female
Child	(n=1516)	(n=758)	(n=758)
Age, years	11.9 (0.4)	11.9 (0.4)	11.9 (0.4)
Disadvantage index	1026 (60)	1026 (60)	1025 (60)
Self-reported hearing impairment (%)	1.3	1.2	1.5
Parent	(n=1520)	(n=200)	(n=1320)
Age, years	43.8 (5.1)	45.9 (6.4)	43.4 (4.8)
Self-reported hearing impairment (%)	3.9	4.0	3.9

variable, adjusted for age, disadvantage index (and sex where appropriate). Correlations and regression analyses were repeated using weighted multi-level survey analyses; as these yielded similar results, unweighted results are displayed.

Inter- and intra-rater reliability

The majority of the Recalling Sentence scoring was undertaken by four scorers. Inter- and intra-rater reliability were assessed for each of the 32-sentence numerical scores. Each of these four scorers rescored recalled sentences for the same 40 randomly selected participants, 10 of which each scorer had previously scored in original assessments.

For inter-rater reliability we used weighted kappa to examine the strength of agreement between scores made by two different scorers. The mean weighted kappa across the 32 numerical scores was 0.89 (range 0.47, $p=0.16$ to 1.0, $p<0.001$). To assess agreement of scores across all four scorers, we used Fleiss' kappa with quadratic weights given the ordinal nature of the sentence numerical scores

(ie, 0 to 3). Mean Fleiss' kappa was 0.89 (range 0.65, $p=0.03$ to 1.0, $p<0.001$). We also examined the per cent of agreement across the 32 numerical scores, showing an average per cent agreement between scorers of 98% (range 98 to 100). For intra-rater reliability, weighted kappa ranged from 0.80 to 0.89, and the per cent of agreement between two scores obtained by the same scorer ranged from 96% to 98%.

RESULTS

Sample characteristics

Of the 1874 families who participated in the CheckPoint, 1516 children and 1520 parents had data for at least one of the hearing and language assessments, including 1502 biological parent-child pairs (figure 1). While there were equal numbers of girls and boys, only 13% of parents were fathers (table 1). The sample's mean socioeconomic status as measured by neighbourhood disadvantage index was on average approximately a quarter of a SD above the national mean, and with a narrower spread (1026, SD 60 vs national mean 1000, SD 100) such that on average families who participated were from more advantaged neighbourhoods. Hearing impairment was self-reported by 3.9% of parents and 1.3% of children.

Population epidemiology of hearing and language measures

Weighted distributions of the four measures are summarised in table 2. Online supplementary table 1 shows extended weighted percentiles (from the 5th to 95th) for reference purposes. As shown in the density plots (figure 2), all measures showed approximately normal distributions although with some skewing (positive skewing, parent and child hearing and

Table 2 Distribution of hearing and language measures in Australian children and parents

Hearing/language measures	All			Males			Females		
	N	Mean (SD)	95% CI	N	Mean (SD)	95% CI	N	Mean (SD)	95% CI
Children									
Hearing threshold, dB HL	1488	8.3 (6.3)	7.9 to 8.6	741	7.7 (5.9)	7.2 to 8.2	747	8.8 (6.9)	8.2 to 9.4
Speech reception threshold, dB	1483	-14.3 (2.4)	-14.4 to -14.1	740	-14.3 (2.3)	-14.4 to -14.1	743	-14.3 (2.5)	-14.5 to -14.1
Receptive vocabulary	1443	1.4 (1.5)	1.3 to 1.5	727	1.5 (1.5)	1.4 to 1.6	716	1.3 (1.5)	1.1 to 1.4
Sentence repetition (raw)	1441	69.7 (12.9)	68.8 to 70.6	717	69.0 (13.5)	67.7 to 70.4	724	70.4 (12.5)	69.2 to 71.6
Sentence repetition (scaled)	1441	9.8 (2.9)	9.6 to 9.9	717	9.6 (3.0)	9.3 to 9.9	724	9.9 (2.8)	9.6 to 10.1
Parents									
Hearing threshold, dB HL	1493	13.4 (7.0)	13.0 to 13.8	198	14.4 (8.4)	13.2 to 15.7	1295	13.3 (6.8)	12.8 to 13.7
Speech reception threshold, dB	1482	-14.9 (3.2)	-15.1 to -14.7	195	-14.4 (2.8)	-14.8 to -13.9	1287	-15.0 (3.2)	-15.2 to -14.8
Receptive vocabulary	1457	5.5 (2.1)	5.4 to 5.7	197	5.8 (2.4)	5.3 to 6.2	1260	5.5 (2.1)	5.3 to 5.6
Sentence repetition (raw)	1446	80.0 (14.5)	79.0 to 81.0	188	79.1 (14.6)	76.4 to 81.7	1258	80.1 (14.5)	79.0 to 81.2
Sentence repetition (scaled)	1446	9.1 (3.3)	8.8 to 9.3	188	8.8 (3.4)	8.1 to 9.4	1258	9.1 (3.3)	8.9 to 9.3

All values are weighted. n=number of participants in cohort with this measure (denominator).
dB HL, decibels hearing level.

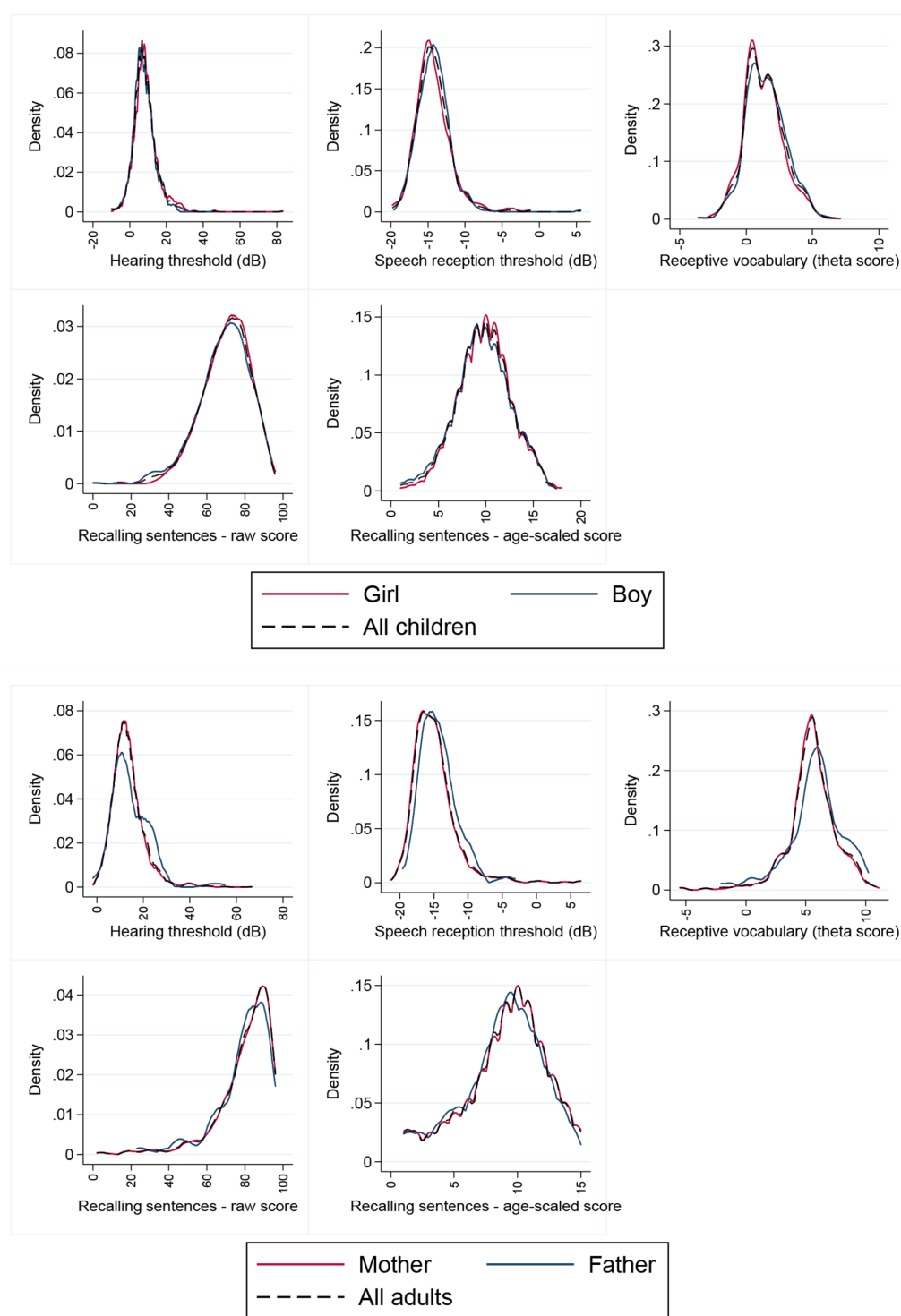


Figure 2 Density plots for hearing and language measures for children and parents. Males (blue), females (red) and both sexes (thin dotted black line) plotted on the same graph for each outcome. dB, decibels .

speech reception thresholds; negative skewing, parent receptive vocabulary).

The children's speech reception threshold approached that of the parents (mean -14.3 dB, SD 2.4 vs -14.9 dB, SD 3.2) but children had better hearing acuity (mean 8.3 dB HL, SD 6.3 vs 13.4 dB HL, SD 7.0). As expected, parents

had higher receptive vocabulary scores. Similarly, parents had higher raw sentence repetition scores (mean 80.0 of a possible 96) than children (mean 69.7), although their standardised sentence repetition scores were on average marginally lower (parents 9.1, SD 3.3; children 9.8, SD 2.9). While we did not conduct statistical comparisons

by sex, we note that there were no striking differences between males and females in either children or parent groups.

Parent-child concordance

Table 3 shows unweighted correlation (CC) and estimated regression (RC) coefficients for concordance of each measure between biological parents and children. Correlation coefficients were smaller for hearing (0.18, 95% CI 0.13 to 0.23 for acuity; 0.18, 95% CI 0.13 to 0.22 for speech reception threshold) than for language measures (0.29, 95% CI 0.24 to 0.34 for receptive vocabulary; 0.34, 95% CI 0.29 to 0.38 for speech reception threshold). Concordance was slightly higher for father-child than mother-child pairs across all measures, noting, however, the much smaller sample of fathers and correspondingly wider and overlapping confidence intervals. All values attenuated somewhat in the linear regression models adjusted for age, sex and disadvantage index (table 3). Estimated regression coefficients for parent-child concordance ranged from 0.14 to 0.29, with similar patterns at the mother-child and father-child level.

DISCUSSION

Principal findings

We showed that it is feasible to include multiple objective hearing and language measures for children and parents in a broadly-focused population-based field study despite limited space for measures not overtly tapping into health. We provide population reference values spanning the auditory-verbal communication system for 11 to 12-year-old children and mid-life adults. Concordance values were consistent with previous twin and adult-offspring studies, in keeping with polygenic heritability that is modest for audition but around 60% for language by late childhood.

Strengths and weaknesses in relation to other studies

Given the large sample size and national focus, our values are robust (particularly given application of survey weights) and concordance estimates precise. We used standardised objective measures and computerised delivery, with each parent-dyad assessed using the same protocols and equipment, on the same day and in the same conditions. Our novel sentence repetition iPad App dramatically reduced the completion time with high inter- and intra-observer reliability.

Most parents were mothers, so fathers' estimates are less precise. Inclusion of more disadvantaged families might have reduced mean language scores somewhat, although applying study weights should have partly mitigated this; it is difficult to predict what effect this might have on concordance values. Some participants were tested before we acquired soundproof booths, which could have yielded slightly higher (worse) hearing thresholds, although we saw no evidence of this.⁴¹ Time constraints required omitting the intra-octave 3 and 6 kHz (whose

Table 3 Parent-child concordance, as correlations and regression adjusted for covariates

	Parent-child				Mother-child				Father-child			
	N	CC	95% CI		N	CC	95% CI		N	CC	95% CI	
Pearson's correlation												
Hearing threshold	1467	0.18	0.13 to 0.23		1271	0.17	0.12 to 0.22		196	0.23	0.09 to 0.35	
Speech reception threshold	1453	0.18	0.13 to 0.22		1260	0.16	0.11 to 0.22		193	0.23	0.10 to 0.36	
Receptive vocabulary	1389	0.29	0.24 to 0.34		1201	0.29	0.24 to 0.34		188	0.28	0.14 to 0.40	
Sentence repetition, raw score	1402	0.31	0.26 to 0.36		1217	0.30	0.25 to 0.35		185	0.38	0.25 to 0.50	
Sentence repetition, scaled score	1402	0.34	0.29 to 0.38		1217	0.33	0.28 to 0.38		185	0.42	0.30 to 0.54	
Adjusted linear regression*												
Hearing threshold	1463	0.16	<0.001		1269	0.16	<0.001		194	0.14	0.009	
Speech reception threshold	1449	0.14	<0.001		1258	0.12	<0.001		191	0.20	0.002	
Receptive vocabulary	1385	0.21	<0.001		1199	0.21	<0.001		186	0.16	0.003	
Sentence repetition, raw score	1398	0.28	<0.001		1215	0.27	<0.001		183	0.32	<0.001	
Sentence repetition, scaled score	1398	0.29	<0.001		1215	0.28	<0.001		183	0.35	<0.001	

Non-biological parent-child pairs were excluded from these analyses (n=13).

*Covariates in adjusted linear regression models include parent and child age, disadvantage index and parent and child sex in models including both sexes. CC, estimated Pearson's correlation coefficient; N, number of biological child-parent pairs with this measure; RC, estimated regression coefficient.

thresholds are typically intermediate to those of the two bounding octave frequencies⁴²) and 0.5 kHz (both the least relevant to spoken speech and the most affected by residual background noise) frequencies. Bone-conduction audiometry would have more accurately classified sensorineural and conductive losses, but was infeasible and would have reduced comparability with other population studies.^{11 43} As the CELF instrument is not validated beyond age 21 years, its norms may not apply well to mid-life adults (see below). Other hearing and language constructs could have been chosen; while essentially pragmatic, our measures provide a robust snapshot of the auditory-verbal spectrum and measures despite their brevity.

Interpretation and meaning in relation to other studies

Our mean hearing thresholds in children (8.3 dB HL, SD 6.3) and adults (13.4 dB HL, SD 7.0) were similar to other population studies of similarly-aged participants in the Netherlands, Australia, the USA and Germany.^{9 44–46} Better hearing thresholds in children also resembled previous findings,^{11 47} suggesting that hearing loss is a progressive condition with early life origins. Mean speech reception threshold was comparable to normative LiSN-S data (−16.0 dB for child and adult groups)¹² and Besser *et al*'s study using the LiSN-S (−17.8 dB in 18 to 27 year olds, −12.5 dB in 66 to 82 year olds).⁴⁸

The higher parent vocabulary scores align with known increases in individuals' vocabulary ability which progresses rapidly in childhood, more slowly during adolescence,⁴⁹ which then continues through adulthood (whether measured by definitions, pictures, synonyms or antonyms) into the seventh decade of life.⁵⁰ Mean child standardised receptive vocabulary scores were comparable to published norms,³⁶ but our parent scores were higher than the normative 40 to 49-year-old values (mean 5.52 vs 3.23). This could reflect our sample's relative socioeconomic advantage,⁵¹ with advantage perhaps conferring a disproportionate rise in vocabulary beyond childhood. Alternatively, the American NPVT normative sample may not apply well to Australian parents.

Our children's sentence repetition scaled score (mean 9.8, SD 2.9) resembled both this subtest's published norms (mean 10.0, SD 3.0),³⁷ and 11 year old Australian scores on the full CELF-4 receptive (mean 99.4, SD 12.8) and expressive (mean 101.6, SD 13.2) scales (unpublished results, Early Language in Victoria Study, Australia – personal communication, Professor Sheena Reilly). This supports the validity of our iPad App, which greatly shortened its one-on-one administration time. Parent scaled scores (mean 9.1) were slightly lower than published standard scores for the CELF-4 normative ceiling subgroup (10.0 in 17 to 21 year olds), even though our parent raw scores were higher than their children's (mean 80.0 vs 69.7). The advantage of using the same measure in parents and children was its ability to support concordance analyses. However, using 17 to 21 year old norms for our parent sample at a mean age of 43.8 years may not

take into account known life course changes in auditory memory, which peaks at 20 years and begins declining from the fourth decade.⁵²

Children from 11 to 12 years resembled their parents more for language than hearing. Our concordance for hearing acuity (0.18) is remarkably similar to Framingham study values of 0.10, 0.19 and 0.13 for middle-, high- and low-frequency averages, respectively, at the much older mean offspring age of 60 to 70 years.²⁵ Thus, although hearing declines steadily from late childhood, children may not come to resemble their parents more with age. Our values also align with previous estimates of moderate polygenic heritability (0.26 to 0.32).²⁵ No studies have reported concordance for speech reception threshold, which here was virtually identical to that for acuity. Our child-parent concordance for receptive vocabulary (0.29) and sentence repetition (0.34) align almost exactly with heritability estimates (taking into account contributions from both parents) for latent oral language factors in the Twins Early Development Study of 60% at 12 years²⁴ and for vocabulary in the International Longitudinal Twin Study of 57% in fourth grade.⁵³ Therefore, despite their poorer language than singletons,⁵⁴ estimates of language heritability derived from twins most likely also apply at the unselected population level.

Unanswered questions and future research

This hearing and language battery could be used widely and efficiently in population studies. At a later date, we will be able to examine predictors (eg, ethnicity, education level, primary language spoken, inflammation, genome-wide association data) as these become available. For oral language, like other cognitive abilities,⁵⁵ heritability has been shown to increase across childhood to early adulthood. Future studies could further explore the shared genetic and environmental factors contributing to performance at different ages. This would allow a better understanding of what determines trajectories of hearing and language ability throughout life, potentially offering opportunities for innovative prevention.

Acknowledgements This paper uses unit record data from Growing Up in Australia, the Longitudinal Study of Australian Children. The study was conducted in partnership between the Department of Social Services (DSS), the Australian Institute of Family Studies (AIFS) and the Australian Bureau of Statistics (ABS). Research Electronic Data Capture (REDCap) tools were used in this study. More information about this software can be found at: www.project-redcap.org. The iPad Apps to administer the NPVT and Recalling Sentence were developed with support from the Centre of Research Excellence in Child Language (Australian National Health & Medical Research Council grant 1023493). The authors thank the LSAC and CheckPoint study participants, staff and students for their contributions.

Contributors JS and JW contributed to interpretation of results, drafting of the initial manuscript, critically revised further drafts and approved the final manuscript as submitted. ACG and KL contributed to interpretation of results, performed the statistical analysis, revised further drafts and approved the final manuscript as submitted. SAC revised further drafts and approved the final manuscript as submitted. MW is the Principle Investigator of the Child Health CheckPoint, planned the analyses and provided critical review of this manuscript.

Funding This work was supported by the National Health and Medical Research Council (NHMRC) of Australia (Project Grants 1041352, 1109355), The Royal Children's Hospital Foundation (2014–241), the Murdoch Children's Research

Institute (MCRI), The University of Melbourne, the National Heart Foundation of Australia (100660), Financial Markets Foundation for Children (2014-055, 2016-310), and the Victoria Deaf Education Institute. MW is supported by NHMRC Senior Research Fellowship 1046518 and Cure Kids New Zealand. The MCRI administered the research grants for the study and provided infrastructural support (IT and biospecimen management) to its staff and the study, but played no role in the conduct or analysis of the trial. DSS played a role in study design; however, no other funding bodies had a role in the study design and conduct; data collection, management, analysis and interpretation; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication. Research at the MCRI is supported by the Victorian Government's Operational Infrastructure Support Programme.

Disclaimer The findings and views reported in this paper are those of the authors and should not be attributed to DSS, AIFS or the ABS.

Competing interests MW received support from Sandoz to present at a symposium outside the submitted work.

Patient consent for publication Not required.

Ethics approval The CheckPoint data collection protocol was approved by The Royal Children's Hospital (Melbourne, Australia) Human Research Ethics Committee (3322D) and the Australian Institute of Family Studies Ethics Committee (14-26).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The Longitudinal Study of Australian Children datasets and technical documents are available to researchers at no cost via a licence agreement. Data access requests are co-ordinated by the National Centre for Longitudinal Data. More information is available at <https://dataverse.ada.edu.au/dataverse/lscac>.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Emmett SD, Francis HW. The socioeconomic impact of hearing loss in U.S. adults. *Otol Neurotol* 2015;36:545–50.
- Noble W. Preventing the psychosocial risks of hearing loss. *Aust Fam Physician* 2009;38:591–3.
- Mason A, Mason M. Psychologic impact of deafness on the child and adolescent. *Prim Care* 2007;34:407–26.
- Stothard SE, Snowling MJ, Bishop DV, et al. Language-impaired preschoolers: a follow-up into adolescence. *J Speech Lang Hear Res* 1998;41:407–18.
- Moore JK, Linthicum FH. The human auditory system: a timeline of development. *Int J Audiol* 2007;46:460–78.
- Stollman MH, van Velzen EC, Simkens HM, et al. Development of auditory processing in 6-12-year-old children: a longitudinal study. *Int J Audiol* 2004;43:34–44.
- Wassenberg R, Hurks PP, Hendriksen JG, et al. Age-related improvement in complex language comprehension: results of a cross-sectional study with 361 children aged 5 to 15. *J Clin Exp Neuropsychol* 2008;30:435–48.
- Baiduc RR, Poling GL, Hong O, et al. Clinical measures of auditory function: the cochlea and beyond. *Dis Mon* 2013;59:147–56.
- Cruikshanks KJ, Wiley TL, Tweed TS, et al. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The Epidemiology of Hearing Loss Study. *Am J Epidemiol* 1998;148:879–86.
- Gopinath B, Rochtchina E, Wang JJ, et al. Prevalence of age-related hearing loss in older adults: Blue Mountains Study. *Arch Intern Med* 2009;169:415–6.
- Shargorodsky J, Curhan SG, Curhan GC, et al. Change in prevalence of hearing loss in US adolescents. *JAMA* 2010;304:772–8.
- Cameron S, Glyde H, Dillon H. Listening in Spatialized Noise-Sentences Test (LiSN-S): normative and retest reliability data for adolescents and adults up to 60 years of age. *J Am Acad Audiol* 2011;22:697–709.
- Smits C, Kramer SE, Houtgast T. Speech reception thresholds in noise and self-reported hearing disability in a general adult population. *Ear Hear* 2006;27:538–49.
- Clegg J, Hollis C, Mawhood L, et al. Developmental language disorders—a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *J Child Psychol Psychiatry* 2005;46:128–49.
- Law J, Boyle J, Harris F, et al. The feasibility of universal screening for primary speech and language delay: findings from a systematic review of the literature. *Dev Med Child Neurol* 2000;42:190–200.
- Nelson HD, Nygren P, Walker M, et al. Screening for speech and language delay in preschool children: systematic evidence review for the US Preventive Services Task Force. *Pediatrics* 2006;117:e298–319.
- Law J, Boyle J, Harris F, et al. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int J Lang Commun Disord* 2000;35:165–88.
- McLeod S, McKinnon DH. Prevalence of communication disorders compared with other learning needs in 14,500 primary and secondary school students. *Int J Lang Comm Dis* 2007;42(Suppl 1):37–59.
- Morris MA, Meier SK, Griffin JM, et al. Prevalence and etiologies of adult communication disabilities in the United States: Results from the 2012 National Health Interview Survey. *Disabil Health J* 2016;9:140–4.
- Kvestad E, Czajkowski N, Krog NH, et al. Heritability of hearing loss. *Epidemiology* 2012;23:328–31.
- Wingfield A, Panizzon M, Grant MD, et al. A twin-study of genetic contributions to hearing acuity in late middle age. *J Gerontol A Biol Sci Med Sci* 2007;62:1294–9.
- Hendrickx JJ, Huyghe JR, Topsakal V, et al. Familial aggregation of pure tone hearing thresholds in an aging European population. *Otol Neurotol* 2013;34:838–44.
- Morell RJ, Brewer CC, Ge D, et al. A twin study of auditory processing indicates that dichotic listening ability is a strongly heritable trait. *Hum Genet* 2007;122:103–11.
- Hayiou-Thomas ME, Dale PS, Plomin R. The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev Sci* 2012;15:233–49.
- Gates GA, Couropmitree NN, Myers RH. Genetic associations in age-related hearing thresholds. *Arch Otolaryngol Head Neck Surg* 1999;125:654–9.
- Hoekstra RA, Bartels M, van Leeuwen M, et al. Genetic architecture of verbal abilities in children and adolescents. *Dev Sci* 2009;12:1041–53.
- Wake M, Clifford S, York E, et al. Introducing growing up in Australia's child health check point: a physical health and biomarkers module for the longitudinal study of Australian children. *Family Matters* 2014;95:15–23.
- Sanson A, Johnstone R. 'Growing Up in Australia' takes its first steps. *Family Matters* 2004;67:46–53.
- Edwards B. Growing up in Australia: the longitudinal study of Australian children: entering adolescence and becoming a young adult. *Family Matters* 2014:5.
- Clifford SA, Davies S, Wake M, et al. Child Health CheckPoint: cohort summary and methodology of a physical health and biospecimen module for the Longitudinal Study of Australian Children. *BMJ Open* 2019;9(suppl 3):3–22.
- Teunisse RJ, Olde Rikkert MG. Prevalence of musical hallucinations in patients referred for audiometric testing. *Am J Geriatr Psychiatry* 2012;20:1075–7.
- Cameron S, Dillon H. Development of the Listening in Spatialized Noise-Sentences Test (LiSN-S). *Ear Hear* 2007;28:196–211.
- Clifford SA, Davies Sarah Gillespie A, et al. Longitudinal Study of Australian Children's Child Health CheckPoint Data User Guide – December 2018. 2018 Melbourne: Murdoch Children's Research Institute. <https://doi.org/10.25374/MCRI.5687590>.
- Gershon RC, Slotkin J, Manly JJ, et al. IV. NIH Toolbox Cognition Battery (CB): measuring language (vocabulary comprehension and reading decoding). *Monogr Soc Res Child Dev* 2013;78:49–69.
- Akshoomoff N, Newman E, Thompson WK, et al. The NIH Toolbox Cognition Battery: results from a large normative developmental sample (PING). *Neuropsychology* 2014;28:1–10.
- Slotkin J, Kallen M, Griffith J, et al. NIH Toolbox: Technical Manual. In: *Health Nio*, editor 2012.
- Semel E, Wiig E, Secord W. *Clinical evaluation of language fundamentals – Fourth edition, Australian: Examiner's Manual*. Sydney, Australia: Pearson Clinical and Talent Assessment, 2006.
- Pink B. *Socio-Economic Indexes for Areas (SEIFA): Technical Paper*. Canberra, 2013.
- Heeringa SG WB, Berglund PA. *Applied survey data analysis*. Boca Raton: CRC Press, 2010.
- Ellul S, Hiscock R, Mensah FK, et al. Longitudinal Study of Australian Children's Child Health CheckPoint Technical Paper 1: Weighting and non-response. 2018 <https://doi.org/10.25374/MCRI.5687593>.
- Wang J, le Clercq CMP, Sung V, et al. Cross-sectional epidemiology of hearing loss in Australian children aged 11-12 years old and 25-year secular trends. *Arch Dis Child* 2018;103:579–85.

42. Wilson RH, McArdle R. A treatise on the thresholds of interoctave frequencies: 1500, 3000, and 6000 Hz. *J Am Acad Audiol* 2014;25:171–86.
43. Marcoux AM, Feder K, Keith SE, et al. Audiometric thresholds among a Canadian sample of 10 to 17 year old students. *J Acoust Soc Am* 2012;131:2787–98.
44. Williams W, Carter L, Seeto M. Hearing threshold levels for a population of 11 to 35 year old Australian females and males. *Int J Audiol* 2014;53:289–93.
45. le Clercq CMP, van Ingen G, Ruytjens L, et al. Prevalence of hearing loss among children 9 to 11 years old: the generation R study. *JAMA Otolaryngol Head Neck Surg* 2017;143:928–34.
46. Twardella D, Perez-Alvarez C, Steffens T, et al. The prevalence of audiometric notches in adolescents in Germany: The Ohrkan-study. *Noise Health* 2013;15:412–9.
47. Agrawal Y, Platz EA, Niparko JK. Prevalence of hearing loss and differences by demographic characteristics among US adults: data from the National Health and Nutrition Examination Survey, 1999–2004. *Arch Intern Med* 2008;168:1522–30.
48. Besser J, Festen JM, Goverts ST, et al. Speech-in-speech listening on the LiSN-S test by older adults with good audiograms depends on cognition and hearing acuity at high frequencies. *Ear Hear* 2015;36:24–41.
49. Rice ML, Hoffman L. Predicting vocabulary growth in children with and without specific language impairment: a longitudinal study from 2;6 to 21 years of age. *J Speech Lang Hear Res* 2015;58:345–59.
50. Salthouse TA. Quantity and structure of word knowledge across adulthood. *Intelligence* 2014;46:122–30.
51. Wolf MS, Curtis LM, Wilson EA, et al. Literacy, cognitive function, and health: results of the LitCog study. *J Gen Intern Med* 2012;27:1300–7.
52. Salthouse TA. Shared and unique influences on age-related cognitive change. *Neuropsychology* 2017;31:11–19.
53. Olson RK, Keenan JM, Byrne B, et al. Genetic and environmental influences on vocabulary and reading development. *Sci Stud Read* 2011;15:26–46.
54. Rutter M, Thorpe K, Greenwood R, et al. Twins as a natural experiment to study the causes of mild language delay: I: Design; twin-singleton differences in language, and obstetric risks. *J Child Psychol Psychiatry* 2003;44:326–41.
55. Haworth CM, Wright MJ, Luciano M, et al. The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* 2010;15:1112–20.