Backward and Simultaneous Masking in Children With Grammatical Specific Language Impairment: No Simple Link Between Auditory and Language Abilities

Article *in* Journal of Speech Language and Hearing Research · March 2009

DOI: 10.1044/1092-4388(2009/08-0114) · Source: PubMed

Heather K J van der Lely
Harvard University

Backward and Simultaneous Masking in Children With Grammatical Specific Language Impairment: No Simple Link Between Auditory and Language Abilities

Stuart Rosen
Alan Adlard
Heather K. J. van der Lely
University College London

Purpose: We investigated claims that specific language impairment (SLI) typically arises from nonspeech auditory deficits by measuring tone-in-noise thresholds in a relatively homogeneous SLI subgroup exhibiting a primary deficit restricted to grammar (Grammatical[G]-SLI).
Method: Fourteen children (mostly teenagers) with G-SLI were compared to age-, vocabulary-, and grammar-matched control children on their abilities to detect a brief tone in quiet and in the presence of a masking noise. The tone occurred either simultaneously with the noise or just preceding it (backward masking). Maskers with and without a spectral notch allowed estimates of frequency selectivity.
Results: Group thresholds for the G-SLI children were never worse than those obtained for younger controls but were higher in both backward and simultaneous masking than in age-matched controls. However, more than half of the G-SLI group (8/14) were within age-appropriate limits for all thresholds. Frequency selectivity in the G-SLI group was normal. Within control and G-SLI groups, no threshold correlated with measures of vocabulary, grammar, or phonology. Nor did the language deficit in the G-SLI children vary with the presence or absence of auditory deficits.
Conclusion: The auditory processing deficits sometimes found in children with SLI appear unlikely to cause or maintain the language impairment.

KEY WORDS: specific language impairment, auditory processing disorder, masking, backward masking, simultaneous masking

Specific language impairment (SLI) is a developmental disorder affecting about 7% of the population, in which language acquisition is impaired in an otherwise apparently typically developing (TD) child (Bishop, 1997; Leonard, 1998). Two general types of explanations have been popular in attempting to account for SLI. The first posits that SLI arises from linguistic deficits (the *domain-specific* view), in particular in systems and processes relating to grammar: syntax, morphology, and possibly phonology, too (Rice, 2003; van der Lely, 2005; van der Lely & Battell, 2003; van der Lely, Rosen, & McClelland, 1998). In contrast to this is a *domain-general* explanation, inspired by earlier work of Efron (1963), that impaired *auditory* processing is the core deficit of SLI (Tallal & Piercy, 1973a, 1973b, 1974, 1975). This view posits that an auditory deficit (especially affecting the perception of rapidly changing or transient sounds) causes impoverished input to the developing language system, leading to impaired speech perception, poor phonological representations and

1092-4388/09/5202-0396

processing, and in turn poor language acquisition. Our focus here will be to evaluate this latter hypothesis—in particular, the extent to which an auditory deficit of a particular kind is evident in teenagers with persisting SLI.

There are at least two main reasons for the recent upsurge of interest in the auditory explanation for SLI. First is the development and application of a computer-run rehabilitation program that claimed remarkable results in ameliorating SLI primarily through improving auditory processing (Merzenich et al., 1996; Tallal et al., 1996). Second is the demonstration of a profound auditory deficit in SLI children using a more-or-less standard psychoacoustic task: backward masking. Wright et al. (1997) investigated the degree of masking obtained for a variety of temporal and spectral relationships between a noise masker and a short probe tone. Most strikingly, although the SLI children were claimed to have normal thresholds for simultaneous masking (when the probe is temporally in the middle of the masker) and forward masking (when the probe follows the masker), they were dramatically impaired in backward masking (when the probe precedes the masker). In this last condition, there was no overlap between the performance of the two groups, with a greater than 40 dB difference in median thresholds.

Not only did backward masking lead to such spectacular differences in thresholds, it also promised a clearer theoretical link between the nonspeech and speech discrimination deficits. If later arriving sounds can perceptually interfere with recently arriving earlier ones, it is easy to see why performance in the tasks of Tallal and Piercy (1973a) would only be impaired for short interstimulus intervals. It also leads naturally to the explanation of why discriminating a */ba/–/da/* contrast (as demonstrated by Tallal & Piercy, 1974) would be difficult, insofar as the following vowel might be expected to exert backward masking on the initial distinctive formant transitions.

However, no other study of backward masking has found such a striking difference in thresholds. Marler, Champlin, and Gillam (2002) found significantly higher backward-masked thresholds in a group of children with language impairment compared to controls but with considerable overlap in obtained thresholds. The only other reports of relatively high thresholds in backward masking in people with developmental language disorders concern dyslexia. Dyslexia too has been claimed to arise from impaired auditory processing (but with no account of why the same auditory deficit leads to dyslexia in some children but SLI in others; Tallal, 1980). Rosen and Manganari (2001) reported a group difference (with overlapping distributions) in backward masking between dyslexic and control teenagers in the absence of differences in forward and simultaneous masking. Montgomery, Morris, Sevcik, and Clarkson (2005) reported a similar result for 7- to 10-year-old dyslexic children and controls. Even stronger evidence against the domain-general hypothesis has been reported. A large study of dyslexic adults found them to have normal performance in backward masking even though they were impaired in other auditory tasks (Ahissar, Protopapas, Reid, & Merzenich, 2000). Particularly relevant to our current focus, Bishop, Carlyon, Deeks, and Bishop (1999) found no significant differences in backward masking performance for 8- to 11-year-old language-impaired twins compared to normally developing control twins matched for age and nonverbal IQ.

At least one other finding of Wright et al. (1997) requires comment. As backward masking has long been claimed to rely much more heavily on central auditory processes than forward or simultaneous masking (Elliott, 1962, 1971; Puleo & Pastore, 1980), an association with language disorders, also presumed to occur from a central deficit, appears plausible. But Wright et al. also reported (without statistical backing) that control children showed a greater difference in thresholds between a broadband and notched noise in simultaneous masking than did the SLI children. Such a result is perplexing because this difference in thresholds is presumed to reflect the operation of a frequency selective mechanism in the cochlea at the very periphery of the auditory system (Rosen, Baker, & Darling, 1998), where SLI children would not be expected to exhibit any disorder. To complicate matters further, this index of frequency selectivity differed less between the two groups under conditions of forward masking, even though the same peripheral frequency analysis is thought to underlie it.

In an attempt to clarify the situation, we measured thresholds in simultaneous and backward masking with broadband and notched noises in TD children and a subgroup of SLI children characterized as having Grammatical(G)-SLI. van der Lely and colleagues have claimed, on the basis of extensive investigations, that G-SLI children have a relatively pure developmental domain-specific persistent deficit in the grammatical components of language—syntax, morphology, and phonology—core to the human language faculty (Marshall, Ebbels, Harris, & van der Lely, 2002; van der Lely, 2005; van der Lely & Battell, 2003). Such children therefore provide a unique testing ground for the claim that auditory deficits underlie all forms of SLI.

We also assessed the language skills of the participants in three main ways. Two of these, for grammar and vocabulary, use standardized tests: the Test of Reception of Grammar (TROG; Bishop, 1983) and the British Picture Vocabulary Scales (BPVS; Dunn, Dunn, Whetton, & Pintilie, 1982). However, the auditory deficit hypothesis implies a more direct link between auditory and phonological skills than with other language skills. Therefore, we also evaluated phonological abilities with a new nonword repetition task that systematically varies phonological complexity (Gallon, Harris, & van der Lely, 2007; van der Lely & Harris, 1999).

In addition to the obvious control group of age- and nonverbal IQ-matched TD children, we also used the fairly standard approach of assessing TD children matched to the G-SLI participants in language performance. In fact, because language skills develop asynchronously in the G-SLI group, we assessed *two* younger TD groups, one matched on grammar and one on vocabulary. Having such an extensive group of TD children also allowed the exploration of the role of auditory skills in language development in TD children, apart from the role they might play in the genesis of SLI.

This study thus allowed us to investigate a number of issues:

1. To what extent do G-SLI listeners exhibit higher thresholds in a backward masking task than do children who are developing language normally?
2. If auditory deficits cause G-SLI, as the domain-general approach claims, then there should be a fairly uniform deficit for the G-SLI group. We should also expect a correlation between the severity of the auditory deficit and the severity of the language problems, especially in phonology.
3. If auditory deficits are associated with, but do not cause G-SLI, then we might expect to see deficits in a subgroup of G-SLI children. No correlations between the severity of the auditory deficit and the severity of the language problems would be further evidence against a causal link, but even significant correlations would not be strong evidence of causality.
4. G-SLI listeners performing better or worse than their language-matched controls, indicating weak links between auditory performance and grammatical development, will be taken as evidence against a strong causal role for auditory processing deficits in the genesis of G-SLI. But equivalent performance is difficult to interpret. It could be that language experience influences auditory processing skills as much as the other way around.
5. G-SLI listeners performing comparably to age-matched controls would rule out any role for a persisting auditory deficit in the maintenance of SLI.

In addition to these general questions, we can also answer more specific questions about the nature of any auditory deficits found: (a) If auditory deficits involve a specific difficulty in temporal analysis, we would expect deficits to be evident for backward but not for simultaneous masking, and (b) if G-SLI listeners have normal peripheral auditory processing, measures of frequency selectivity should not differ among the groups.

Experiment 1: Simultaneous and Backward Masking With a Bandpass Noise

Method

Participants. Four groups of children participated (see Tables 1 and 2). Fourteen mostly teenaged children (mean age of 15;8 [years;months]) previously diagnosed with G-SLI formed the SLI group (van der Lely et al., 1998; van der Lely, Rosen, & Adlard, 2004). All showed a significant impairment (more than 1.5 *SD*s below the mean) on one or more standardized tests of grammar assessing comprehension and expression, and their vocabulary knowledge too was impaired. Furthermore, on two assessments of aspects of grammar core to their deficit (verb tense and agreement and passive sentences), they exhibited an error rate greater than 20% when virtually no errors would be expected after the age of 5 or younger in TD children (van der Lely, 1996a, 1996b, 1999). However, they were of average nonverbal intelligence as measured by Raven's Progressive Matrices and the block design subtest of the British Ability Scales (BAS; Elliott, Murray, & Pearson, 1978). They showed no abnormalities in other aspects of their development nor did any have a hearing impairment.

A further 14 children (mean age of 16;2) served as chronological age- and nonverbal IQ-matched controls (CA). Two groups of younger children developing language normally were matched to the G-SLI participants on different language tests. Because different language components develop asynchronously in SLI children, we adopted a now well-tested strategy of matching on different aspects of language. Consistent with our previous studies, the most pertinent linguistic levels for this disorder are measures of vocabulary, expressive morphology, and syntax (via sentence understanding). The younger, language ability control group (LA1) consisted of 11 children (mean age of 8;2) whose raw scores on two tests of grammar did not differ statistically ($p > .15$) from those of the G-SLI participants (the TROG and the Grammatical Closure subtest of the Illinois Test of Psycholinguistic Abilities [ITPA; Kirk, McCarthy, & Kirk, 1968]). A further 12 children, the LA2 controls (mean age of 9;3), were matched to the G-SLI participants on single word vocabulary raw scores (BPVS, with $p > .8$). Note that in TD children, there is a strong correlation (.60–.79) between vocabulary scores and nonverbal IQ

Table 1. Characteristics of the grammatical specific language impairment (G-SLI) listeners and their measured thresholds.

Listener	Sex	Age	Backward	Simultaneous	In quiet	Block design	BPVS	TROG	TOPhS
AD	F	12.5	82.4	79.4	30.9	93	75	83	30
AT	M	19.0	68.4	76.9	27.4	85	68	82	39
AZ	M	17.1	28.9	76.9	25.4	135	89	98	66
B3	F	19.0	89.9	96.4	31.4	94	58	71	52
BP	M	15.3	56.4	78.9	34.9	113	68	55	63
CG	M	14.6	58.4	83.4	29.4		61	60	35
CP	M	13.0	91.4	92.4	36.9	114	69	69	71
G2	M	15.7	76.9	78.4	25.9	99	65	76	41
JW	M	16.2	88.9	76.9	34.9	127	71	71	
ML	M	12.4	83.9	79.4	34.4	111	68	79	68
MP	M	19.8	71.4	77.4	26.4	79	76	67	79
SI	F	15.9	52.9	79.4	28.9		57	71	82
SM	M	13.0	38.4	76.4	28.9	113	74	69	94
WL	M	16.3	88.9	83.4	33.4	103	80	71	65
G-SLI	11 M/3 F	15.7 (2.5)	69.8 (20.1)	81.1 (6.1)	30.7 (3.8)	106 (16.5)	70 (8.6)	73 (10.5)	60 (19.7)
CA	13 M/1 F	16.2 (1.6)	51.7 (15.4)	76.7 (1.7)	28.4 (3.6)	103 (10.2)	102 (10.0)	109 (18.0)	
LA2	7 M/5 F	9.2 (0.6)	73.2 (10.1)	78.2 (3.8)	33.9 (3.9)		96 (28.6)	102 (27.9)	
LA1	5 M/6 F	8.1 (0.4)	77.0 (10.8)	82.7 (6.8)	35.6 (7.2)		101 (9.6)	102 (12.8)	89 (5.1)

Note. Also shown are the means and standard deviations (in parentheses) from all four groups tested. The scores from the block design, British Picture Vocabulary Scales (BPVS), and Test of Reception of Grammar (TROG) are standardized to the age of the participant, whereas the Test of Phonological Structure (TOPhS) is a raw score. Note that the TOPhS was not applied to the chronological age- and nonverbal IQ-matched control (CA) group, as children of that age are typically at ceiling (with scores of 96). TOPhS scores for the two younger, language ability control groups (statistically indistinguishable at $p > .8$) were combined for a total of 14 listeners (8 LA1 and 6 LA2).

tasks (Dunn et al., 1982; Elliott, 1983). Therefore we can also take the BPVS scores as an approximate measure of these children's nonverbal abilities. Most of the participants in this study (in all four groups) also participated in a separate study of auditory discrimination abilities (van der Lely et al., 2004). Informed consent for all was obtained by written parental permission.

Measures of language ability. Two standardized measures of language ability, the TROG and BPVS, were available for all the participants and so could be used to explore relations between language and auditory skills.

Table 2. Summary statistics for the raw scores from the three language tests used to match listener groups.

	BPVS		TROG		ITPA		
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	Number tested
G-SLI	**78.2**	16.8	**14.4**	1.60	**21.9**	3.88	14
LA1	66.7	11.8	**14.7**	2.28	**24.3**	4.43	11
LA2	**79.8**	15.5	17.3	1.15	27.8	3.79	12
CA	123.9	14.3	18.5	1.34			14

Note. Values used for matching are in bold. ITPA = Illinois Test of Psycholinguistic Abilities.

Due to the special role of phonological processing in the auditory deficits hypothesis, we also assessed participants using a nonword repetition task, as these are known to be sensitive to phonological problems. The Test of Phonological Structure (TOPhS; Gallon et al., 2007; van der Lely & Harris, 1999) is different from other similar tasks in systematically varying the phonological structure of the test items (van der Lely, 2004). Scores are expressed as the number of words correctly repeated out of a total of 96. The TOPhs was administered during the same time period as the auditory tests (within 2–3 months), whereas the two standardized measures of language ability were administered within a prior 6- to 18-month period. Previous test/retest scores using the standardized measures with delays of over 2 years have revealed hardly any change over time for children of this age. Indeed, a recent study of 6 G-SLI children over some 12 years from around 9 years old to adulthood revealed a remarkable consistency on core tests of grammar (such as passive sentences and pronominal reference) with no significant changes in scores over this long time period (Tang, 2004). Such consistency is also evident in our work that has reported results from the same children over many years (see, e.g., van der Lely, 1996a, 1997; van der Lely & Stollwerck, 1997).

Measurements of probe tone thresholds. Thresholds for a short probe tone were determined in three main conditions: in quiet, in simultaneous masking, and in backward masking. The tasks were modeled closely on those described by Wright et al. (1997), with identical stimuli but some differences in the adaptive tracking procedure. Note that better performance is indicated by lower probe tone thresholds.

Measurements were made monaurally in the right ear over Sennheiser HD475 headphones using a two-interval two-alternative forced-choice task. A maximum-likelihood adaptive procedure was used to track 90% correct by varying the level of the 20-ms 1-kHz sinusoidal probe tone. The probe was allowed to change by a maximum of 8.6 dB between trials to allow the listener a few trials at the start of the session in which the probe was clearly audible. For the masking conditions, two 300-ms bursts of masking noise were presented on each trial with a 340-ms interstimulus interval. Masking noises were bandpass (0.6–1.4 kHz) at a spectrum level of 40 dB SPL. The 20-ms probe tone could occur either simultaneously with the masking noise (200 ms after masker onset: simultaneous masking) or with its onset 20 ms prior to the start of the masker (backward masking). Thus, in backward masking, there was no overlap between the probe tone and the masker (nonsimultaneous masking). All stimuli were gated on and off with 10-ms cosine-squared envelopes.

The listener sat in front of a computer monitor, which presented two cartoon faces, side by side. At the start of the trial, both of these had closed mouths. The mouth of the cartoon face on the left side opened and closed in synchrony with the first observation interval, whereas the mouth of the cartoon on the right side opened and closed in synchrony with the second observation interval. The probe tone occurred along with one of the noise bursts. The listener indicated which of the noise bursts was associated with the tone by clicking with a mouse on the appropriate cartoon face. Feedback was given by a "happy" or a "sad" face replacing the face chosen, depending upon the correctness of the response.

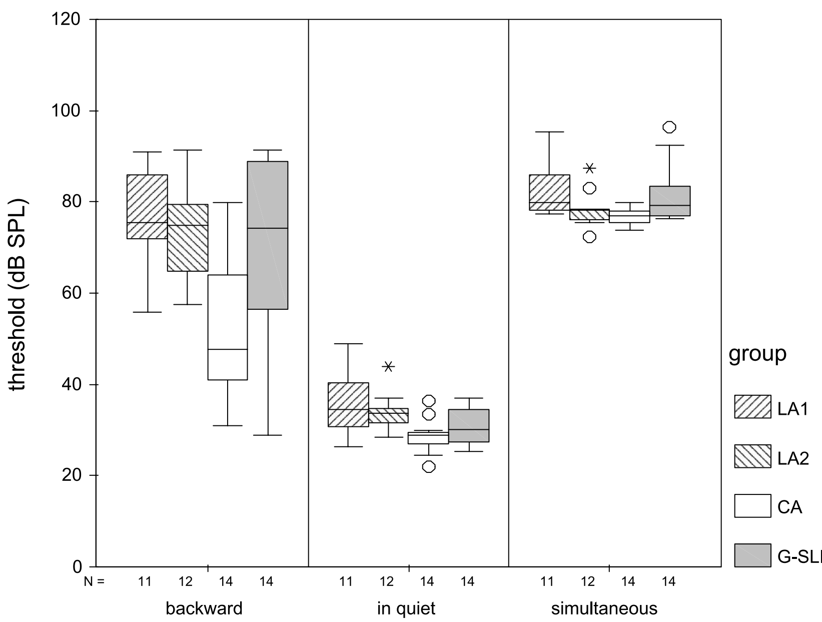
Listeners were first acquainted with the experimental situation by being tested twice with the probe tone in quiet. This provided training for the masking tasks and also established the listener's absolute threshold. Then, two tests of one of the masking conditions (simultaneous or backward) were run, followed by two tests of the other. The order of the two masking conditions was randomized across listeners. Testing for a particular condition was terminated when 2 paired, successive thresholds were within 6 dB. When this criterion was not met, a further 2 thresholds were run until within 6 dB, as long as time was available. Listeners participated in 2–6 threshold determinations per condition, with a mean of 2.5. Most (75%) of the participant/condition combinations required only 2. All tests were performed, with appropriate breaks, in a single session. Results were then summarized by calculating the median for all thresholds obtained. The medians thus calculated were based on 2–6 individual thresholds (after any deletions detailed below), where each threshold was based on 26–36 trials ($M = 27.4$, $SD = 3.1$).

During the course of testing, it became apparent that our version of the maximum-likelihood adaptive technique was sensitive to lack of attention, especially during the beginning of the task. In order to minimize the contribution of outlying thresholds, results from each listener and condition were subject to the following procedure: A particular threshold was excised if it was 6 dB higher than any others in the same condition and (a) an error had been made on the first or second trial or (b) only a single error was made during the test. However, thresholds were only excised if there were more than 2 thresholds available in a particular condition. About 5% of all thresholds were excised in this way. In order to minimize the difficulties associated with listener errors on the first or second trial, a modification to the adaptive procedure was implemented, which ignored errors during these trials. This modified procedure was in place for about the last quarter of testing.

Results

Group differences in performance. Box plots of the thresholds obtained in each condition and for each group can be found in Figure 1. Masking conditions were analyzed separately, both because variability in backward masking tends to be substantially greater than in other conditions and because we were not particularly interested in the extent to which thresholds are different across conditions. (We know, for example, that thresholds will be lowest in quiet.) One-way analyses of variance (ANOVAs) showed significant group differences in all three conditions ($p < .02$). Planned comparisons using an independent samples t test (allowing for unequal variances when necessary) were performed to determine whether the G-SLI group differed from any of the control groups on each of the three conditions. Of the six comparisons between the G-SLI group and the two younger control groups (LA1 and LA2), only one was significant: in quiet, for LA2, $p \approx .04$, with the G-SLI group performing better. Note, though, that the mean threshold for the LA1 group was even poorer than for the LA2 group, but greater variability prevented the difference reaching statistical significance. More importantly, the CA group performed significantly better than the G-SLI group in both simultaneous ($p = .019$) and backward masking ($p = .013$) but not in quiet ($p = .12$). In short, the G-SLI group performed at least as well as the younger controls in all threshold

Figure 1. Box plots of the thresholds obtained in Experiment 1 for each of the four groups of listeners. The box indicates the interquartile range of values obtained, with the median indicated by the solid horizontal line. Whiskers indicate the range of measurements except for points more than 1.5 (indicated by an unfilled circle) or 3 (indicated by an asterisk) box lengths from the upper or lower edge of the box. LA1 and LA2 = the younger, language ability control groups. CA = chronological age- and nonverbal IQ-matched control group; G-SLI = grammatical specific language impairment.



tasks but were inferior performers to the CA group in both simultaneous and backward masking.

To preclude the possibility that differences in the abilities of the different groups of listeners to learn the task were responsible for the differences found, a repeated-measures ANOVA was used to determine the extent to which thresholds changed over the first two measurements, for which all listeners provided data. A 4 × 2 analysis (Group × Threshold) for each condition showed none of the interaction terms to be significant, implying that any changes in threshold for the first two tests were the same for all groups of listeners. There was statistical evidence for overall improvements in threshold in simultaneous masking and in quiet (as evidenced by a significant main effect, $p < .05$), but these changes were small (2.3 and 1.1 dB, respectively). Of course, we cannot say whether extended practice would have resulted in the groups being more (or less) similar, but it seems clear from this analysis that any important differential effects of practice do not seem to be operating over a short time scale.

Individual analyses. It is evident, though, that there is a great deal of overlap among the groups, even for the G-SLI and CA groups. In order to characterize the performance of each individual listener without regard to age, we first quantified how thresholds changed with age in the control listeners. Figure 2 shows the thresholds obtained in each condition as a function of age for the control listeners only, along with the best-fitting lines obtained from linear regressions. All three conditions

Figure 2. Thresholds obtained in Experiment 1 in quiet and under conditions of simultaneous and backward masking, as a function of the age of the listener. Best-fitting straight lines from a linear regression are shown for each condition. Control listeners only.

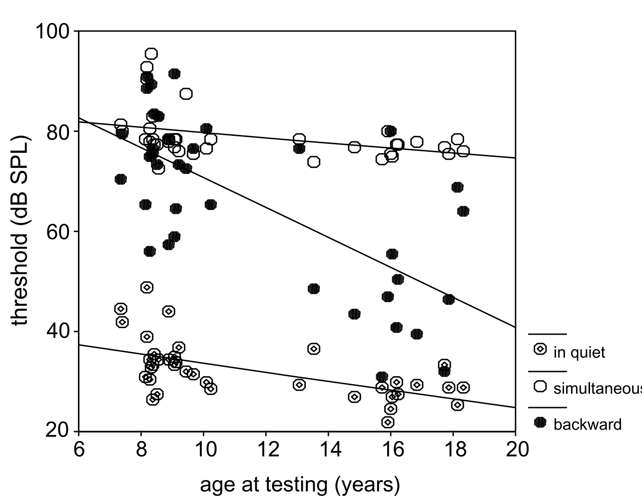


Table 3. Number and percentage of children in each group of listeners who had a standardized residual threshold more than 1.64 *SD*s above the control mean, with age taken into account, for the three main conditions.

Group	Backward		Simultaneous		Backward and simultaneous		In quiet		Number tested
	n	%	*n*	%	*n*	%	*n*	%	
LA1	0	0.0	2	18.2	0	0.0	1	9.1	11
LA2	1	8.3	1	8.3	0	0.0	1	8.3	12
CA	1	7.1	0	0.0	0	0.0	1	7.1	14
G-SLI	5	35.7	4	28.6	3	21.4	2	14.3	14

Note. Also shown is the number of children in each group who performed poorly on both masking tasks. The expected percentage is 5% for the three single conditions.

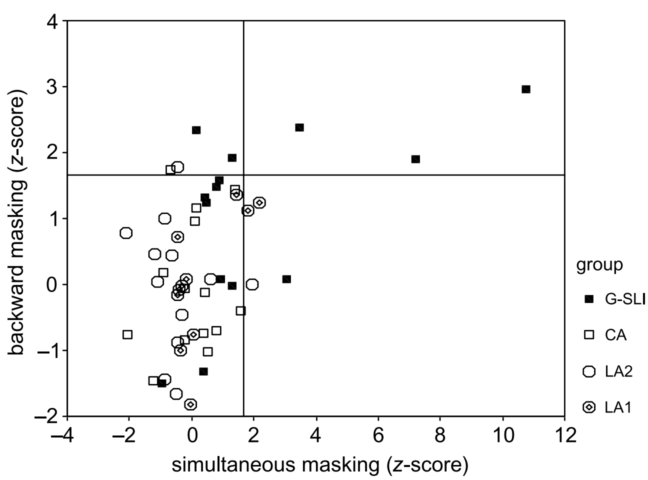
show clear linear decreases in threshold with age (at least to the $p < .02$ level) and no statistical evidence of a quadratic trend.[1] Backward-masked thresholds decreased most with age (3 dB per year), followed by thresholds in quiet (0.9 dB/year) and simultaneous masking (0.5 dB/year). These slopes are reasonably similar to those found in another study when using a similar bandwidth masker (at 3.7, 1.5, and 0.7 dB/year, respectively; Buss, Hall, Grose, & Dev, 1999) and to those of Hartley, Wright, Hogan, and Moore (2000) with identical stimuli and when linear regressions were computed for all listeners aged 6–12 years (4.3 and 1.4 dB/year for backward and simultaneous masking, respectively, excluding one listener with a threshold nearly 4 *SD*s lower than the rest of the group).

These regressions were used to provide an age-corrected z-residual for each listener in each condition by taking the residual of the linear fits and then dividing by the standard deviation of the raw residuals (calculated separately for each of the three control groups). This resulted in a measure of auditory performance with no correlation with age and a mean and standard deviation close to 0 and 1, respectively. Z-residuals were derived for each of the G-SLI listeners on the basis of results from the controls.

Using these z-residuals, we used the conventional cutoff of $z > 1.64$ *SD*s (the poorest performing 5% of the normal population) to calculate the number of listeners in each group exhibiting abnormally poor performance (see Table 3). Although a substantial number of G-SLI children do exhibit impaired auditory processing on both simultaneous and backward masking, the majority do not. In fact, 8 of the 14 (57%) performed within normal limits on all the auditory tasks. A scatter plot of the calculated z-scores for simultaneous and backward masking is shown in Figure 3. Three of the 6 poorly performing G-SLI listeners were poor on both simultaneous and backward masking.

Relationships with measures of language performance. As stated previously, if poor auditory processing is at the root of SLI in these children, we should expect some correspondence between the severity of the auditory and linguistic deficits. In a first look at this possibility, we compared the language performance of the G-SLI listeners with normal masked thresholds ($z \leq 1.64$) to those

Figure 3. A scatter plot of the standardized residuals for simultaneous and backward masking (taking out the effect of age on performance). Reference lines for $z = 1.64$ are drawn on each axis as criteria for abnormally poor performance. A positive z-residual is used because higher thresholds indicate poorer performance.



[1]For the threshold in quiet, one of the LA1 listeners had an extreme outlier T statistic more than twice as big as any other in that condition and more than 40% bigger than any other value in the other two regressions. Cook's D (a measure of the influence of a data point on the estimated regression coefficients) was also large, 70% bigger than any other in all three regressions and more than three times bigger than any in the same condition. Therefore, all the thresholds from this listener were excised from the data set, leaving 37 control listeners. For further details of regression diagnostics, see Cook and Weisberg (1999).

who had poor thresholds ($z > 1.64$). There were no significant differences between the 8 good and 6 poor auditory performers in TROG and BPVS, $t(12) < 1.6$, $p > .14$ for both independent samples t tests, and nearly identical scores for the two groups on the ToPhS, comprising 8 good and 5 poor auditory performers, $t(11) = -0.002$, $p = .998$. It is also interesting to note that these two groups did not differ on block design, our measure of nonverbal IQ: 7 good and 5 poor auditory performers, $t(10) = 0.357$, $p = .73$. (Note that the numbers differ for each test because of missing data; see Table 1.)

Another way to investigate this supposed link between language and auditory skills is through correlations. Within the G-SLI group, neither BPVS, TROG, ToPhS, nor nonverbal IQ correlated with any of the auditory tasks ($p \geq .1$, with $p > .25$ for the crucial ToPhS), but there was a fair degree of correlation within the auditory tasks themselves. The z-residuals for backward masking correlated both with those for simultaneous masking ($r = .54$; $p < .05$) and in quiet ($r = .63$; $p < .02$), whereas the correlation between z-residuals in quiet and for simultaneous masking just missed statistical significance ($r = .51$; $p = .062$).

Finally, we might also expect language and auditory performance to be related in the TD children if the auditory skills measured here are crucial for language development or if component language development affects auditory perception. Bivariate correlations were calculated between all three auditory thresholds (expressed as z-residuals) and the two measures of language ability available for the entire control population (TROG and BPVS). Only one correlation was significant, that between BPVS and thresholds in quiet ($p = .04$, all others $p > .15$), but this arose from a single LA2 listener with undue influence on the regression (indicated by a high value of Cook's D). Excising this single data point makes the regression nonsignificant ($p = .21$). Similarly, for the 14 LA1 and LA2 control listeners with ToPhS scores (one excised for a score more than 3 *SD*s lower than the mean), all correlations with auditory thresholds were nonsignificant ($p > .2$). Thus, there was no relationship within TD children between auditory and language measures. Note, too, that the variability within the control groups seems broadly comparable to what would be expected in the entire population (standard deviations for the BPVS and TROG being 10.5 and 14.1 when 15 is expected), so the lack of correlation did not arise through a lack of variability in the language scores. Perhaps surprisingly, even the auditory measures themselves were uncorrelated ($p > .19$) once the mediating effects of age were accounted for. Nonverbal IQ, too, was uncorrelated with any of the auditory tasks (although only available for the CA group).

Experiment 2: The Effects of a Spectral Notch

As noted previously, there has also been a claim that SLI children suffer from impaired frequency selectivity. In a second testing session, we assessed this claim directly in a smaller number of children by comparing probe thresholds for a bandpass noise with those obtained for a noise with a spectral notch.

Method

Participants. All the participants were drawn from those used in Experiment 1, but fewer were available for this later testing session. The time between the first and second testing sessions varied over periods from a few hours up to 11 months, with a lag of 6–7 months for all the LA2 group. There were 11 children available from each of the G-SLI and LA2 groups but only 3 from the CA group and none from the LA1 group.

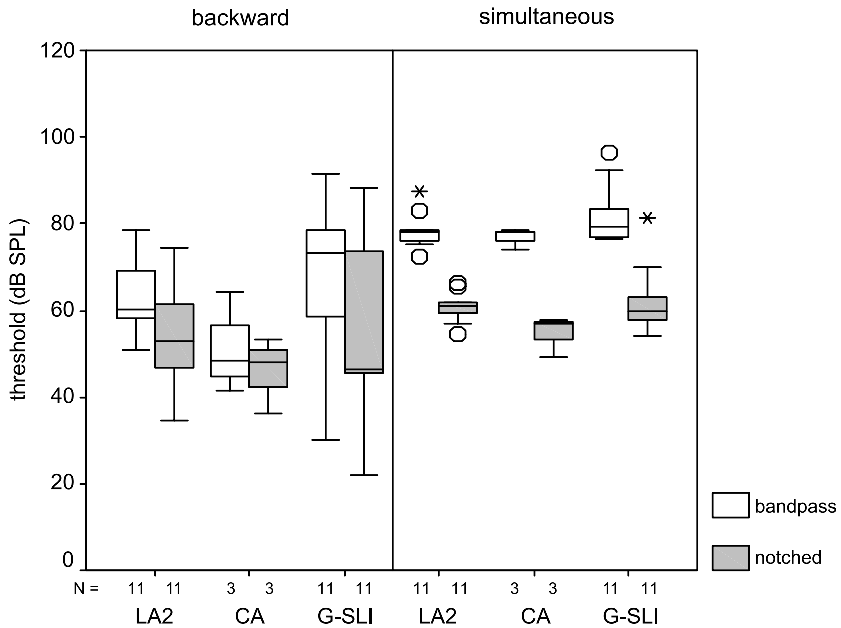
Procedure. The stimulus properties and psychophysical procedures were essentially identical to those described for Experiment 1. The only significant difference was that masking noises could either be either bandpass as before (0.6–1.4 kHz) or notched (0.4–0.8 kHz and 1.2–1.6 kHz).

The testing session began with a measurement of absolute threshold in quiet (mostly to reacquaint the listener with the task), followed by two tests in each of three conditions presented in a randomized order: (a) simultaneous masking with a notched noise, (b) backward masking with a notched noise, and (c) backward masking with a bandpass noise.

Results

Figure 4 compares the thresholds obtained in backward and simultaneous masking with and without a spectral notch. Note that all thresholds depicted were from the second session of testing, except for simultaneous masking with a bandpass noise, which was only run in the first session. Mindful again of the significantly greater variability in backward masking, separate repeated-measures ANOVAs for the effect of a notch in backward and simultaneous masking revealed significant effects of noise type both for backward ($p < .005$) and simultaneous masking ($p < .001$) but no effect involving group. Inspection of the box plots suggests that the release from masking with a notched noise is considerably greater for simultaneous than for backward masking. This was confirmed by a paired-samples t test, collapsed across group, on a measure of selectivity from each masking

Figure 4. Box plots comparing the thresholds obtained in bandpass and notched noises in the conditions of backward and simultaneous masking. Thresholds for simultaneous masking in a bandpass noise were determined in the first testing session, but all other thresholds are from the second testing session.



condition (the difference in thresholds for a bandpass and notched noise). The degree of selectivity differed significantly for the two masking conditions ($p < .001$).

It thus appears that the degree of frequency selectivity, as indexed by the difference in thresholds for the bandpass and notched noise, did not differ among the groups but was greater for simultaneous masking. Clearly, however, the statistical power of the ANOVA was very weak, with only 3 CA listeners. However, further data concerning selectivity in 21 other teenage CA listeners is available from two other studies. As we show in the Discussion, the selectivity evidenced by those 21 listeners is very similar to the results from the 3 listeners reported here.

Comparison With Previous Studies

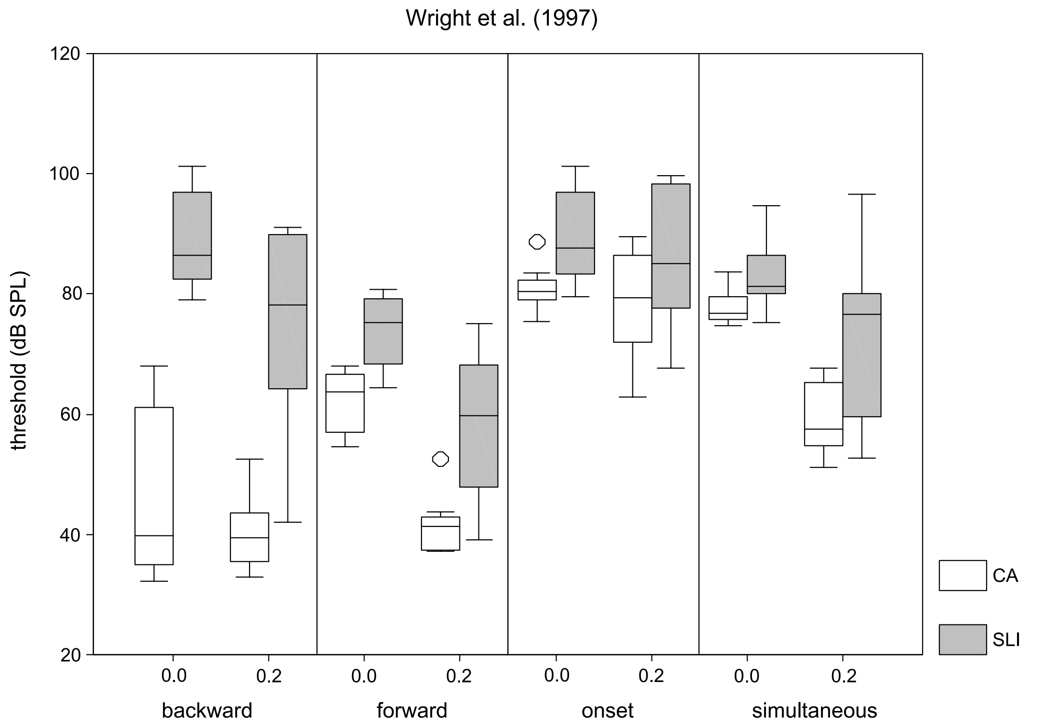
Before going on to discuss the implications of our results, it will be useful to summarize the main findings and compare them to previous reports. First, at least partially in agreement with Wright et al. (1997) but quite different from the findings of Bishop et al. (1999), listeners with SLI had, as a group, higher probe thresholds in backward masking than age-matched controls. Unlike the Wright et al. study, however, in which every SLI child had a higher threshold than every control, there was considerable overlap between our groups. Most G-SLI children performed normally on all the auditory tasks used here. Also in contradiction to Wright et al. is the fact that performance in simultaneous masking distinguished the groups as well as did performance in backward masking.

A Reanalysis of Wright et al.'s (1997) Data

However, a closer look at Wright et al.'s (1997) results suggests that to characterize the deficit displayed by the SLI children as highly specific to backward masking oversimplifies the situation. Figure 5 displays results from Wright et al.'s study, arranged so as to compare most readily, within each condition, the thresholds obtained by the two groups. Note first that the median threshold of the SLI listeners was always higher than that of the controls, in all 8 conditions (4 Masking Configurations × 2 Notches). The extent to which these differences were significant was addressed through a 2 × 2 (Notch × Group) repeated-measures ANOVA[2] for each masker configuration separately (large differences in variability across the conditions preclude a single ANOVA). In no case was the Notch × Group interaction significant ($p > .1$), indicating that the effect of the notch was never different for the two listener groups. The notched noise, however,

[2]Wright et al. (1997) used a standard ANOVA in a 2 × 2 × 4 design (Notch × Group × Condition), taking account neither of the fact that a single listener was assessed in all eight conditions nor of the vast differences in variability across the different cells (e.g., the standard deviation for the CA group for backward masking in a bandpass noise was nearly 5 times that for simultaneous masking).

Figure 5. Box plots of the data obtained by Wright et al. (1997) comparing the performance of SLI and control children in a variety of masking tasks with a short probe tone, in noises without (0.0) and with (0.2) a spectral notch. "Onset" refers to a condition in which the probe tone started simultaneously with the masker onset; in "forward" masking, the probe appeared immediately after the masker. The other two conditions used temporal configurations identical to those used in this study.



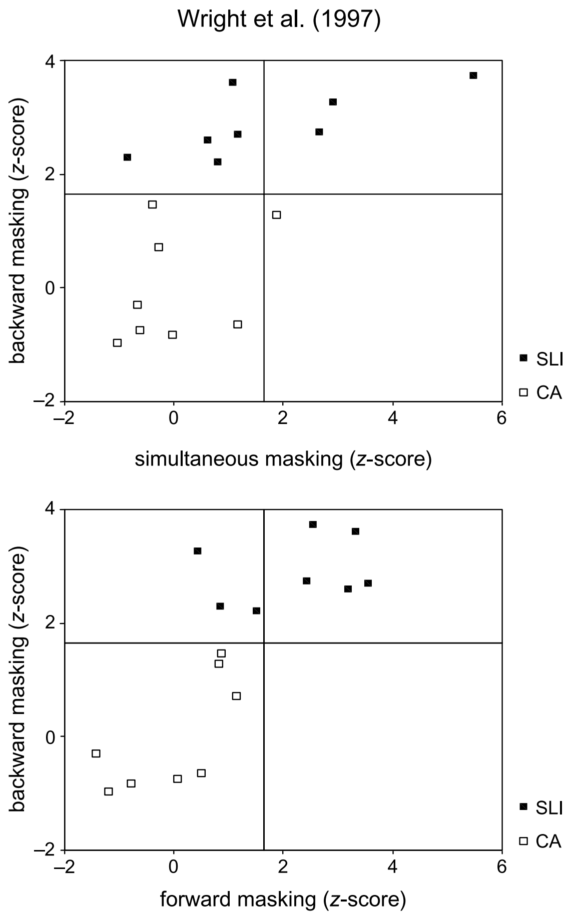
always led to lower thresholds ($p < .01$) except in the onset condition ($p > .1$). Finally, the effect of group was highly significant for forward and backward masking ($p < .001$) and still reached significance for simultaneous masking ($p < .03$). Only for the onset condition was the group difference not significant, but this was nearly so ($p = .054$).

It is also informative to examine the performance of the listeners in Wright et al. (1997) when they are expressed as standardized residuals as calculated from the statistics of the control listeners, in order to classify each listener's performance as impaired or unimpaired (see Figure 6). Again taking a z-score of 1.64 as indicating abnormally poor performance, we see that all SLI, but no CA, listeners were impaired at backward masking. However, although no CA listener was impaired in forward masking, 5 of the SLI listeners were. We would only expect 5% of the population (0.4 of a case on average in a sample this size) to fail this criterion, if the SLI listeners performed similarly to controls. By the binomial test, an outcome this extreme is highly unlikely ($p < 10^{-6}$). Even for simultaneous masking, 3 SLI listeners failed the criterion, an event whose probability is also small ($p < .001$). One CA listener did fail at simultaneous masking, but this outcome is reasonably consistent with a Gaussian distribution of thresholds ($p \approx .06$). The analysis of residuals thus supports the ANOVA-derived claim above that Wright et al.'s SLI listeners had a deficit in more than backward masking. From this point of view, it is perhaps less surprising that in the present study we found simultaneous masking to distinguish G-SLI listeners from their age-matched controls roughly as well as backward masking. But this should not cloud the extent to which backward masking was much more effective in distinguishing SLI children from controls in Wright et al.'s study.

Do High Masked Thresholds Result From Inefficient Processing?

The finding that some SLI listeners may be impaired in more than just backward masking is in agreement with some aspects of the hypothesis put forward by Hartley and Moore (2002). They argued that the deficit associated with language disorders, at least as evidenced in masking studies, is better characterized as poor processing efficiency as opposed to poor temporal acuity. In this view, some listeners with a language disorder require a better signal-to-noise ratio (SNR) for signal detection at the output of a temporal smearing window than is typical. If this were the case, a deficit should be found for all masking situations but will be harder to detect in some situations than in others. The differences between backward,

Figure 6. Scatter plots of the standardized residuals for thresholds in simultaneous, backward, and forward masking for a bandpass noise reported in Wright et al. (1997). Reference lines for $z = 1.64$ are drawn on each axis as criteria for abnormally poor performance.



forward, and simultaneous masking are attributed to changes in the level of the probe signal, which then undergoes different degrees of amplitude compression in the cochlea. Therefore, small differences in efficiency at relatively uncompressed high probe levels (typical for simultaneous masking, with group differences more difficult to detect) are expanded for medium probe levels where compression is greatest (typical for backward masking, where the deficit would appear greater).

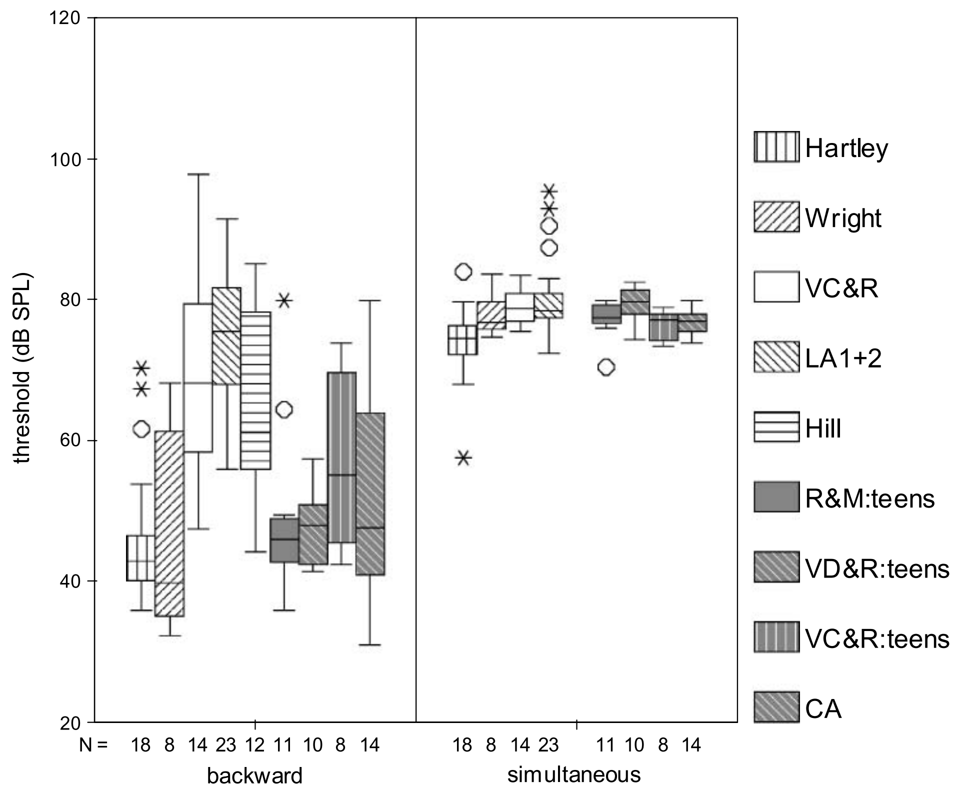
The model of Hartley and Moore (2002) is extremely parsimonious, perhaps too much so, in that it implies that simultaneous, forward, and backward masking are essentially identical, as opposed to the longstanding claims (mentioned previously) that backward masking appears to involve more central auditory processes. For example, it is difficult to see how such an approach could account for the finding that normal adults exhibit very little forward, but significant backward, masking with maskers contralateral to the probe (Elliott, 1962, 1971). On the other hand, the model makes some clear predictions that the performance differences between control and language-disordered listeners can be readily manipulated by changes in the level of the probe. Therefore, by varying the level of the masker noise, or the temporal separation of the probe and signal, it should be possible to reduce the degree of impairment for the language-disordered listeners in backward masking.

Hartley and Moore (2002) have also argued that these same factors may explain the greater degree of variability in backward masking, both within and across listeners. However, this argument cannot be used to support a difference in the statistical reliability of masked thresholds between disordered and normal listener groups across different masker conditions. Because cochlear compression acts as a kind of multiplier of differences in processing efficiency, we would expect not only the absolute difference in threshold across listener groups to vary but also the measurement error. Therefore, we would expect the statistical significance of differences between listener groups to be approximately the same for backward and simultaneous masking, even though the absolute difference was larger for backward masking. Ironically, that is exactly the result reported here, although Hartley and Moore's study was originally aimed at accounting for the claim that significant differences between listener groups are only found in backward masking (as claimed by Wright et al., 1997).

Masked Thresholds in Control Groups Across Studies

As pointed out by Bishop et al. (1999), an important part of the striking difference in backward masking performance between the control and SLI groups in Wright et al. (1997) appears to arise from the low thresholds evidenced by that study's control listeners. Figure 7 compares results across a number of studies that used the same stimulus conditions for simultaneous and backward masking as Wright et al. Note first the high degree of consistency across studies in simultaneous masking but the large differences found in backward masking even for similarly aged children. Although Wright et al.'s 8-year-old children do appear to have very low thresholds, such a result has been replicated by Hartley et al., 2000. However, three studies using essentially identical stimuli have tended to find higher thresholds (i.e., the present study; Hill, Hartley, Glasberg, Moore, & Moore, 2004; and Vanniasegaram, Cohen, & Rosen, 2004). Also worth noting are two further studies, which are not displayed in Figure 7 because their methods differed somewhat from the rest. Montgomery et al. (2005) used identical stimuli to Wright et al. (1997) but a single-interval yes/no procedure to assess 26 TD 7- to 10-year-old children. Thresholds for both backward and simultaneous

Figure 7. Box plots comparing the thresholds obtained in backward and simultaneous masking for typically developing (TD) children in a bandpass noise as obtained in a number of different studies. The five boxes with white backgrounds to the left of each panel represent results from TD children aged 7 to 10.5 years (the middle box is missing for simultaneous masking, as Hill et al. [2004] did not test this condition). The four boxes with light gray backgrounds to the right of each panel represent results from teenagers. The mean age of the children tested varies somewhat from study to study. Hartley = Hartley et al., 2000; Wright = Wright et al., 1997; VC&R = younger children from Vanniasegaram et al., 2004; LA1+2 = the younger control groups from the present study combined; Hill = Hill et al., 2004, backward masking only; R&M:teens = Rosen & Manganari, 2001; VD&R = Vance, Dry, & Rosen, 1999; VC&R:teens = older children from Vanniasegaram et al., 2004; CA = age-matched controls from the present study.



masking were very similar, with a mean of about 70 dB SPL. This is somewhat low for simultaneous masking (so the procedure does not seem to overestimate thresholds) but closely in line for the backward masked thresholds found in the three studies with higher thresholds. Also, Bishop et al. (1999) tested 8- to 11-year-old TD children and found thresholds in backward masking to average 72 dB SPL on first testing but then drop to about 50 dB SPL in subsequent sessions. These too are relatively high compared with results in Wright et al. and Hartley et al., even though Bishop et al. used a masker level 10 dB lower than all the other studies but with an otherwise identical stimulus configuration. Although thresholds in backward masking would not be expected to change linearly with masker level, we would still expect Bishop et al.'s thresholds to be higher had they used a more intense masker. In short, there appear to be two studies (comprising a total of 26 children) that obtained relatively low thresholds for backward masking in children younger than 10 years old and five studies (with a total of about 86 children) that found high thresholds.

Such variability across studies makes comparisons extremely difficult, and it is difficult to know to what to ascribe such differences. Some possible factors are biases in listener recruitment (Bishop et al., 1999) and differences in methodology and calibration. Given that backward masking is so highly variable even within studies, chance may also be crucial especially when test groups are small.

Measures of Frequency Selectivity

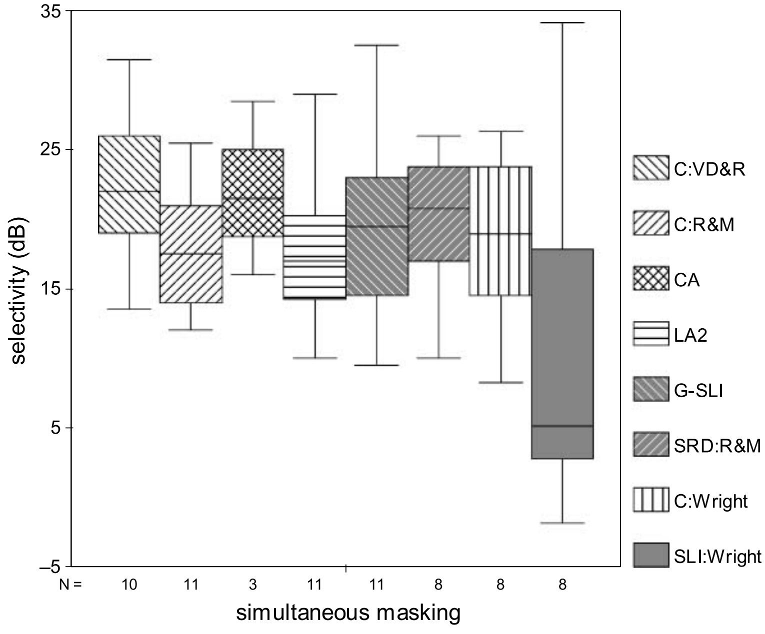
The final crucial difference between the findings of our study and those of Wright et al. (1997) concerns the measure of frequency selectivity as indexed by the difference between thresholds in a bandpass and notched noise. Recall that Wright et al. claimed that their SLI children exhibited reduced selectivity, especially in simultaneous masking. Although our reanalysis of their data failed to reveal any statistical evidence for this (the effect of the notch was not different for any masking condition for the two listener groups), Figure 8 suggests some evidence for their claim. At the same time, it is clear from the only two other studies to have tested language-disordered children with a notched noise (this study and that of dyslexic teenagers by Rosen & Manganari, 2001) that frequency selectivity appears not to differ, on average, from that shown by control listeners. Again, it is difficult to explain this discrepancy from Wright et al.'s results. The discrepancy may arise from the differences in the ages of the SLI children tested, the characteristics of the subgroups tested, or by chance. The lack of statistical significance even in Wright et al.'s data implies, in fact, that there is nothing to explain. Language-disordered children do not differ from controls in the degree of frequency selectivity, a predominantly peripherally determined phenomenon.

Discussion

Perhaps the central theoretical issue for the present study to address concerns the extent to which impaired auditory processing can cause SLI. Researchers sympathetic to this point of view will point to the fact that the G-SLI listeners, as a group, performed more poorly than age-matched controls for both masking conditions (see Table 3 and Figure 1) and are much more likely than the controls as a whole to evidence some auditory deficit. For example, considering only the two masking tasks, the incidence of poor performance (lowest 5 percentile) on one or both of these is some three times the rate in the G-SLI group (6/14 or nearly 43%) as compared with the control population as a whole (5/37 or 13.5%). Furthermore, G-SLI listeners tended to perform similarly to younger language-matched controls, implying some kind of developmental relationship across domains.

But there are many impediments to this simplistic conclusion. First, most of the G-SLI listeners were within normal limits on both masking tasks (8/14). In fact, all 14 G-SLI listeners used here also participated in a study investigating discrimination performance for isolated single formants differing in initial spectral transitions and in a variant of the original Tallal and Piercy (1973a) task (van der Lely et al., 2004). Of these, 4 were within normal limits on those auditory tasks in addition to the thresholds measured here. Therefore, nearly 30% of the G-SLI listeners exhibited no signs of an auditory deficit, even using a number of tasks that have been proven to

Figure 8. Box plots comparing the degree of frequency selectivity (the difference in masked thresholds for a bandpass and notched noise) obtained in simultaneous masking from a number of different studies. Higher values imply greater selectivity. TD children are displayed using boxes with white backgrounds and "C:" in the legend; children with language disorders are portrayed with shaded boxes. SRD:R&M = children with specific reading difficulty/dyslexia from Rosen and Manganari (2001). Legend notation is otherwise as in Figure 7. Note that SLI children from Wright et al. (1997) appear to have somewhat lower values than those found otherwise.



distinguish language-disordered from control listeners in other studies. Furthermore, G-SLI listeners with good and bad auditory performance did not differ in their language abilities, including phonology, nor were there any correlations between auditory and language skills in the G-SLI group; therefore, poor auditory performance does not even appear to exacerbate SLI. It is difficult to see how a theory that posits an auditory deficit as the cause of SLI can handle this complexity unless it is argued that the SLI listeners are of different types. Given that the G-SLI listeners here constitute a much more homogeneous group with respect to their grammar and nonverbal abilities than is typically the case, the claim of a heterogeneous disorder seems harder to defend. Therefore, auditory deficits appear to be associated with SLI but not to cause it.

At least as strong evidence against a proposed auditory cause for language impairment comes from the control listeners, from whom no evidence for a relationship between auditory and language skills was found. That backward masking specifically cannot measure an auditory ability crucial for the development of language is supported by the fact that performance in this task appears to improve after age 10, long after the fundamentals of grammar are meant to be fully developed (Hartley et al., 2000). Some control listeners also failed one or more auditory tasks yet exhibited normal language skills. Although at a much reduced incidence, some 9 out of 37 (about 24%) control listeners who participated in both the masking and nonspeech discrimination tasks referred to previously (van der Lely et al., 2004) performed within the lowest 5 percentile for at least one of the four tasks. Thus, as Bishop et al. (1999) have claimed, an auditory deficit seems neither sufficient nor necessary for a language impairment.

Supposing for the moment that the auditory deficits found here are crucial to the development of language, the exact nature of the deficit is also of great importance. As is well known, Tallal and colleagues have claimed the primacy of a deficit in processing of temporal features, but there is no support for that stance here nor in much of the rest of the literature (for a review, see Rosen, 2003). In fact, the results here, which show simultaneous as well as backward masking to be impaired in some G-SLI listeners, can be seen to support the notion of Hartley and Moore (2002). They claim that, insofar as an auditory deficit can be found in some SLI listeners, it is better characterized as an impairment in "processing efficiency" as opposed to temporal analysis. The relation, however, between this narrowly defined notion of processing efficiency and language is far from clear. Furthermore, it must be stressed that Hartley and Moore's model only attempts to characterize any deficit that might exist but does not address the issue of why many SLI listeners appear to have normal auditory processing, whereas some people with normal language have a deficit. Nor does it resolve claims about the extent to which this particular deficit can or cannot account for the language disorder. Perhaps less controversial is the clear evidence that frequency selectivity is not associated with language disorders, a claim consistent with the understanding that this is primarily a peripheral cochlear function, which is not expected to be deficient in SLI listeners.

From our standpoint, among other evidence, the wide variability in auditory skills in the G-SLI group in the face of the relative homogeneity of the language disorder makes the auditory deficit an unlikely explanation for the language problem. The situation appears to be exactly the same in dyslexia. Impaired auditory processing appears to be more common in dyslexics, but far from universal, and causally unrelated to it (Ramus, 2003; Rosen, 2003).

It could, of course, be argued that an auditory deficit is a crucial factor in language development only for younger children than those studied here, perhaps only in the period when language is developing rapidly, within the first 3 years of life. Clearly, this study cannot address that possibility directly. Even assuming for the moment that such a claim is true, why is it that the auditory deficit resolves in some and not in others? Furthermore, it is hardly consistent with the notion that a language deficit can be remediated primarily through remediation of the auditory deficit (Merzenich et al., 1996; Tallal et al., 1996). In fact, a recent large randomized controlled trial of such auditory training found it to be ineffective in remediating SLI (Cohen et al., 2005), as would be consistent with the notion that auditory deficits are not a crucial factor in its genesis.

There is also the distinct possibility that the associations between deficits in language and audition flow the other way, with impaired language processing impacting on the development of nonspeech auditory abilities. The effect of language component learning on auditory perception in people with language disorders has hardly begun. The effect of language component learning on auditory perception in people with language disorders has hardly begun to be studied.

Conclusions

Masked thresholds were measured in a group of teenagers with G-SLI, displaying a relatively homogeneous pattern of language deficits. The teenagers with G-SLI, as a group, were found to exhibit higher thresholds in backward masking than an age-matched control group. Although this might be thought to support the notion of a temporal processing deficit as being an important factor in the genesis of SLI, other aspects of our results prevent such a simple conclusion. First, higher thresholds were also found for simultaneous masking, so the auditory deficit cannot be characterized as a temporal one. At least as importantly, 60% of the G-SLI listeners performed within normal limits on both tasks. Furthermore, within control and G-SLI groups, no auditory threshold correlated with measures of vocabulary, grammar, or phonology. Therefore, the auditory processing deficits we found in some children with SLI appear unlikely to cause or maintain their language impairment.

Acknowledgments

This study and the second and third authors were supported by The Wellcome Trust Grants 063713 and 044179/Z/95 to the third author, with further support provided by Grant 046823/Z/96 to the first author. We thank the teachers, therapists, and children at the following schools for their help and cooperation: Moor House School, Hurst Green; Dawn House School, Nottingham; Thornhill Primary School, London; The London Oratory School, London; and Nower Hill High School, Pinner, Middlesex. Much appreciation is expressed to Bev Wright, Doug Hartley, Jeff Marler, and Penny Hill, who graciously supplied their original data for reanalysis.

References

Ahissar, M., Protopapas, A., Reid, M., & Merzenich, M. M. (2000). Auditory processing parallels reading abilities in adults. *Proceedings of the National Academy of Sciences of the United States of America, 97,* 6832–6837.

Bishop, D. V. M. (1983). *Test of Reception of Grammar*. Manchester, UK: Manchester University.

Bishop, D. V. M. (1997). *Uncommon understanding: Development and disorders of language comprehension in children*. Hove, UK: Psychology Press.

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.

Buss, E., Hall, J. W., Grose, J. H., & Dev, M. B. (1999). Development of adult-like performance in backward, simultaneous, and forward masking. *Journal of Speech, Language, and Hearing Research, 42,* 844–849.

Cohen, W., Hodson, A., O'Hara, A., Boyle, J., Durram, T., et al. (2005). Effects of computer-based intervention through acoustically modified speech (Fast ForWord) in severe mixed receptive-expressive language impairment: Outcomes from a randomized controlled trial. *Journal of Speech, Language, and Hearing Research, 48,* 715–729.

Cook, R. D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. New York: Wiley.

Dunn, L., Dunn, L., Whetton, C., & Pintilie, D. (1982). *The British Picture Vocabulary Scales*. Windsor, UK: nferNelson.

Efron, R. (1963). Temporal perception, aphasia and déjà vu. *Brain, 86,* 403–424.

Elliott, C., Murray, D., & Pearson, L. (1978). *British Ability Scales*. Windsor, UK: nferNelson.

Elliott, L. L. (1962). Backward masking: Monotic and dichotic conditions. *The Journal of the Acoustical Society of America, 34,* 1108–1115.

Elliott, L. L. (1971). Backward and forward masking. *Audiology, 10,* 65–76.

Gallon, N., Harris, J., & van der Lely, H. (2007). Non-word repetition: An investigation of phonological complexity in children with grammatical SLI. *Clinical Linguistics & Phonetics, 21,* 435–455.

Hartley, D. E. H., & Moore, D. R. (2002). Auditory processing efficiency deficits in children with developmental language impairments. *The Journal of the Acoustical Society of America, 112,* 2962–2966.

Hartley, D. E. H., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *Journal of Speech, Language, and Hearing Research, 43,* 1402–1415.

Hill, P. R., Hartley, D. E. H., Glasberg, B. R., Moore, B. C. J., & Moore, D. R. (2004). Auditory processing efficiency and temporal resolution in children and adults. *Journal of Speech, Language, and Hearing Research, 47,* 1022–1029.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1968). *Illinois Test of Psycholinguistic Abilities*. Urbana, IL: University Press.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

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.

Marshall, C., Ebbels, S., Harris, J., & van der Lely, H. (2002). Investigating the impact of prosodic complexity on the speech of children with specific language impairment. *UCL Working Papers in Linguistics, 14,* 43–66.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996, January 5). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271,* 77–81.

Montgomery, C. R., Morris, R. D., Sevcik, R. A., & Clarkson, M. G. (2005). Auditory backward masking deficits in children with reading disabilities. *Brain and Language, 95,* 450–456.

Puleo, J. S., & Pastore, R. E. (1980). Contralateral cueing effects in backward masking. *The Journal of the Acoustical Society of America, 67,* 947–951.

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13,* 212–218.

Rice, M. L. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 63–95). Mahwah, NJ: Erlbaum.

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.

Rosen, S., Baker, R. J., & Darling, A. M. (1998). Auditory filter nonlinearity at 2 kHz in normal hearing listeners. *The Journal of the Acoustical Society of America, 103,* 2539–2550.

Rosen, S., & Manganari, E. (2001). Is there a relationship between speech and nonspeech auditory processing in children with dyslexia? *Journal of Speech, Language, and Hearing Research, 44,* 720–736.

Tallal, P. (1980). Auditory temporal perception, phonics and reading disabilities in children. *Brain and Language, 9,* 182–198.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X. Q., et al. (1996, January 5). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271,* 81–84.

Tallal, P., & Piercy, M. (1973a, February 16). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature, 241,* 468–469.

Tallal, P., & Piercy, M. (1973b). Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia, 11,* 389–398.

Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia, 12,* 83–94.

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

Tang, S. (2004). *Linguistic profiles in grammatical–specific language impairment (G-SLI) in adults*. Unpublished master's thesis, University College London.

van der Lely, H. K. J. (1996a). Specifically language impaired and normally developing children: Verbal passive vs adjectival passive sentence interpretation. *Lingua, 98,* 243–272.

van der Lely, H. K. J. (1996b). *The Test of Active and Passive Sentences (TAPS)*. Available from www.dldcn.com.

van der Lely, H. K. J. (1999). *Verb Agreement and Tense Test (VATT)*. Available from www.dldcn.com.

van der Lely, H. K. J. (1997). Narrative discourse in grammatical specific language impaired children: A modular language deficit? *Journal of Child Language, 24,* 221–256.

van der Lely, H. K. J. (2004). Evidence for and implications of a domain-specific grammatical deficit. In L. Jenkins (Ed.), *The genetics of language* (pp. 117–144). Oxford, UK: Elsevier.

van der Lely, H. K. J. (2005). Domain-specific cognitive systems: Insight from grammatical specific language impairment. *Trends in Cognitive Sciences, 9,* 53–59.

van der Lely, H. K. J., & Battell, J. (2003). WH-movement in children with grammatical SLI: A test of the RDDR hypothesis. *Language, 79,* 153–181.

van der Lely, H. K. J., & Harris, J. (1999). *The Test of Phonological Structure (TOPhS)*. Available from www.dldcn.com.

van der Lely, H. K. J., Rosen, S., & Adlard, A. (2004). Grammatical language impairment and the specificity of cognitive domains: Relations between auditory and language abilities. *Cognition, 94,* 167–183.

van der Lely, H. K. J., Rosen, S., & McClelland, A. (1998). Evidence for a grammar-specific deficit in children. *Current Biology, 8,* 1253–1258.

van der Lely, H. K. J., & Stollwerck, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition, 62,* 245–290.

Vance, M., Dry, S., & Rosen, S. (1999). Auditory processing deficits in a teenager with Landau-Kleffner syndrome. *Neurocase, 5,* 545–554.

Vanniasegaram, I., Cohen, M., & Rosen, S. (2004). Evaluation of selected auditory tests in school-aged children suspected of auditory processing disorder (APD). *Ear and Hearing, 25,* 586–597.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997, May 8). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature, 387,* 176–178.

Received June 16, 2008

Accepted August 13, 2008

DOI: 10.1044/1092-4388(2009/08-0114)

Contact authors: Stuart Rosen, Speech, Hearing and Phonetic Sciences, University College London, 2 Wakefield Street, London WC1N 1PF, England. E-mail: stuart@phon.ucl.ac.uk.

Heather K. J. van der Lely, Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, 2 Wakefield Street, London WC1N 1PF, England. E-mail: h.vanderlely@dldcn.org.