Available online at www.sciencedirect.com
SCIENCE DIRECT®
Brain and Language 94 (2005) 260–273
Brain and Language
www.elsevier.com/locate/b&l

Speech and non-speech processing in people with specific language impairment: A behavioural and electrophysiological study

G.M. McArthur [a,*], D.V.M. Bishop [b]

[a] Macquarie University, Macquarie Centre for Cognitive Science, Sydney, NSW 2109, Australia
[b] Department of Experimental Psychology, South Parks Road, Oxford, OX13UD, UK

Accepted 13 January 2005
Available online 16 February 2005

Abstract

McArthur and Bishop (2004) found that people with specific language impairment (SLI) up to 14 years of age have poor behavioural frequency discrimination (FD) thresholds for 25-ms pure tones, while people with SLI upto 20 years of age have abnormal auditory N1–P2–N2 event-related potential (ERP) responses to the same tones. In the present study, we extended these findings to more complex non-speech and speech sounds by comparing younger (around 13 years) and older (around 17 years) teenagers with SLI and controls for their behavioural FD thresholds and N1–P2 ERPs to 25 and 250-ms pure tones, vowels, and non-harmonic complex tones. We found that a subgroup of people with SLI had abnormal responses to tones and vowels at the level of behaviour and the brain, and that poor processing was associated with the spectral complexity of auditory stimuli rather than their phonetic significance. We suggest that both the age of listeners and the sensitivity of psychoacoustic tasks to age-related changes in auditory skills may be crucial factors in studies of sound processing in SLI.

Keywords: Specific language impairment; Frequency discrimination; Vowel discrimination; Event-related potentials

1. Introduction

Around 5% of children have problems understanding or producing spoken language despite having normal general cognitive abilities and no medical problems. The cause of this condition, known as specific language impairment (SLI), is not yet known. In fact, there may be no single cause. SLI may result when a number of risk factors co-occur (Bishop, Carlyon, Deeks, & Bishop, 1999). These risk factors may include deficient short-term verbal memory (Gathercole, 1993), a limited processing capacity (Leonard, 1996), and an impaired ability to acquire grammatical rules (Rice, Wexler, & Redmond, 1999). Another potential risk factor for SLI is poor auditory processing, which could affect the ability to discriminate between speech sounds. This may result in less stable or specific neural representations of speech sounds, which may ultimately interfere with perceiving and producing spoken language.

One aspect of auditory processing that is receiving increasing attention in studies of language impairment is frequency discrimination (FD). Although most research has focused on discrimination of brief or rapid stimuli (see Tallal, 2000; for review), some recent studies have found clear evidence for poor FD in children with SLI when there is no stress on rapid processing (Korpilahti, 1995; Ors et al., 2002; Uwer, Albrecht, & von Suchodoletz, 2002). For example, Mengler, Hogben, Michie, and Bishop (in press) found that a group of 9- to 12-year-olds with SLI were significantly poorer than age-matched controls at discriminating between the frequencies of 100-ms pure tones. When most of these children were re-tested $2\frac{1}{2}$ years later by Hill, Hogben, and Bishop (in press), the FD thresholds of the SLI group

* Corresponding author. Fax: +61 2 9850 6059.
E-mail address: gmcarthu@maccs.mq.edu.au (G.M. McArthur).

0093-934X/$ - see front matter
doi:10.1016/j.bandl.2005.01.002

had improved, but were still significantly poorer than the control group.

Other behavioural studies have found less clear evidence for poor FD in SLI. McArthur and Bishop (2004) found that only five out of sixteen 9- to 20-year-old teenagers with SLI had poorer FD thresholds than controls. When they re-tested most of these people $1\frac{1}{2}$ years later, two people with poor FD thresholds were still impaired while two others had improved to reach the normal range (Bishop & McArthur, in press). These variable findings could reflect heterogeneity of SLI: An FD deficit may be only one of a number of risk factors for SLI, so only a subgroup of people with SLI would perform poorly on behavioural tests of FD. However, another possibility suggested by our data (Bishop & McArthur, in press; Hill et al., in press) is that poor FD improves with age. A similar argument was recently put forward by Wright and Zecker (2004) to account for varied behavioural findings on auditory deficits in children with language and learning disabilities. They suggest that these children have a maturational lag in auditory discrimination of about 3–4 years. Whether an auditory deficit is seen in SLI will depend both on the age of the sample and the developmental trajectory of the specific auditory function that is tested. Wright and Zecker (2004) further postulated that auditory maturation ceases at puberty. This would explain why delayed auditory processing in some people with SLI does not catch up by early adulthood.

One limitation of behavioural studies of FD in people with SLI is that poor scores on behavioural FD tests could result from poor attention or low motivation as well as an FD deficit. Event-related potentials (ERPs) can be used to measure auditory processing without the attention of a listener. Auditory ERPs represent the average pattern of electrical activity emitted by groups of neurons in response to a sound. This activity, which is thought to reflect post-synaptic activation, is measured using electrodes that are placed on a listener's scalp. These electrodes detect the electrical activity that is present immediately after the onset of a sound. When many exemplars of the same sound are presented, the ratio of activity generated by a sound (signal) compared to activity generated by other factors (noise) becomes great enough for an auditory ERP waveform to emerge for that sound (see Fig. 2 for examples of ERP waveforms). The first and second positive peaks in the auditory ERP waveform are called P1 and P2, respectively. The first and second negative peaks are called N1 and N2, respectively. The locations of the neurons that produce these peaks are not fully understood. Some evidence suggests that P1 stems from the secondary auditory cortex (Liegeois-Chauvel, Musolino, Badier, Marquis, & Chauval, 1994); N1 originates from multiple sources that include the primary auditory cortex, the posterior–superior temporal plane, and non-specific frontal regions (Bruneau & Gomot, 1998; though cf Liegeois-Chauvel et al., 1994); P2 originates from the mesencephalic reticular activating system (Ponton, Eggermont, Kwong, & Don, 2000); and N2 is produced by cells in the superior temporal gyrus and medial temporal lobes (O'Donnell et al., 1993).

A handful of studies have tested people with SLI for their auditory ERPs to non-speech sounds. Some have found that people with SLI, as a group, had abnormal N1 responses (Lincoln, Courchesne, Harms, & Allen, 1995; Tonnquist-Uhlen, 1996; Tonnquist-Uhlen, Borg, Persson, & Spens, 1996) and P2 responses to tones (Adams, Courchesne, Elmasian, & Lincoln, 1987; Tonnquist-Uhlen, 1996). Neville, Coffey, Holcomb, and Tallal (1993) found that a subgroup of their children with SLI, who had poor performance on a behavioural auditory processing measure, had abnormal N1 and N2 responses to tones under certain conditions. However, other studies of SLI have found normal N1 or P2 responses to tones (Courchesne, Lincoln, Yeung-Courchesne, Elmasian, & Grillon, 1989; Marler, Champlin, & Gillam, 2002; Mason & Mellor, 1984; Ors et al., 2002).

It is difficult to explain these contradictory findings. There is no obvious difference between the studies that did and did not find abnormal auditory ERPs people with SLI in terms of stimulus type, inter-stimulus interval (ISI), or number of stimuli. However, these studies do differ slightly in sample size, with studies finding abnormal ERPs in people with SLI using larger groups. There may also be a difference in the types of people recruited for the SLI groups. Three of the five studies that found abnormal ERPs in SLI noted that their participants had particularly severe learning disabilities (i.e., Tonnquist-Uhlen, 1996; Tonnquist-Uhlen et al., 1996; Neville et al., 1993). For example, the two studies by Tonnquist-Uhlen and colleagues tested children with mild to moderate delays in mental function, whose reference values for their intelligence quotient were adjusted to the 8-year-old level (the mean age of the children was 11.9 years). Nine of their 20 children with SLI had pathological EEGs and two were border-line. Thus, this study may have tested children with more global cognitive impairments who would not usually meet the typical criteria for SLI.

Age is another factor that might explain why some studies find abnormal ERPs in people with SLI while others did not. It has been suggested that an auditory processing deficit present early in life may resolve with age (Farmer & Klein, 1995). We know that the studies that did and did not find abnormal ERPs in people with SLI had comparable age-ranges across studies (8–15 years and 7–16 years, respectively). However, we do not know the proportion of children at each age in each study. It is possible that the studies that found abnormal ERPs in children with SLI had a higher proportion of young children in their SLI group than studies that found normal ERPs in children with SLI.

Another factor to consider when measuring ERPs in children with SLI is that auditory ERP components change throughout childhood and adolescence. These changes make it difficult to study ERP components such as the N1 and P2 because they are often missing in children when relatively fast presentation rates are used (Albrecht, von Suchodoletz, & Uwer, 2000; Ponton et al., 2000). Fortunately, it is possible to compare ERPs with missing components using intra-class correlation coefficients (ICC). The ICC provides a global index of how similar two ERP waveforms are in their shape and amplitude. ICC values are like Pearson *r* values in that they range from 0 (the two waveforms are completely different) to 1.0 (the waveforms are exactly the same) to −1.0 (the waveforms are inverted). However, ICC values are typically much lower than Pearson *r* values because they take both the shape and the absolute amplitude of the two waveforms into account, while Pearson *r* values only consider the shape of the two waveforms.

McArthur and Bishop (2004) used ICCs to measure how well the auditory ERPs of people with SLI with poor and normal behavioural FD thresholds matched a grand average auditory ERP for controls the same age. They found that people with SLI with poor FD thresholds *and* people with SLI with normal FD thresholds had inappropriate auditory N1–P2–N2 ERPs for their age compared to controls. This was surprising as only the people with poor FD thresholds were expected to have abnormal auditory ERPs. However, it transpired that the people with SLI with poor FD thresholds tended to be younger than the people with SLI with normal FD thresholds. The interpretation offered by McArthur and Bishop (2004) for this pattern of results was similar to Wright and Zecker's (2004) explanation for conflicting auditory behavioural findings. That is, because auditory ERPs continue to mature into adulthood (Albrecht et al., 2000; Courchesne, 1990; Ponton et al., 2000), a teenager or young adult who is four years behind in auditory maturation will have an immature auditory ERP. However, behavioural FD thresholds typically reach adult levels at around 9 years of age (Thompson, Cranford, & Hoyer, 1999). Thus, a 4-year delay would not be evident in people with SLI above the age of around 13 years.

If people with SLI do have poor FD due to immature auditory processing, the question arises as to whether this is linked to abnormal speech perception (Rosen, 2003). It is possible that abnormal FD is a marker for a general developmental delay, but is not directly implicated in leading to observed language difficulties (Bishop & McArthur, 2004).

We are aware of only two studies that have compared FD and speech discrimination in the same group of people with SLI. Ors et al. (2002) measured how quickly 10- to 14-year-old children with SLI and controls responded to 1000-Hz rare deviant tones presented amongst frequent standard tones (the tones and ISIs were 50 and 1200 ms long, respectively). They were also tested for how quickly they could respond to 40 (Condition A) or 70 (Condition B), 350-ms, rare deviant speech sounds (the Swedish word "buss") presented amongst 160 (Condition A) or 140 (Condition B) frequent standard speech sounds (the Swedish word "puss"). The speech sounds were separated by an ISI of either 3000 ms (Condition A) or 1500 ms (Condition B). The children's electroencephalograms (EEGs) were measured while they did the discrimination tasks. The results revealed that the SLI group had slower behavioural responses for the 1000-Hz tones and for "buss." They also had significantly smaller and later P3 responses to both tones and the speech sounds. This suggested that there was an association between poor non-speech and speech processing in children with SLI at the level of behaviour and the brain. It also suggested that children with SLI may be slower to categorise sounds and update representations in working memory.

In a similar study, Uwer et al. (2002) measured how well 5- to 10-year-old children with receptive and expressive SLI and controls could discriminate between 175-ms, 1000-, and 1200-Hz tones, and between 175-ms consonant–vowels (CVs; /da/, /ga/, and /ba/), using a same-different behavioural task and an ERP mismatch response. The children with expressive SLI made significantly more errors than the other groups on the behavioural discrimination task for both the tones and CVs. The ERP data were less clear-cut: The differential brain response to standard and deviant stimuli (the mismatch response) was normal in the SLI group when tones were used, but was attenuated for CVs. However, the difference between standard and deviant tones in this study was substantial, leaving open the possibility that the method was too insensitive to detect more subtle FD deficits.

The first aim of the present study was to compare non-speech (pure tones) and speech (vowels) processing in people with SLI to controls using more sensitive threshold estimation methods than the methods of constant stimuli used by Ors et al. (2002) and Uwer et al. (2002). We also compared the auditory N1–P2 ERPs of people with SLI and controls to steady-state pure tones and vowels. If SLI is caused by abnormal processing of sound frequency, we might expect to find that people with SLI with abnormal FD thresholds and N1–P2 ERPs to pure tones would also have abnormal response thresholds and ERPs to vowels, since vowels are distinguished principally by sound frequencies that are relatively stable across time (Tallal & Piercy, 1975). Parenthetically, the integrity of vowel processing in people with SLI has not yet been resolved. Tallal, Stark, and Curtiss (1976) and Tallal and Stark (1981) found that 5- to 9-year-old children with SLI were able to discriminate between 250-ms vowels that differed in formant frequency as well as controls. However, Frumkin and Rapin (1980) and Stark and Heinz (1996) found that at least some 6- to 12-year-old children with SLI were less able than controls to discriminate between vowels that varied in formant frequency.

The second aim of the study was to identify whether poor vowel processing in people with SLI stemmed from a speech-specific deficit that is only seen when stimuli have phonetic status (i.e., are perceived as speech sounds), or a general FD deficit that is stressed by the spectral complexity of vowels. The vowels used in this study were composed of a fundamental formant (F0) with a frequency that was fixed between vowels (200 Hz), a higher formant (F1) with a frequency that was varied between vowels (600–1000 Hz), and two higher formants (F2 and F3) with frequencies that were fixed between vowels (2200 and 3000 Hz, respectively). To discriminate between vowels, a listener had to discriminate between F1 frequencies that were embedded amongst the F0, F2, and F3 frequencies. Wright et al. (1997) found that children with SLI were less able to discriminate a tone followed by a masking noise containing similar frequencies compared to a tone followed by a masking noise containing a spectral notch around the frequency of the tone. They suggested that the children with SLI were less able to separate the frequencies of the tone from the noise. The presence of the F0, F2, and F3 formant frequencies in the vowels used in this study could place extra pressure on discriminating F1 in the same way. This might make vowel processing harder for a person with an FD deficit than a person with normal FD.

We tested whether poor vowel processing resulted from a speech-specific deficit or an FD deficit that is stressed by spectral complexity by comparing the processing of vowels to non-harmonic complex tones. These non-harmonic complex tones resembled the vowels insofar as two higher frequency bands were present in the signal (i.e., F2 and F3) in addition to the frequency to be distinguished (i.e., F1). However, they had no harmonics, no F0, and no formant bandwidths, so they were not at all speech-like. If people with SLI had poor vowel processing because they had a speech-specific deficit, then they should have poorer responses to the vowels than the non-harmonic complex tones, and their vowel thresholds should correlate less well with their complex-tone thresholds. However, if they had poor vowel processing because they had an FD deficit that was stressed by spectral complexity, then they should have poor responses to both the vowels and the non-harmonic complex tones, and their vowel thresholds should correlate well with their complex tone thresholds.

2. Methods

Methods were approved by the Ethics Committee of the Department of Experimental Psychology at the University of Oxford. Each listener and their parent or guardian gave their informed consent to participate in the research.

2.1. Listeners

The 16 people with SLI (seven females) were aged from 12 to 21 years, and were recruited from language development centres and support groups. All these participants had received special educational support for SLI, and had previously had thorough screening to ensure that no exclusionary conditions (hearing loss, autistic disorder, neurological damage, etc.) were present. The 16 people with normal spoken language skills (controls; nine females) were recruited from scout and guide groups, a college, and a high school. They were matched to the people with SLI for age and non-verbal IQ. Twenty-four of these participants (11 people with SLI and 13 controls) had taken part in the study by McArthur and Bishop (2004) some 18 months previously. All participants had non-verbal IQ scores within the average range, had no reported auditory, physiological, or neurological problems. They also had normal hearing sensitivity for tones (i.e., were able to detect a tone at 20 dB HL) that were similar in frequency (i.e., 750 Hz) to the components of the experimental sounds that were adjusted to calculate thresholds.

The people with SLI scored more than 1 *SD* below the age-mean level on at least two of four key spoken language tests (see below). Controls scored within the average range on at least three of the four spoken language tests. Group statistics are illustrated in Table 1.

Because we had previously found that FD thresholds varied with age in people with SLI (McArthur & Bishop, 2004), we subdivided participants into two age bands. We used a cut-off point of 14.5 years to subdivide the groups because it was the central point that was closest to the cut-off point used in McArthur and Bishop (2004, 14 years) that produced the best balance in group sizes. There were six younger teenagers with SLI (SLI-young group; 12.39–14.42 years; $M = 13.37$ years, $SD = 0.90$; 2 females), 10 older teenagers with SLI (SLI-old group; 14.92–20.88 years; $M = 17.28$ years, $SD = 2.32$; 5 females), seven younger teenagers with normal language skills (control-young group; 12.34–13.67 years; $M = 12.84$ years, $SD = 0.47$; 4 females), and nine older teenagers with normal language skills (control-old group; 14.71–21.01 years; $M = 16.71$ years, $SD = 2.24$; 5 females).

2.2. Psychometric tests

The psychometric test battery is described in detail in McArthur and Bishop (2004). We used the Standard Progressive Matrices (Raven, Raven, & Court, 1998) to ensure that all participants had non-verbal IQ of 75

Table 1
Mean age, non-verbal IQ, and spoken language scores of the SLI and control groups

	SLI ($N = 16$)			Control ($N = 16$)			Group Effect	(η^2)[b]
	M	*SD*	Range	*M*	*SD*	Range		
Age	15.82	2.70	12–20	15.02	2.58	12–21	$t(30) = 0.85, p = .40$	.02
Non-verbal IQ	91.87	13.36	75–113	97.50	9.70	75–113	$t(30) = 1.36, p = .18$	.06
BPVS	79.12	15.88	56–110	109.94	9.68	93–128	$t(30) = 6.63, p < .001$[a]	.59
Figurative language	4.38	1.78	3–9	11.12	3.05	7–16	$t(30) = 7.64, p < .001$[a]	.66
Recreating sentences	4.31	2.15	3–6	8.43	2.13	6–13	$t(30) = 5.45, p < .001$[a]	.50
Recalling sentences	4.25	1.34	3–10	10.25	1.98	6–13	$t(30) = 10.02, p < .001$[a]	.77

[a] $p < .05$.
[b] Effect size is represented by Eta squared (η^2), which is calculated in the same way as r^2, but is reported as η^2 when one variable is non-linear.

or above, plus four oral language tests that are widely used in the diagnosis of SLI: the British Picture Vocabulary Scale (BPVS), Long Form (Dunn, Dunn, Whetton, & Pintilie, 1982), the Recalling Sentences sub-test of the Clinical Evaluation of Language Fundamentals-Revised (Semel, Wiig, & Secord, 1987), and the Recreating Sentences and Figurative Language sub-tests of the Test of Language Competence-Expanded Edition (TLC-EE; Wiig & Secord, 1989). Standard scores on the Raven's Matrices and BPVS have mean of 100 and *SD* of 15, and standard scores on the other subtests have a mean of 10 and standard deviation of 3.

2.3. Behavioural thresholds

2.3.1. Stimuli

There were six stimulus conditions: 25- and 250-ms pure tones, 25- and 250-ms vowels, and 25- and 250-ms non-harmonic complex tones. We used the 25-ms sounds to obtain behavioural responses to the same sounds that were used to trigger the ERPs. We used the 250-ms sounds to check that the 25-ms FD thresholds of the SLI groups were not disproportionately confounded by the brevity of the 25-ms sounds compared to the controls. This is important because a popular theory holds that people with SLI are have a specific problem processing rapid or brief auditory information (Tallal, 2000). Each condition presented standard and deviant sounds at 80 dB SPL (see Procedure). The conditions were counterbalanced between listeners.

The standard and deviant sinusoidal pure tones had 2.5-ms onset and offset ramps. Standard tones had a frequency of 600 Hz while deviant pure tones had a higher frequency that was adjusted between trials (602–1000 Hz).

The standard and deviant vowels, which fell along the /ɛ/-/a/ continuum (i.e., s*e*t to s*a*t), were generated with the Cascade branch of a Klatt synthesiser. We used synthesised vowels rather than spoken vowels because we wanted to measure the discrimination thresholds of vowels using the same adaptive psychophysical procedures that were used to measure thresholds for pure tones and non-harmonic complex tones. It is difficult to generate spoken speech stimuli that vary in the systematic way that is required by these adaptive procedures. We generated the vowels using the formant frequencies and bandwidths outlined by Kewley-Port and Watson (1994). The standard vowel (/ɛ/) set f_0, F1, F2, and F3 at 200, 600, 2200, and 3000 Hz, respectively, and used bandwidths of 70, 90, and 170 Hz for F1, F2, and F3, respectively. The deviant vowels were the same except that the frequency of F1 was higher than the standard vowel and was adjusted between trials (602–1000 Hz). The 250-ms vowels increased to half their amplitude from 0 to 30 ms, increased to their full amplitude from 31 to 230 ms, and then fell back to 0 from 231 to 250 ms. The 25-ms vowels simply increased in amplitude from 0 to 25 ms. In both the 25-ms and 250-ms conditions, the gain of each deviant vowel was adjusted to ensure that they were all the same amplitude (i.e., 80 dB SPL).

The 25- and 250-ms standard and deviant non-harmonic complex tones had rise and fall times of 2.5 ms, and were composed of three pure tones with frequencies similar to the F1, F2, and F3 frequencies of the vowel formants (i.e., 600, 2207, and 3001 Hz, respectively). The two highest pure tones were set at 2207 and 3301 Hz rather than 2200 and 3000 Hz (respectively) to avoid harmonics. Thus, the non-harmonic complex tones were less spectrally complex than the vowels but more spectrally complex than the pure tones. The deviant non-harmonic complex tones were the same as the standard non-harmonic complex tone except that the lowest pure tone was set at a higher frequency that was adjusted between trials (602–1000 Hz).

2.3.2. Procedure

Behavioural FD thresholds for the 25- and 250-ms pure tones, vowels, and non-harmonic complex tones were calculated using the same psychophysical paradigm. Each trial in the paradigm was composed of three sounds of the same type (e.g., 25-ms pure tones) that were each separated by 500-ms silence. The second sound was the "standard" (i.e., the frequency of the pure tones, the F1-frequency of the vowels, and the frequency of the lowest pure tone of the non-harmonic complex tones was set at 600 Hz) and was visually represented by a dinosaur that jumped in the centre of the PC monitor when the standard was presented. Either the first sound (visually represented by a yellow ball, to the left of the dinosaur, which jumped when the first stimulus was presented) or the third sound (represented by a red ball, to the right of the dinosaur, which jumped when the third stimulus was presented) was the same as the standard sound. This was randomly allocated between trials. The remaining "deviant" sound was different to the standard stimulus (i.e., the frequency of the pure tone, the F1-frequency of the vowel, and the frequency of the lowest pure tone of the non-harmonic complex tone was higher than 600 Hz). The listener used a PC mouse to click on the ball that represented the deviant sound. If the response was correct, a thumbnail image was added to a column of images on the left edge of the PC screen. When the column reached the top of the screen, the dinosaur and the two balls jumped up and down to a short tune.

The parameter estimation by sequential tracking (PEST) procedure was used to adjust the frequency of the deviant sound between trials to the level where the listener correctly identified the different sound 79% of the time (Taylor & Creelman, 1967). The frequency was initially set at 1000 Hz (ceiling value) and was adjusted in 80-Hz steps. These were reduced to 2 Hz after the first six reversals in response adjustment. The task finished when the listener made 10 reversals in response adjustment or had completed 60 trials. A listener's threshold score was the mean frequency after the sixth reversal in response adjustment. Higher threshold scores represented poorer discrimination. These scores were transformed into logarithmic units to normalise their distribution for statistical analyses.

2.4. Auditory ERPs

2.4.1. Stimuli

The three stimulus conditions—25-ms pure tones, 25-ms vowels, and 25-ms non-harmonic complex tones—were all presented at 80 dB SPL. It was not feasible to measure ERPs to the 250-ms sounds because we wanted to minimise the length of the testing session (i.e., 3 h) for the comfort of our volunteers. We used 25-ms stimuli rather than 250-ms stimuli because we had already established that 25-ms sounds produce reliable auditory ERPs in both adults (McArthur, Bishop, & Proudfoot, 2003) and people with SLI (McArthur & Bishop, 2004).

Each condition was composed of 1700 standard and 300 deviant trials, which were divided into eight blocks of 250 trials. Four of the eight blocks in each condition used 600-Hz standards (i.e., set the frequency of the pure tones, the F1 of the vowels, and the lowest pure tone of the non-harmonic complex tones at 600 Hz) and 700-Hz deviants (i.e., set the frequency of the pure tones, the F1 of the vowels, and the lowest tone of the non-harmonic complex tones at 700 Hz). The remaining four blocks reversed the stimuli. (*Note*. we presented standard and deviant stimuli in each condition to calculate the mismatch negativity (MMN). However, the MMN was not reliable enough to be included in the analysis.) The 24 blocks of trials (eight blocks for each of the three stimulus types) were presented in random order. The mean stimulus-onset-asynchrony (SOA) between each trial was 734 ms. The SOA was randomly jittered between trials (the standard deviation was 78 ms) to avoid anticipatory ERP effects.

2.4.2. Procedure

Volunteers were seated in a lounge chair in an electrically and acoustically shielded testing booth. The sounds were presented diotically through headphones while listeners watched a video on a small television 1.3 m away. The soundtrack of the television was played at a low-level (approximately 50 dB SPL) to alleviate boredom and maintain attention away from the experimental sounds. Playing a video soundtrack at this volume has little effect on the reliability of the auditory P1, N1, and P2 components in adults (McArthur et al., 2003).

The EEG was recorded from sintered electrodes that were placed in line with the 10–20 International system: two sites along the midline of the head (Fz and FCz), four sites over the left hemisphere (FP1, F3, FC3, and F7), and four sites over the right hemisphere (FP2, F4, FC4, and F8). The ground electrode was positioned on the midline between FPz and Fz. Linked mastoids were used as the online reference. The vertical electro-oculogram (VEOG) was recorded from above and below the right eye; the horizontal electro-oculogram (HEOG) was recorded 1 cm from the outside of the outer canthi of each eye. The signal was amplified 20,000 times and sampled at 250 Hz (i.e., once every 4 ms) with an online bandpass filter of 0.05–30 Hz.

Each participant's EEG was processed offline. The influence of VEOG activity was removed from the EEG sites (ocular artefact reduction) using an algorithm of an average 'blink' that was calculated from at least 20 VEOG epochs of 400 ms that were triggered by a 10% increase in VEOG activity (Neurosoft Inc., 1999). The EEG was divided into 536-ms epochs with a 50 ms prestimulus interval. Epochs were baseline corrected from −50 to 0 ms. Epochs with changes in HEOG or EEG activity greater than 150 μV from baseline were rejected.

There were high ICCs between each listener's entire ERP (i.e., −50 to 486 ms) to 600-Hz stimuli (tones, vowels, or complex tones) and their ERP to 700-Hz stimuli (tones, vowels, and complex tones, respectively), which did not differ between groups (see Section 3.2.1 and Table 3 for statistics). Therefore, ERPs were calculated by averaging the epochs of all the standard stimuli (600 and 700 Hz) bar those that fell immediately after a deviant stimulus (i.e., a maximum of 1400 standard stimuli in each condition; see Section 3.2.1 and Table 3 for statistics on the number of accepted epochs for each stimulus in each group). We used the activity at Fz to represent auditory ERPs as it appears to be the most reliable single site for measuring auditory ERPs. It is the site that produces the largest response, it is the site most commonly used to represent auditory ERPs, and is one of the few sites where adults and children produce analogous ERP components (Ponton et al., 2000). We did not analyse activity at the other channels because we used linked mastoids to record the EEG so our data were not appropriate for comparing activity across hemispheres. The N1 response measured at fronto-central sites is often referred the N1b subcomponent of the N1 response. However, in the absence of any data for discriminating between the N1 subcomponents in this study, we shall simply refer to this component as N1.

Seven listeners did not have a N1–P2 complex to either pure tones, vowels, or non-harmonic complex tones (three people in the SLI-young group, two people in the SLI-old group, one person in the control-young group, and one person in the control-old group). Consequently, we used ICCs to measure how appropriate each listener's auditory ERP was for their age in the N1–P2 ERP region, which was taken as the interval 100–228 ms post-stimulus-onset for comparability with our previous analyses (Bishop & McArthur, in press). For listeners with SLI, we calculated the ICC between their own ERP waveform and the average ERP waveform of controls either younger than 14.5 years (if the listener with SLI was younger than 14.5 years) or the average ERP waveform of controls older than 14.5 years (if the SLI listener was older than 14.5 years) in the N1–P2 ERP region. We did the same for controls except that we compared their own individual ERP waveform to the mean ERP waveform of all the other controls in their appropriate age group (i.e., above or below 14.5 years). We did not include a control listener's ERP in the mean ERP waveform for their age group because this would have artificially inflated each control listener's ICC.

We calculated ICC values using the formula: (MS between − MS within)/(MS between + MS within), where MS between $= (((\sum(X^2) + \sum(Y^2) + 2\sum(XY))/2 - (\sum(X) + \sum(\mathrm{Y}))\text{^}2/(2N)))/(N-1)$, MS within $= (0.5(\sum(X^2) + \sum(Y^2)) - \sum(XY))/N$, and N is the number of pairs of data points. This is equivalent to the formula used to compute ICC values in the SPSS one-way random model reliability analysis (Note: this is not equivalent to the ICC analysis in the Neurosoft Inc. (1999) software). The ICC values were transformed to Fisher z values to normalise the data for statistical analyses. The larger the ICC score, the better the match between a listener's auditory ERP and the auditory ERP expected for their age.

3. Results

3.1. Behavioural FD thresholds

A Levene Median test indicated that there was no significant difference between the variance in group scores for each stimulus condition. Kolmogorov–Smirnov tests of normality indicated that the distributions of all bar two of the 4 (groups) by 3 (stimuli) by 2 (durations) variables did not differ significantly from normal (the exceptions were 250-ms vowels for the SLI-old group: $D(10) = .27$; $p = .03$; and 25-ms complex tones for the SLI-old group: $D(10) = .28$; $p = .02$; Note: At least one of these significant effects may be due to a Type I error resulting from multiple (i.e., 24) comparisons for normality). Thus, parametric statistics were used to test the data for significant main effects and interactions. A criterion of $p < .05$ was used to test whether differences between means were statistically significant in all behavioural and ERP analyses.

The thresholds of the SLI-young, SLI-old, control-young, and control-old groups for 25-ms (left) and 250-ms (middle) pure tones, vowels, and non-harmonic complex tones are illustrated in Fig. 1. A 2 (group: SLI versus control) by 2 (age: young versus old) by 3

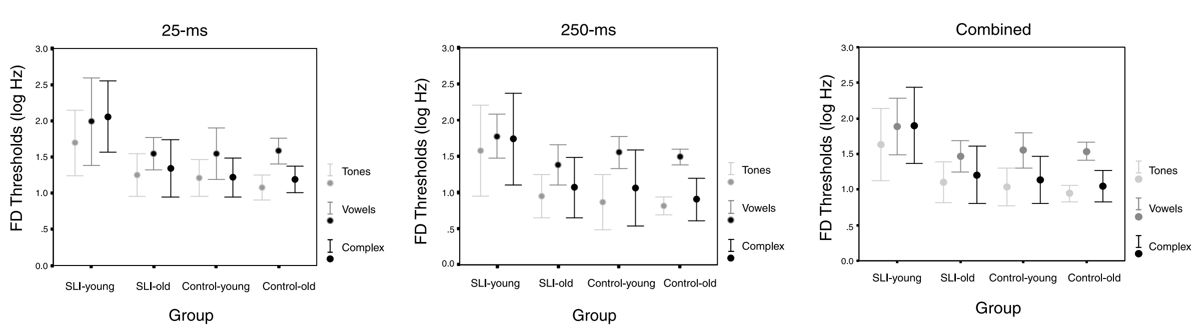

Fig. 1. Mean thresholds (log Hz) and 95% confidence intervals for tones (light grey), vowels (dark grey), and non-harmonic complex tones (black) of the SLI-young, SLI-old, control-young, and control-old groups for 25-ms stimuli (left), 250-ms stimuli (middle), and 25- and 250-ms stimuli combined (right).

(stimulus: pure tones versus vowels versus non-harmonic complex tones) by 2 (duration: 25-ms versus 250-ms) repeated measures ANOVA revealed a main effect of duration, with lower (i.e., better) thresholds for 250-ms stimuli than 25-ms stimuli in all four groups (see Table 2 for statistics). There was no duration by group interaction ($F(1,56) = 0.35$, $p = .56$, $\eta^2 = .01$), duration by age group interaction ($F(1,56) = 0.34$, $p = .57$, $\eta^2 = .01$), or duration by group by age interaction ($F(1,56) = 0.2$, $p = .88$, $\eta^2 = .001$). Given (1) stimulus duration did not have a disproportionate effect on the SLI groups compared to the control groups, and (2) the need to maximise statistical power to compensate for small sample sizes, we combined the data for the 25 and 250-ms sounds. The mean discrimination thresholds for tones, vowels, and complex tones are illustrated in Fig. 1 (right-hand side).

There was a main effect of group due to the higher (i.e., poorer) FD thresholds of the listeners with SLI compared to the controls. There was also a main effect of age, with younger participants having higher FD thresholds than older participants. There was a significant group by age interaction, which was examined further by separate ANOVAs for the younger and older groups. For the young participants (below 14.5 years), there was a significant effect of group ($F(1,11) = 8.29$, $p = .02$, $\eta^2 = .43$), with the SLI-young group having higher FD thresholds than the control-young group. There was also a significant interaction between group and stimulus ($F(2,22) = 4.24$, $p = .03$, $\eta^2 = .28$). To explore this interaction, paired t tests were conducted to compare average thresholds for pure tones, vowels, and non-harmonic complex tones for the SLI-young and control-young groups separately. While the control-young group had significantly higher thresholds for vowels compared to non-harmonic complex tones ($t(6) = 3.96$, $p = .007$) and compared to pure tones ($t(6) = 5.63$, $p = .001$), the SLI-young group did not (vowels versus non-harmonic complex tones: $t(5) = 0.11$, $p = .92$; vowels versus pure tones: $t(5) = 1.70$, $p = .14$). For the old groups (aged above 14.5 years), the effect of group was not significant ($F(1,17) = 0.37$, $p = .55$), and no interactions with group were significant.

There was also a main effect of stimulus, with higher thresholds for vowels than non-harmonic complex tones, which in turn were higher than thresholds for pure tones. This stimulus effect interacted significantly with group. To examine this further, paired t tests were used to compare average thresholds for vowels, non-harmonic complex tones and pure tones for the SLI and control groups separately. In the control group, thresholds for vowels were significantly higher than for non-harmonic complex tones ($t(15) = 7.31$, $p < .001$) but the difference between non-harmonic complex tones and pure tones just fell short of significance ($t(15) = 1.87$, $p = .08$). Conversely, in the SLI group, thresholds for vowels and non-harmonic complex tones did not differ ($t(15) = 1.33$, $p = .20$), but thresholds for non-harmonic complex tones were significantly higher than for pure tones ($t(15) = 2.81$, $p = .01$).

The Pearson r correlation coefficient between thresholds for vowels and pure tones was .70 for the SLI-young group, .49 for the SLI-old group, .62 for the control-young group, and .52 for the control-old group. The Pearson r correlation coefficient between thresholds for vowels and complex tones was .71 for the SLI-young group, .44 for the SLI-old group, .65 for the control-

Table 2
Repeated measures ANOVAs on behavioural FD thresholds, ERP reliability measures (number of epochs and split-half reliability), and N1–P2–N2 ICC values

Measure	Main effects	Significant interactions[b]
Behavioural FD thresholds	Group[a]: $F(1,28) = 7.89$, $p = .009$, $\eta^2 = .22$[c] Age[a]: $F(1,28) = 7.15$, $p = .01$, $\eta^2 = .20$ Stimulus[a]: $F(2,56) = 32.12$, $p < .001$, $\eta^2 = .53$ Duration[a]: $F(1,28) = 34.18$, $p < .001$, $\eta^2 = .55$	Group × Age[a]: $F(1,28) = 4.39$, $p = .04$, $\eta^2 = .14$ Stimulus × Group[a]: $F(2,56) = 5.01$, $p = .01$, $\eta^2 = .15$
Split-half reliability ICC (Fisher z)	Group: $F(1,27) = 5.58$, $p = .03$, $\eta^2 = .17$ Age: $F(1,27) = 0.80$, $p = .38$, $\eta^2 = .03$ Stimulus: $F(2,54) = 1.98$, $p = .15$, $\eta^2 = .07$	
Number of epochs	Tones: $H(3) = 9.57$, $p = .02$ Vowels: $H(3) = 3.16$, $p = .37$ Complex tones: $H(3) = 7.13$, $p = .07$	
N1–P2–N2 ICC (Fisher z)	Group[a]: $F(1,27) = 15.75$, $p < .001$, $\eta^2 = .37$ Age: $F(1,27) = 2.00$, $p = .17$, $\eta^2 = .07$ Stimulus: $F(2,54) = 2.25$, $p = .12$, $\eta^2 = .08$	Group × Age[a]: $F(1,27) = 4.72$, $p = .04$, $\eta^2 = .15$

[a] $p < .05$.
[b] Only values for significant interactions are included.
[c] Effect size is represented by Eta squared (η^2), which gives the proportion of variance in the dependent variable accounted for by an independent variable.

young group, and .57 for the control-old group. These moderate to strong correlation coefficients did not reach statistical significance because of low statistical power due to small sample sizes.

3.2. Auditory ERPs

3.2.1. Reliability

Our first consideration was whether the ERPs of the younger and older SLI and control groups were equally reliable. We tested this in two ways. First, we used ICCs to calculate the split-half reliability between each participant's mean ERP to the 600-Hz stimuli (pure tones, vowels, and non-harmonic complex tones, respectively) compared to their mean ERP to 700-Hz stimuli (pure tones, vowels, and non-harmonic complex tones, respectively). Second, we compared the mean number of epochs that were included in each listener's ERP. We found that one young control had remarkably low split-half ICCs (−0.21, 0.03, 0.31) and far fewer epochs (around 400 out of a possible 1400) than the other participants. The ERPs of this child were excluded from all subsequent analyses, so the control-young group was composed of six, rather than seven, participants.

Kolmogorov–Smirnov tests of normality on the split-half ICCS to tones, vowels, and complex tones indicated that the distributions of all 4 (groups) by 3 (stimuli) variables did not differ significantly from normal. Levene Median tests indicated that there was no significant difference between the variance of the four groups for each stimulus condition. Thus, a 2 (group) by 2 (age) by 3 (stimulus) repeated measures ANOVA was used to test the split-half ICCs for significant main effects and interactions (see Table 2 for ANOVA results and Table 3 for means and standard deviations).

The main effects of age and stimulus, and the group by age interaction, were not statistically significant. There was a main effect of group, with higher spilt-half ICCs in the control group. This stemmed from unusually high spilt-half ICCs of the control-young group compared to the control-old, SLI-young and SLI-old groups. Post hoc t tests revealed that the spilt-half ICCs of the SLI and control groups only differed for vowels ($t(29) = 32.59$, $p = .02$). Pearson r correlation coefficients were used to test whether lower split-half ICCs for vowels might confound N1–P2 ICCs for vowels. There were only very low or inverse coefficients in the SLI group ($r = -.39$, $p = .14$), the control group ($r = -.13$, $p = .64$), and the two groups combined ($r = .06$, $p = .73$) between split-half ICCs and N1–P2 ICCs. Thus, the (robust) split-half reliability of the SLI and control groups' ERPs were independent of their N1–P2 ICCs.

Turning to the number of accepted epochs in each listener's ERP, Levene Median tests indicated that there was no significant difference between the variance of the four groups for each stimulus condition. Kolmogorov–Smirnov tests of normality indicated that the distributions of the four groups did not differ for tones or complex tones. However, they did differ for vowels. Thus, a Kruskal–Wallis test was used to compare the number of accepted epochs across the four groups for the three types of sounds (see Table 2 for ANOVA results and Table 3 for means and standard deviations). There was no significant difference across the four groups for vowels or complex tones. There was a significant difference for tones, with particularly high ranks for the control-old group compared to the control-young, SLI-young and SLI-old groups. Post hoc Mann–Whitney U tests revealed that the elevated rankings of the control-old group were significantly higher

Table 3
ERP reliability and N1–P2–N2 ICC measures to tones, vowels, and non-harmonic complex tones in the SLI-young, SLI-old, Control-young, and Control-old groups

	SLI-young ($N = 6$)		SLI-old ($N = 10$)		Control-young ($N = 6$)		Control-old ($N = 9$)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Split-half reliability ICC (Fisher z)[a]								
Tones	0.86	0.27	0.99	0.46	1.58	0.46	0.94	0.43
Vowels	1.02	0.33	1.09	0.63	1.59	0.31	1.39	0.30
Complex	0.97	0.62	1.20	0.54	1.41	0.24	1.10	0.52
Number of epochs[a]								
Tones	1321.00	115.72	1303.80	79.33	1348.67	50.62	1385.89	9.44
Vowels	1319.00	131.34	1309.70	72.62	1344.50	62.59	1382.00	16.76
Complex	1317.50	117.39	1324.50	76.38	1384.33	13.78	1361.78	40.97
N1–P2–N2 ICC (Fisher z)								
Tones	0.13	0.51	0.22	0.29	0.79	0.25	0.39	0.55
Vowels	0.33	0.27	0.29	0.55	0.97	0.43	0.51	0.47
Complex	−0.09	0.28	0.16	0.24	1.01	0.47	0.48	0.51

[a] $p < .05$.

than those of the SLI-young ($U = 10$, $p = .04$) and SLI-old ($U = 9.5$, $p = .004$) groups, which did not differ from the control-young group. Spearman rank correlation coefficients were used to test if the number of accepted epochs for tones might confound N1–P2 ICCs for tones. There was no relationship between the number of accepted epochs for tones and N1–P2 ICCs to tones in the SLI group ($r_s = .16$, $p = .54$), the control group ($r_s = -.07$, $p = .82$), or the two groups combined ($r_s = .22$, $p = .23$). Thus, the (substantial) number of accepted epochs in both the SLI and control groups was not related to their N1–P2 ICCs.

3.2.2. N1–P2 ERPs

The N1–P2 ICC values reflect the extent to which an individual's waveform in the N1–P2 region resembles the grand average of controls in the same age band. The auditory ERPs and mean ICCs of the SLI-young, SLI-old, control-young, and control-old groups for pure tones, vowels, and non-harmonic complex tones are shown in Fig. 2 and Table 3, respectively. (*Note.* The apparent group difference in P1 in Fig. 2 is misleading: The difference between the groups in the P1 region were not statistically significant.)

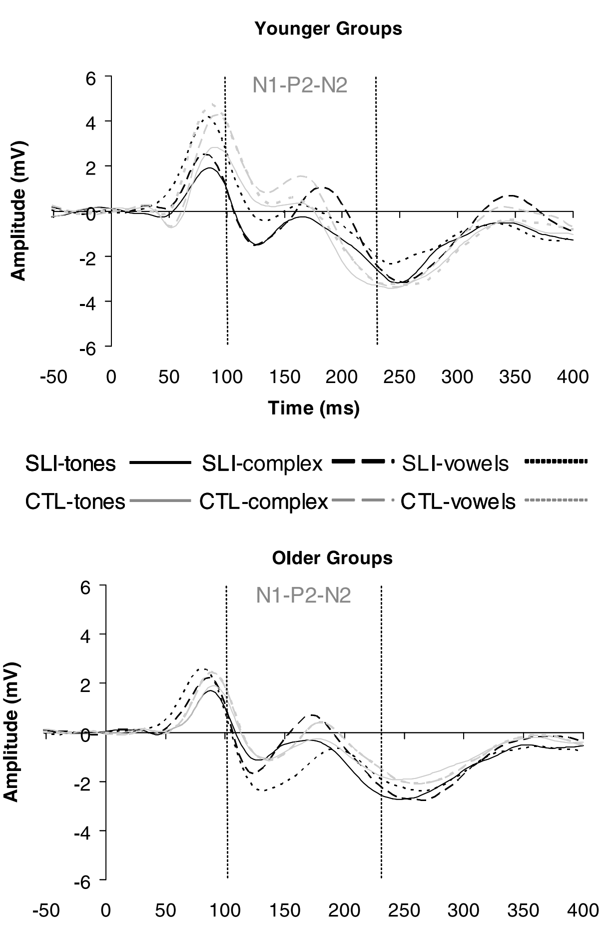

Fig. 2. Mean auditory ERPs of the younger (top) and older (bottom) SLI groups (black lines) and control groups (grey lines) to tones (unbroken line), vowels (dotted line), and complex tones (unbroken line).

Levene Median tests indicated that there was no significant difference between each group's variance for each stimulus condition. Kolmogorov–Smirnov tests of normality indicated that the distributions of all 4 (groups) by 3 (stimuli) variables did not differ significantly from normal. Thus, a 2 (group) by 2 (age) by 3 (stimulus) repeated measures ANOVA (see last row of Table 2) was used to test the data for significant main effects and interactions. There was a significant main effect of group, due to the lower ICCs of the participants with SLI compared to controls. Thus, the N1–P2 ERPs of the SLI participants were less appropriate for their age than the N1–P2 ERPs of the control participants. There was no significant main effect of age. However, there was a significant group by age interaction, with younger participants having higher ICCs than older participants in the control group. Further exploration of this effect revealed that it was driven by four people in the control-old group who had low ICCs. Three of these four people fell close to the 14.5 year cut-off point (i.e., were 14.71, 15.42, and 15.53 years of age). Reclassifying any one of these participants into the control-young group reduced the group by age interaction to a non-significant level. Thus, the reliability of this interaction is questionable. There were no main effects or interactions between the ICC values to the different types of stimuli.

3.3. Behavioural thresholds and auditory ERPS combined

The data of the SLI and control groups was separated into younger listeners (i.e., SLI-young and control-young groups combined) and older listeners (SLI-old and control-old groups combined). The Pearson r correlation coefficients between the behavioural thresholds and ERP ICCs for pure tones, vowels, and non-harmonic complex tones in the younger group ($N = 12$) were $-.42$ ($p = .18$), $-.44$ ($p = .15$), and $-.77$ ($p = .003$), respectively. The correlation coefficients between the same variables in the older group ($N = 19$) were $-.05$ ($p = .84$), $-.01$ ($p = .97$), and $-.29$ (.22), respectively.

4. Discussion

4.1. Is non-speech processing associated with poor speech processing in people with SLI?

In our previous study (McArthur & Bishop, 2004) we found two lines of evidence for abnormal auditory processing in people with SLI. First, behavioural FD thresholds for pure tones were abnormally high in the SLI-young group. Second, ERPs to pure tones in the N1–P2–N2 region were age-inappropriate in both the younger and older SLI groups, regardless of their FD thresholds. If these non-verbal auditory deficits underlie language problems, we might expect to see analogous difficulties with vowels that vary in formant frequency.

This expectation was confirmed in the present study. We found deficits in processing pure tones and vowels in the people with SLI. On the behavioural tasks, poor discrimination for pure tones was associated with poor discrimination of vowels in the younger SLI listeners. In addition, there was a strong relationship between pure tone and vowel thresholds in the SLI-young group, which was more robust than the same relationship in the SLI-old, control-young, and control-old groups. In the ERP measures, we saw that younger and older people with SLI had age-inappropriate N1–P2 ERPs to pure tones and vowels compared to controls. This could be taken to suggest that many people with SLI are less like to detect (Hyde, 1997) or switch their attention (Naatanen & Picton, 1987) to changes in the sound environment.

The conjunction of impaired tone and vowel processing in the same groups of people with SLI are in line with Uwer et al. (2002), who found an association between non-speech and speech processing in people with SLI at the level of both behaviour and the brain. Further, the unusually high vowel thresholds of the SLI-young group concur with previous findings of impaired vowel processing in children with SLI (Frumkin & Rapin, 1980; Stark & Heinz, 1996). Finding poor behavioural discrimination thresholds for tones and vowels in younger people with SLI and poor ERPs to tones and vowels in both younger and older people with SLI replicate the results of McArthur and Bishop (2004). This adds weight to the idea that people with SLI have a 4-year delay in auditory maturation. As previously mentioned, this may be detected by the ERPs in both younger and older people with SLI because ERPs reflect maturation changes in processing into adulthood. However, it may not be detected by FD tasks in older people with SLI because FD thresholds may hit ceiling level at 9 years of age, so FD tasks may only be sensitive to a 4-year auditory processing delay in people younger than 14. This idea is supported by the different correlation coefficients between the behavioural thresholds and ERP ICC measures to pure tones, vowels, and non-harmonic complex tones in the younger and older listeners. In young listeners (i.e., SLI and controls combined) there were moderate to strong relationships between behavioural and ERP measures to the sounds. This makes sense if both are valid measures of auditory processing. However, the correlations between the behavioural and ERP measures in the older listeners (i.e., SLI and controls combined) were virtually non-existent. This makes sense if the behavioural thresholds in older listeners had hit ceiling, reducing their validity as measures of auditory function.

Although an auditory maturational account of SLI is consistent with the current pattern of data, alternative interpretations must be kept in mind. For example, the results could be taken as evidence for deviant auditory processing resulting from abnormal neural organisation that is not characteristic of typical development at any age. This would explain why both younger and older people with SLI have abnormal auditory ERPs for their age at Fz. However, it does not so easily explain why poor discrimination thresholds were seen only younger people with SLI, although, it may be that by late adolescence, the brain has had enough time to compensate for its unusual organisation, allowing for normal discrimination thresholds. One way of distinguishing a maturational account from a deviance account would be to compare ERPs at different electrode sites. Previous work on typically developing children has demonstrated a shift in the pattern of activity in temporal, central, and frontal sites with age (Bruneau & Gomot, 1998; Ponton et al., 2000). Because we needed to minimize set-up time to allow for a prolonged session of data acquisition in our ERP study, we focused only on frontal sites, using Fz as the electrode giving the strongest auditory ERP. However, the maturational account could be subjected to stronger test by considering whether the distribution of activity across the brain in those with SLI resembled that of younger children. Another point to bear in mind is that the stimulus presentation conditions were different for the behavioural and electrophysiological sessions, with stimuli being presented against a soft background of video soundtrack in the latter case. Detection of the tones and vowels in the ERP session therefore involved separation of signal from background noise, and it could be that this additional auditory stress made the ERPs more sensitive to mild auditory deficits in the older participants.

To distinguish these possibilities we need to test larger groups of younger and older people with SLI on discrimination tasks that are more sensitive to individual differences in auditory processing at older ages and in varying levels of background noise. We also need to test many more people with normal spoken language skills for their auditory ERPs to produce ERP norms for individual ages. It will be important to consider ICC measures at multiple scalp sites to see the distribution of activity changes with age in people with SLI. This spatial information may lead to some conclusions about the structures underlying abnormal brain responses in the SLI population, once the sources of auditory ERPs are better understood.

4.2. Is poor vowel discrimination related to spectral complexity or phonetic significance?

The pattern of behavioural responses to vowels compared to non-harmonic complex tones and pure tones differed for people with normal vowel discrimination (control-young, control-old, and SLI-old groups) and those with poor vowel discrimination (SLI-young group). The people with normal discrimination had poorer thresholds for vowels than non-harmonic complex tones, and only moderate correlations between their thresholds for vowels and complex tones, suggesting that the phonetic status of the vowels lead them to process vowels somewhat differently to other non-speech sounds. In contrast, the SLI-young group had equally poor discrimination for vowels and non-harmonic complex tones, and strong correlations between their vowel and complex-tone thresholds, suggesting that their poor vowel processing related to problems with processing the spectral information of vowels rather than phonetic status.

4.3. Current issues and recommendations revised

In McArthur and Bishop (2001), we suggested a number of reasons why previous studies of rapid auditory processing in people with SLI have produced conflicting results. In light of the present results and McArthur and Bishop (2004), three of these ideas need revising. First, we suggested that studies finding impaired rapid auditory processing in people with SLI may have used rapid auditory processing tasks that taxed FD to a greater degree than studies that did not find impaired rapid auditory processing in people with SLI. However, it now seems unlikely that FD alone is the root of the problem. The age-inappropriate brain responses of people with SLI to individual sounds have the potential to interfere with the processing of sounds presented individually, sequentially, or simultaneously. Thus, poor performance on rapid auditory processing tasks may not result from poor FD. Instead, poor performance on rapid auditory processing tasks and FD tasks may both be indicators of age-inappropriate brain responses to sounds in general (see also Bishop & McArthur, 2004).

A second explanation that we considered for the conflicting findings of previous auditory processing studies of SLI was that an auditory processing deficit resolves with age. We dismissed this idea because our review of the literature revealed no systematic age differences for studies that did and did not find evidence for impaired processing in people with SLI. However, the present results indicate that age may indeed be an important factor when combined with the age-sensitivity of a psycho-acoustic task. If a task is relatively insensitive to age-related changes in auditory processing (e.g., FD tasks) then performance could asymptote at adult levels at a relatively young age (e.g., 9 years). If listeners are older than that age (e.g., 12–21 years), then the task will be sensitive to immature auditory processing only in the youngest people with SLI, whose immature auditory processing falls below the 9-year-old asymptote level. This means that behavioural studies of auditory processing in SLI need to select psycho-acoustic tasks and stimuli that are sensitive enough to detect abnormal processing in the oldest participants of their samples.

A third reason we suggested for the conflicting results of previous studies of auditory processing in SLI was that only a subgroup of people with SLI have poorer auditory processing scores than controls. The present results continue to support this idea because not all people with SLI had abnormal N1–P2 ERPs to sounds for their age. However, the size of the subgroup may be larger than we thought. Behavioural thresholds in the present experiment and previous studies suggest that 30–40% of people with SLI have poorer thresholds than controls (Heath, Hogben, & Clark, 1999; McArthur & Bishop, 2004; McArthur & Hogben, 2001). However, both the present study and McArthur and Bishop (2004) found immature N1–P2 ERPs to sounds in people with SLI with normal behavioural thresholds (i.e., the SLI-old group) as well as those with poor behavioural FD thresholds (the SLI-young group). This suggests that the subgroup of people with SLI who have abnormal auditory processing for their age may be in the majority rather than the minority.

4.4. Overview and conclusions

We tested younger (around 13 years) and older (around 17 years) people with SLI and controls for their FD thresholds and N1–P2 ERPs to 25- and 250-ms pure tones, vowels, and non-harmonic complex tones. The results showed that people with SLI with poor discrimination thresholds or abnormal auditory N1–P2 ERPs to pure tones had the same for vowels, and that abnormal vowel processing related to the spectral complexity rather than the phonetic significance of isolated vowels. We suggest that age may be a crucial factor in studies of auditory processing in SLI, and that the pattern of behavioural deficits will depend upon the age-sensitivity of a psychoacoustic task.

Acknowledgments

This research was supported by the Wellcome Trust. We would like to Faith Ayre for organising the volunteers for the study, Paul Sutcliffe for his help with data collection, and all our volunteers and their parents for their invaluable contributions to this research.

References

Adams, J., Courchesne, E., Elmasian, R. O., & Lincoln, A. J. (1987). Increased amplitude of the auditory P2 and P3b components in adolescents with developmental dysphasia. In R. Johnson, R.

Parassuraman, & J. W. Rohrbaugh (Eds.), *Current trends in evoked potential research* (pp. 577–583). New York: Elsevier.

Albrecht, R., von Suchodoletz, W., & Uwer, R. (2000). The development of auditory evoked dipole source activity from childhood to adulthood. *Clinical Neurophysiology, 111*, 2268–2276.

Bishop, D. V. M., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language, and Hearing Research, 42*, 1295–1310.

Bishop, D. V. M., & McArthur, G. M. (in press). Individual differences in auditory processing in specific language impairment: A follow-up study using event-related potentials and behavioural thresholds. *Cortex*.

Bishop, D. V. M., & McArthur, G. M. (2004). Immature cortical responses to auditory stimuli in specific language impairment: Evidence from ERPs to rapid tone sequences. *Developmental Science, 7*, F11–F18.

Bruneau, N., & Gomot, M. (1998). Auditory evoked potentials (N1 wave) as indices of cortical development. In B. Garreau (Ed.), *Neuroimaging in child neuropsychiatric disorders* (pp. 113–123). Paris: Springer.

Courchesne, E. (1990). Chronology of postnatal human brain development: Event-related potential, positron emission tomography, myelinogenesis and synaptogenesis studies. In J. W. Rohrbaugh, R. Parasuraman, & R. Johnson (Eds.), *Event-related brain potentials: Basic issues and applications* (pp. 210–241). New York: Oxford University Press.

Courchesne, E., Lincoln, A. J., Yeung-Courchesne, R., Elmasian, R., & Grillon, C. (1989). Pathophysiologic findings in non-retarded autism and receptive developmental language developmental disorder. *Journal of Autism and Developmental Disorders, 19*, 1–17.

Dunn, L. M., Dunn, L. M., Whetton, C., & Pintilie, D. (1982). *British picture vocabulary scale*. Windsor, UK: NFER-Nelson.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin and Review, 2*, 460–493.

Frumkin, B., & Rapin, I. (1980). Perception of vowels and consonant–vowels of varying duration in language impaired children. *Neuropsychologia, 18*, 443–454.

Gathercole, S. (1993). Word learning in language-impaired children. *Child Language Teaching and Therapy, 9*, 187–199.

Heath, S. M., Hogben, J. H., & Clark, C. D. (1999). Auditory temporal processing in disabled readers with and without oral language delay. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 637–647.

Hill, P., Hogben, J. H., & Bishop, D. V. M. in press. Auditory frequency discrimination in children with specific language impairment: A longitudinal study. *Journal of Speech, Language, and Hearing Research*.

Hyde, M. (1997). The N1 response and its applications. *Audiology and Neuro-Otology, 2*, 281–307.

Kewley-Port, D., & Watson, C. S. (1994). Formant-frequency discrimination for isolated English vowels. *Journal of the Acoustical Society of America, 95*, 485–496.

Korpilahti, P. (1995). Auditory discrimination and memory function in SLI children: A comprehensive study with neurophysiological and behavioural methods. *Scandinavian Journal of Logopedics and Phoniatrics, 20*, 131–139.

Leonard, L. (1996). Characterizing specific language impairment: A cross-linguistic perspective. In M. Rice (Ed.), *Toward a genetics of language* (pp. 243–256). Hillsdale, NJ: Lawrence Erlbaum.

Liegeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauval, P. (1994). Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology, 92*, 204–214.

Lincoln, A. J., Courchesne, E., Harms, L., & Allen, M. (1995). Sensory modulation of auditory stimuli in children with autism and receptive developmental language disorder: Event-related brain potential evidence. *Journal of Autism and Developmental Disorders, 25*, 521–539.

Marler, J. A., Champlin, C. A., & Gillam, R. B. (2002). Auditory memory for backward masking signals in children with language impairment. *Psychophysiology, 39*, 767–780.

Mason, S. M., & Mellor, D. H. (1984). Brain-stem middle latency and late cortical evoked potentials in children with speech and language disorders. *Electroencephalography and Clinical Neurophysiology, 59*, 297–309.

McArthur, G. M., & Bishop, D. V. M. (2001). Auditory perceptual processing in people with reading and oral language impairments: Current issues and recommendations. *Dyslexia, 7*, 150–170.

McArthur, G. M., & Bishop, D. V. M. (2004). Which people with specific language impairment have auditory processing deficits?. *Cognitive Neuropsychology, 21*, 79–94.

McArthur, G. M., Bishop, D. V. M., & Proudfoot, M. (2003). Proudfoot. Do video sounds interfere with auditory event-related potentials? *Behavioural Research Methods, Instrumentation, and Computers, 35*, 329–333.

McArthur, G. M., & Hogben, J. H. (2001). Auditory backward recognition masking in children with a specific language impairment and children with a specific reading disability. *Journal of the Acoustical Society of America, 109*, 1092–1100.

Mengler, E. D., Hogben, J. H., Michie, P. T., & Bishop, D. V. M. (in press). Poor frequency discrimination is related to oral language disorder in children: *A psychoacoustic study*. Dyslexia.

Naatanen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of component structure. *Psychophysiology, 24*, 375–425.

Neurosoft, Inc. (1999). *SCAN: User guide*. Virgina, USA: Neurosoft, Inc.

Neville, H. J., Coffey, S. A., Holcomb, P. J., & Tallal, P. (1993). The neurobiology of sensory and language processing in language-impaired children. *Journal of Cognitive Neuroscience, 5*, 235–253.

O'Donnell, B. F., Shenton, M. E., McCarley, R. W., Faux, S. F., Smith, R. S., Salisbury, D. F., et al. (1993). The auditory N2 component in schizophrenia: Relationship to MRI temporal lobe gray matter and to other ERP abnormalities. *Biological Psychiatry, 34*, 26–40.

Ors, M., Lindgren, M., Blennow, G., Nettelbladt, U., Sahlen, B., & Rosen, I. (2002). Auditory event-related brain potentials in children with specific language impairment. *European Journal of Paediatric Neurology, 6*, 47–62.

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: Evidence from multi-channel evoked potentials. *Clinical Neurophysiology, 11*, 220–236.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Standard progressive matrices*. Oxford, UK: Oxford Psychologists' Press.

Rice, M. L., Wexler, K., & Redmond, S. M. (1999). Grammaticality judgements of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 943–961.

Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment. Is there a deficit? What is its nature? Does it explain anything?. *Journal of Phonetics, 31*, 509–527.

Semel, E., Wiig, E. H., & Secord, W. (1987). *Clinical evaluation of language fundamentals-revised.* San Antonio, TX: The Psychological Association.

Stark, R. E., & Heinz, J. M. (1996). Vowel perception in children with and without language impairment. *Journal of Speech and Hearing Research, 39*, 860–869.

Tallal, P. (2000). Experimental studies of language learning impairments: From research to remediation. In D. V. M. B. Bishop & L. Leonard (Eds.), *Speech and language impairments in children* (pp. 131–156). Hove, UK: Psychology Press.

Tallal, P., & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia, 13*, 69–74.

Tallal, P., & Stark, R. E. (1981). Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *Journal of the Acoustical Society of America, 69*, 568–574.

Tallal, P., Stark, R. E., & Curtiss, B. (1976). Relation between speech perception and speech production impairment in children with developmental dysphasia. *Brain and Language, 3*, 305–317.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America, 41*, 782–787.

Thompson, N. C., Cranford, J. L., & Hoyer, E. (1999). Brief-tone frequency discrimination by children. *Journal of Speech, Language, and Hearing Research, 42*, 1061–1068.

Tonnquist-Uhlen, I. (1996). Topography of auditory evoked long-latency potentials in children with severe language impairment: The P2 and N2 components. *Ear and Hearing, 17*, 314–326.

Tonnquist-Uhlen, I., Borg, E., Persson, H. E., & Spens, K. E. (1996). Topography of auditory evoked cortical potentials in children with severe language impairment: The N1 component. *Electroencephalography and Clinical Neurophysiology: Evoked Potentials, 100*, 250–260.

Uwer, R., Albrecht, R., & von Suchodoletz, W. (2002). Automatic processing of tones and speech stimuli in children with specific language impairment. *Developmental Medicine and Child Neurology, 44*, 527–532.

Wiig, E. H., & Secord, W. (1989). *Test of language competence-expanded edition*. USA: The Psychological Corporation.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature, 387*, 176–178.

Wright, B. A., & Zecker, S. G. (2004). Learning Problems, Delayed Development and Puberty. *Proceedings of the National Academy of Sciences, 101*, 9942–9946.