Deficits in speech perception predict language learning impairment

Johannes C. Ziegler*†, Catherine Pech-Georgel‡, Florence George‡, F.-Xavier Alario*, and Christian Lorenzi§

*Centre National de la Recherche Scientifique, Université de Provence, 13331 Marseille, France; ‡Centre Hospitalier Universitaire de la Timone, Université de la Méditerranée, 13005 Marseille, France; and §Centre National de la Recherche Scientifique, Université Paris 5, Institut Universitaire de France, 75005 Paris, France

Edited by Edward E. Smith, Columbia University, New York, NY, and approved August 10, 2005 (received for review May 31, 2005)

Specific language impairment (SLI) is one of the most common childhood disorders, affecting 7% of children. These children experience difficulties in understanding and producing spoken language despite normal intelligence, normal hearing, and normal opportunities to learn language. The causes of SLI are still hotly debated, ranging from nonlinguistic deficits in auditory perception to high-level deficits in grammar. Here, we show that children with SLI have poorer-than-normal consonant identification when measured in ecologically valid conditions of stationary or fluctuating masking noise. The deficits persisted even in comparison with a younger group of normally developing children who were matched for language skills. This finding points to a fundamental deficit. Information transmission of all phonetic features (voicing, place, and manner) was impaired, although the deficits were strongest for voicing (e.g., difference between/b/and/p/). Children with SLI experienced perfectly normal ''release from masking'' (better identification in fluctuating than in stationary noise), which indicates a central deficit in feature extraction rather than deficits in low-level, temporal, and spectral auditory capacities. We further showed that speech identification in noise predicted language impairment to a great extent within the group of children with SLI and across all participants. Previous research might have underestimated this important link, possibly because speech perception has typically been investigated in optimal listening conditions using non-speech material. The present study suggests that children with SLI learn language deviantly because they inefficiently extract and manipulate speech features, in particular, voicing. This result offers new directions for the fast diagnosis and remediation of SLI.

phonetic deficit | auditory deficit | speech intelligibility | masking noise | specific language impairment

Many children experience unexpected difficulties in understanding and producing spoken language despite normal intelligence, normal hearing, normal opportunities to learn language, and in the absence of any obvious neurological problems (for review, see ref. 1). This disorder is typically called specific language impairment (SLI). In the past years, research on SLI has experienced a growth of interest partially because it has become clear that more children than initially thought show language learning difficulties. Indeed, recent epidemiological studies estimate the incidence of SLI to be ≈7.4% in a population of monolingual English-speaking kindergarten children (2). Children with SLI exhibit deficits in several aspects of language, including phonology, morphology, and syntax (1, 3). One of the hallmarks of SLI is a deficit in the use of function morphemes (e.g., *the*, *a*, and *is*) and other grammatical morphology (e.g., plural *-s*, past tense *-ed*). Children with SLI also are at high risk for subsequent literacy problems (4).

The causes of SLI are still hotly debated. Current theories of SLI fall into two categories: those that attribute SLI to a specifically linguistic deficit and those that attribute SLI to general processing limitations (for a review, see ref. 5). Linguistic deficit theories typically assume that children with SLI have difficulty acquiring linguistic mechanisms, such as past tense rules or the grammatical principle of inflection (6, 7). Children with SLI are thought to be "stuck" at an early stage of grammatical development. Such a delay could actually reflect a general maturational delay of language and other cognitive systems (8, 9).

In contrast, general processing deficit theories assume that it is not the specific nature of the material that is important but rather how it is processed in the brain. Nonlinguistic deficits in either perception or memory are thought to be responsible for language disorder (10–12). The most prominent theory of this kind, also called the fast temporal-processing deficit hypothesis, maintains that SLI is a consequence of a deficit in processing brief and/or rapidly changing auditory information and/or in remembering the temporal order of auditory information (13–16). For example, Tallal and Piercy (13) found that some children with SLI have difficulty reporting the order of pairs of high- and low-frequency sounds when these sounds are brief in duration and presented rapidly. Such a deficit may underlie difficulties in perceiving grammatical forms (e.g., *the* or *is*), which are generally brief and unstressed (17).

The auditory deficit account has been criticized because many children with SLI perform normally on a variety of auditory tasks (18) and because auditory deficits do not predict much of the variance within the group of language-impaired children (19). Also, a number of studies have shown that auditory deficits are not restricted to rapid auditory processing (20). To the contrary, many "slow" tasks, such as 4-Hz amplitude modulation or 2-Hz frequency modulation detection, seem to be difficult for children with developmental language learning disorders (21–25). Given that speech intelligibility depends heavily on the integrity of its low-frequency amplitude modulations (e.g., ref 26), a slow temporal-processing deficit might offer a viable explanation of speech perception deficits in SLI.

The goal of the present study was to investigate the acoustic/phonetic nature of potential speech perception deficits in children with SLI. Our aim was to increase the power of detecting speech perception deficits in SLI by shifting the focus of attention from purely nonlinguistic auditory tasks to speech identification of phonetic categories in ecologically valid listening conditions. In the present study, we used a psychophysical technique testing for consonant identification in the presence of masking noise. This technique parallels the standard tone-detection-in-noise tasks that have proved extremely successful in the study of SLI (9, 27). Speech identification of vowel–consonant–vowel (VCV) stimuli (e.g.,/aba/, /aga/, and /ada/)

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: SLI, specific language impairment; A-match, age-matched; L-match, language-matched; VCV, vowel–consonant–vowel; AM, amplitude-modulated; IQ, intelligence quotient.

[†] To whom correspondence should be addressed at: Laboratoire de Psychologie Cognitive, Centre National de la Recherche Scientifique, Université de Provence, Case 66, 3 Place Victor Hugo, 13331 Marseille Cedex 3, France. E-mail: ziegler@up.univ-mrs.fr.

Table 1. Characteristics of children with SLI and of A-match and L-match controls

Group	Age (range), yr	IQ-P (range)	Comp.	Working mem.	Vocab.	Phonol.
SLI	10.4 (8.3–12.5)	99.4 (85–110)	84.7	28.0	50.6	77.3
L-match	8.6 (7.9–9.6)	102.1 (85–129)	87.8	32.0	58.8	92.0
A-match	10.6 (8.6–12.5)	97.0 (83–110)	95.6	56.0	72.4	99.3
Statistical tests						
L-match	$P < 0.01$	ns	ns	ns	ns	$P < 0.08$
A-match	ns	ns	$P < 0.0001$	$P < 0.01$	$P < 0.0001$	$P < 0.001$

Values for comprehension (Comp.) indicate the percent correct on the ECOSSE picture/word comprehension test (36). Values for working memory (mem.) and vocabulary (Vocab.) indicate percent correct on the L2MA language battery (37). Phonology (Phonol.) values indicate percent correct on a word repitition test taken from the L2MA language battery (37). IQ-P, performance IQ (35); ns, not significant.

was measured in optimal conditions (silence) and in conditions of masking noise.

Two features of the present experimental design must be highlighted: First, previous studies investigated phonetic categorization performance only for a limited number of phonetic contrasts (28, 29). In contrast, in the present study, identification performance was studied for all of the 16 consonantal categories of French. This technique allowed us to investigate more systematically the reception of speech consonant/phonetic features, such as voicing, manner, and place of articulation. Secondly, two types of masking noise were used: temporally fluctuating noise and stationary noise. In conditions of temporally fluctuating noise, unimpaired listeners experience "release from masking," that is, better speech identification in fluctuating noise than in stationary noise (30). This effect indicates that the normal auditory system is capable of taking advantage of relatively short temporal minima in the fluctuating background to detect speech cues, a capacity often called "listening in the valleys." Clearly, this capacity requires a certain degree of temporal and spectral resolution (e.g., ref. 31). Temporal resolution is required to follow the background fluctuations to extract speech cues during the background valleys, whereas spectral resolution is required to access parts of the speech spectrum that are not (or less) masked by the background noise. Temporal fine structure cues (amplitude fluctuations faster than ≈500 Hz) also play a role in the perceptual segregation of speech from background noise (32). Indeed, listeners with sensorineural hearing loss after cochlear damage show degraded spectral resolution and poor fine-structure coding. Not surprisingly, such listeners typically show reduced or abolished masking release (31, 33, 34).

In summary, in Exp. 1, we tested speech identification and the transmission of phonetic features in quiet and noisy conditions. The comparison of performance in fluctuating noise and stationary noise conditions allowed us to test whether children with SLI showed abnormal masking release. Two control groups were tested: one that was matched in terms of chronological age and nonverbal ability and another that was matched in terms of language ability (i.e., younger, typically developing children). The purpose of this second group was to control for the possibility that the ability to do a speech perception task under taxing conditions might be influenced by top-down knowledge of language. If so, potential speech perception deficits could be a consequence rather than a cause of poor language. In contrast, if speech perception deficits persist in the L-match comparison, then they are likely to be the cause rather than the consequence of SLI.

Experiment 1

Methods. ***Participants.*** Ten children (seven boys) with SLI were recruited from the neuropediatric service of the La Timone Hospital in Marseille, France. The children were diagnosed as language-impaired by a multidisciplinary team. Diagnosis included a medical assessment (hearing and vision) and neuropsychological and psycholinguistic testing. All children had a nonverbal intelligence quotient (IQ) >85 on the French version of Wechsler Intelligence Scale for Children (35). They had audiometric thresholds of <20 dB hearing level between 0.25 and 8 kHz and no history of hearing difficulty. None of the children had suffered from otitis media, and none of them showed evidence of seizure or brain lesions. Language comprehension and syntactic knowledge were tested with L'ECOSSE (Épreuve de Compréhension Syntaxico–Sémantique), a standardized sentence/picture matching test (36). All other tests were taken from L2MA, a standardized test of spoken language ability (37): Verbal memory was assessed with backward and forward number repetition, vocabulary was assessed with picture naming, and phonology was assessed with word repetition and phonetic fluency measures. Impaired children in this study were at least 2 SD below the age-appropriate mean on at least four subtests. The characteristics of the participants and a summary of the language test results are found in Table 1.

Two control groups with 10 children each were recruited from a local school in Marseille. The first group, the age-matched (A-match) controls, were selected such that the normally developing children had the same chronological age and nonverbal cognitive ability (performance IQ) as children with SLI (see Table 1). The second group, the language-matched (L-match) controls, were selected such that the normally developing children had the same overall language ability. These children were, on average, ≈2 years younger than children with SLI.

Stimuli. One set of 48 unprocessed VCV stimuli was recorded. These speech stimuli consisted of three exemplars of 16 possible /a*C*a/utterances (*C* =/p,t,k,b,d,g,f,s,ʃ,m,n,r,l,v,z,ʒ/) read by a French female speaker in a quiet environment. Each signal was digitized by a 16-bit analog/digital converter at a 44.1-kHz sampling frequency. VCV identification was assessed in silence or noise. In the latter condition, a gated speech-shaped noise masker (i.e., a noise with the long-term power spectrum of running speech) was added to each utterance (and refreshed in each trial of a given session). This speech-shaped noise was either (*i*) steady (i.e., unmodulated) or (*ii*) modulated with a sine-wave modulator. The expression describing the sine-wave modulator, $m(t)$, was

$$m(t) = [1 + m\sin(2\pi f_m t + \phi)]n(t), \qquad [1]$$

where $n(t)$ represents the speech-shaped noise. Modulation depth, m, was fixed at 1 (i.e., 100%); modulation frequency, f_m, was fixed at 32 Hz. The starting phase of the first-order modulation, ϕ, was randomized between 0° and 360° on each trial.

In each experimental condition, the noise masker was added to each speech utterance at a 0 dB (rms) signal-to-noise ratio. This signal-to-noise ratio was determined in a preliminary experiment so as to yield a consonant identification performance of ≈50% correct when the speech-shaped noise was steady in

Table 2. Speech identification performance (% correct) of children with SLI and for L-match and A-match controls in silence, fluctuating noise, and stationary noise

Group	Silence	Fluct. noise	Stat. noise	Masking rel.
SLI	94.6	72.1	61.7	10.4
L-match	96.4	90.8	82.5	8.3
A-match	99.4	94.4	86.3	8.1
Statistical tests				
L-match	ns ($P > 0.50$)	$P < 0.01$	$P < 0.02$	ns ($P > 0.70$)
A-match	$P < 0.05$	$P < 0.001$	$P < 0.001$	ns ($P > 0.60$)

Masking release is the difference between identification performance in fluctuating and stationary noise. Fluct., fluctuating; Stat., stationary; rel., release; ns, not significant.

control children. In each utterance, signal and noise were of identical duration (mean duration = 648 ms; SD = 46 ms). Noise was shaped by using a raised cosine function with 50-ms rise/fall times.

Each stimulus was presented diotically to the listener through headphones (model HD 565, Sennheiser, Old Lyme, CT), and overall levels were calibrated from each combination of parameters to produce an average output level of 70 dB(A) for continuous speech.

Procedure. The children were tested individually with a single-interval, 16-alternative procedure without feedback. A personal computer controlled the course of the experiment. In each experimental condition, the 48 VCV utterances were presented randomly. All children started with the silence condition. Presentation of the other two conditions (stationary noise and fluctuating noise) was counterbalanced. The children were instructed to identify each stimulus. The 16 possible choices were presented on the screen of the computer. The children gave their responses orally, or they pointed to one of the 16 choices on the computer screen. The experimenter entered the responses by clicking on one of the 16 options on the computer screen. All children were initially trained by asking them to read aloud the letter strings in each response box. Children who showed severe production problems during training were asked to always confirm their oral response by pointing to the label on the screen.

Results. ***Identification performance.*** Table 2 presents the percentage of correct identifications of children with SLI and of the L-match and A-match controls for each experimental condition (silence, fluctuating noise, and stationary noise). In silence, the speech perception deficit was only significant in the A-match comparison ($P < 0.05$) but not in the L-match comparison. In contrast, strong deficits were obtained for children with SLI in the two speech-in-noise conditions in the A-match ($P < 0.001$) and the L-match ($P < 0.02$) comparisons.

Masking release. A clear masking release effect was observed for the three groups of participants (see Table 2). That is, performance was ≈10% better in fluctuating noise than in stationary noise. However, the size of the masking release effect was similar for the three groups. An ANOVA confirmed that the masking release effect was present in the A-match and the L-match comparison (A-match: $\Delta F = 14.58$, $P < 0.001$; L-match: $\Delta F = 11.79$, $P < 0.003$). More importantly, the masking release effect did not interact with the group factor in the A-match or L-match comparison (A-match: $\Delta F = 0.22$, $P > 0.60$; L-match: $\Delta F = 0.14$, $P > 0.70$), indicating that the size of the release effect was similar across groups.

Phonetic feature transmission. The specific reception of three speech features (voicing, place, and manner) was evaluated by information transmission analyses (38) that were performed on the basis of individual confusion matrices. In both comparisons, the results were analyzed by using 2 × 3 × 3 ANOVAs with group (SLI vs. controls), phonetic feature (voicing vs. place vs. manner), and condition (silence vs. fluctuating noise vs. stationary noise) as factors. The results showed that children with SLI performed more poorly than controls (A-match: $\Delta F = 12.87$, $P < 0.002$; L-match: $\Delta F = 7.59$, $P < 0.02$), indicating that the reception of speech features was generally impaired in children with SLI. More importantly, the data exhibited a significant interaction between group and phonetic feature (A-match: $\Delta F = 13.13$, $P < 0.0001$; L-match: $\Delta F = 9.91$, $P < 0.001$), reflecting the fact that the deficit was stronger for voicing (A-match: 23%; L-match: 20%) than for place (A-match,15%; L-match,12%) or manner (A-match, 8%; L-match, 7%). These data are presented in Fig. 1. The triple interaction between group, feature, and condition was significant in the A-match comparison ($\Delta F = 4.9$, $P < 0.02$). This interaction reflects the fact that the voicing deficit was exaggerated in noise (31%) compared with silence (8%). Because this triple interaction failed to reach significance in the L-match comparison ($\Delta F = 1.7$, $P > 0.20$), we only present the double interaction in Fig. 1.

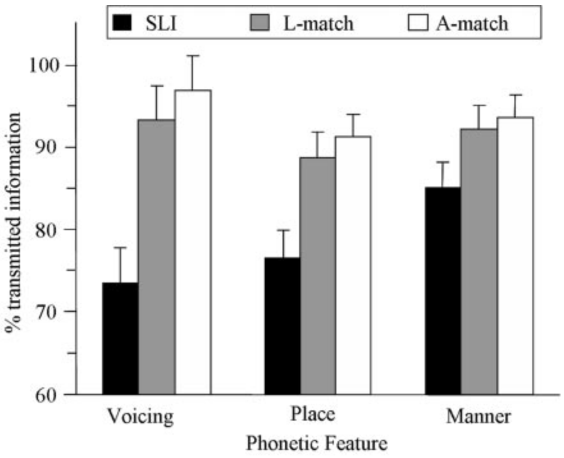

Fig. 1. Percentage of transmitted information for phonetic features in children with SLI and in A-match and L-match controls. Data were pooled across conditions (silence, fluctuating noise, and stationary noise). Error bars indicate SEM.

Discussion. Under optimal listening conditions (silence), children with SLI showed only subtle speech perception deficits that were not significant in the L-match comparison. This finding is consistent with previous reports that have qualified speech perception deficits in developmental language learning disorders as fragile (39). However, under conditions of stationary and fluctuating noise, children with SLI showed substantial speech perception deficits, in the A-match and L-match comparisons. The fact that the speech perception deficit persist in the L-match comparison suggests that it is the cause rather than the consequence of SLI.

Phonetic feature transmission analyses showed that the speech perception deficit was strongest for voicing, although it also affected the reception of all other phonetic features. This finding contrasts with the general pattern of phonetic deficits reported in listeners with sensorineural hearing loss in quiet and noise, for whom reception of place of articulation is mostly degraded, whereas reception of voicing and manner is barely affected (40, 41).

The size of the masking release effect (≈10%) was identical across the three participant groups. That is, all children were apparently able to take advantage of relatively short temporal minima in the fluctuating background to detect and/or restore speech cues. This finding is important because it suggests that the sensory and cognitive processes known to be involved in masking release, such as auditory grouping based on stimulus spectral and fine-structure cues (32), perceptual restoration (42), and informational masking (43), are functional in children with SLI. This pattern of results again contrasts with data from hearing-impaired patients who tend to show reduced or abolished masking release.

Experiment 2

The goal of Exp. 2 was twofold. First, we wanted to provide a replication of our main findings (i.e., speech perception deficits in noise in the presence of intact masking release) with a new group of participants and more intensive psychophysical testing. Second, we wanted to test whether rapid noise fluctuations would result in stronger speech perception deficits than slow fluctuations. Indeed, the fast temporal-processing deficit hypothesis (13) would predict larger deficits under conditions of high-frequency noise modulations (e.g., 128 Hz) than under conditions of low-frequency noise modulations (e.g., 4 Hz). These larger deficits would exist because a rapid processing deficit would prevent the auditory system from taking advantage of the short temporal minima in rapidly fluctuating background noise. In other words, to restore the speech signal in conditions of fluctuating noise, the auditory system needs to go as "fast" as the noise to detect speech cues in the noise valleys. Thus, in case of a rapid-processing deficit, "glimpsing" in noise valleys should be more difficult when modulation frequency is high. Conversely, the "slow" temporal processing deficit hypothesis (e.g., ref. 25) would predict larger deficits under conditions of slow noise modulations (4 Hz) than under conditions of fast noise modulations (128 Hz). In summary, Exp. 2 tested for speech perception deficits in optimal listening conditions and in four masking conditions: stationary, 4-Hz, 32-Hz, and 128-Hz noise.

Methods. ***Participants.*** Ten children (7 boys) with SLI were again recruited from the neuropediatric service of the La Timone Hospital in Marseille, France. The average age was 10.8 years (age range 8.6–12.6 years). Selection procedure was identical to Exp. 1. In particular, all children had a nonverbal IQ of >85 and audiometric thresholds at <20 dB hearing level between 0.25 and 8 kHz. Ten A-match controls were recruited from a local school (mean age, 10.8 years; age range, 8.6–12.9 years). Given that Exp. 1 confirmed the reliability of our results with regard to a L-match group, in the present experiment, we selected only 10 controls that were matched with regard to chronological age (mean age, 10.8 years; age range, 8.6–12.9 years). The controls had normal vision and audition and no history of language or reading disability. Because of time constraints, no language tests were administered for the controls apart from word repetition (phonology) and phonological decoding (nonword reading).

Stimuli and procedure. Stimuli were identical to Exp. 1. To increase the power of the experiment, we made the following changes: First, we reduced the signal-to-noise ratio from 0 to −15 dB (rms). Second, we increased the number of data points by repeating the 48 VCV stimuli four times per condition (196 data points). In the fluctuating noise condition, we added two conditions with a f_m of 4 Hz and 128 Hz.

Results. ***Identification performance.*** Table 3 presents identification performance across all experimental conditions. No significant speech perception deficit was obtained in silence. In contrast, reliable deficits were obtained for all of the speech-in-noise conditions.

To address the issue of whether the speech perception deficit varied as a function of noise modulation frequency, we submitted the data to a 2 × 3 ANOVA with group (SLI vs. controls) and masking condition (4, 32, and 128 Hz) as factors. The ANOVA showed significant main effects of group ($\Delta F = 22.9, P < 0.0001$) and masking condition ($\Delta F = 212.0, P < 0.0001$). The interaction between these effects was not significant ($\Delta F < 1$), suggesting that the speech perception deficit of children with SLI was similar across the three different types of masking conditions.

Table 3. Speech identification performance (% correct) of children with SLI and of A-match controls in silence, stationary (Stat.) noise, and AM noise

			Frequency modulation of AM noise		
Group	Silence	Stat. noise	4 Hz	32 Hz	128 Hz
SLI	92.8	37.4	65.2	58.4	41.7
Controls	99.8	49.2	76.9	73.2	56.6
Diff.	7.0	11.8	11.7	14.8	14.9
t test	$P < 0.07$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$

Diff. indicates the difference between SLI and controls.

Masking release. A clear masking release effect was observed. As in Exp. 1, performance was substantially better in fluctuating noise than in stationary noise. The amount of masking release (i.e., performance in fluctuating noise minus performance in stationary noise) across the three modulation frequencies is presented in Fig. 2*A*.

Statistical analyses confirmed that the size of the masking release varied as a function of frequency: Masking release was best for 4 Hz and worst for 128 Hz ($\Delta F = 136.6, P < 0.0001$). More importantly, children with SLI and controls obtained literally identical masking release effects, with maximum release (≈28%) at the lowest modulation frequency (4 Hz). The interaction between the effects of masking release and group were not significant ($\Delta F < 1$). This result replicates our previous finding that children with SLI show intact masking release.

Phonetic feature transmission. The specific reception of the three speech features was analyzed in a 2 × 3 × 3 ANOVA with group (SLI vs. controls), phonetic feature (voicing vs. place vs. manner), and condition (4, 32, and 128 Hz) as factors. The results showed poorer performance for children with SLI than for controls ($\Delta F = 23.7, P < 0.001$), indicating that the reception of speech features was generally impaired in children with SLI. More importantly, the data exhibited a significant interaction between group and phonetic feature ($\Delta F = 4.1, P < 0.05$), reflecting the fact that the deficit for voicing (24%) was stronger than the deficit for place (12%) or manner (10%). These data are presented in Fig. 2*B*. The triple interaction also reached significance ($\Delta F = 3.2, P < 0.05$) because the deficits of voicing and place were strongest in the 32-Hz condition, whereas the deficit of manner was strongest in the 4-Hz condition. Note, however,

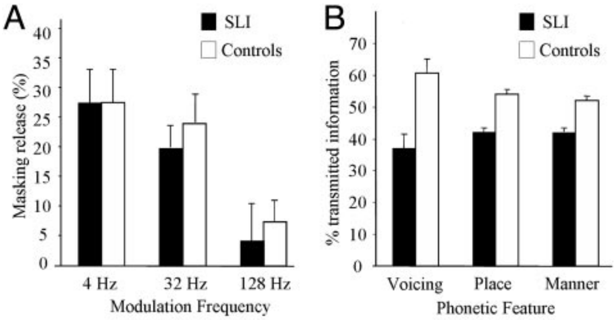

Fig. 2. Masking release (*A*) and the reception of phonetic features (*B*) for children with SLI and for controls. Masking release is the difference between performance in fluctuating noise and stationary noise. Error bars indicate SEM.

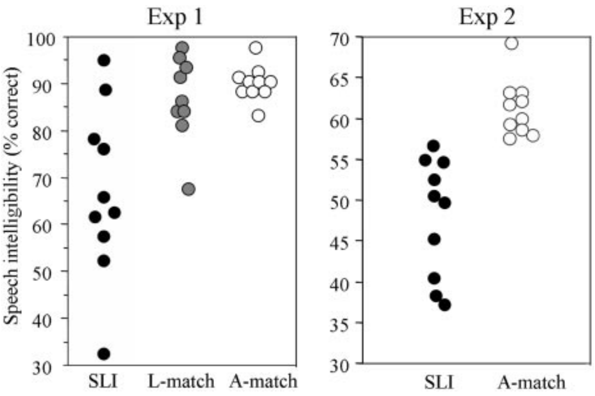

Fig. 3. Scatter plots showing speech-perception-in-noise performance (% correct) under conditions of noise (stationary noise and 32-Hz AM noise) in Exps. 1 and 2.

that in each condition, the strongest deficit was always obtained for voicing.

Discussion. Exp. 2 replicated the main findings of Exp. 1, that is, the existence of a weak speech perception deficit in silence but a robust deficit in noise. The speech perception deficit was equally strong across the different masking conditions (4, 32, and 128 Hz). This finding does not directly support the fast or slow temporal-processing deficit hypothesis (13, 25), which would have predicted a variation of the deficit as a function of modulation frequency. Moreover, as in Exp. 1, masking release was intact in all conditions, which rules out low-level spectral or temporal processing deficits.

Experiments 1 and 2: Individual Performance and Regression Analyses

Two important issues still need to be addressed: (*i*) How general the deficit is, and (*ii*) whether it actually predicts the language deficit. Indeed, it has been previously argued that only a subgroup of children with SLI might have sensory deficits (18, 44) and that the severity of the auditory deficit does not appear to predict the severity of the language impairment (19).

To address the first issue in our data, we analyzed individual scatter plots for children with SLI and controls (Fig. 3). Because both experiments used stationary and 32-Hz amplitude-modulated (AM) noise, we pooled the data from Exps. 1 and 2 across these two conditions. In Exp. 1, seven of 10 children with SLI were at least 1 SD below the mean of the L-match controls, and eight of 10 were at least 1 SD below the mean of the A-match controls. In Exp. 2, nine of 10 children with SLI were at least 1 SD below the mean of the A-match controls. Clearly, the present speech-perception-in-noise deficits should be considered as very general.

To address the issue of predictive power, we calculated Pearson correlations between speech intelligibility in noise and language performance on the word repetition subtest of the L2MA battery (37). This analysis revealed a highly significant correlation between speech intelligibility in noise and word repetition (Exp. 1: $r = 0.74$, $P < 0.0001$; Exp. 2: $r = 0.89$, $P < 0.0001$). Note that the size of the correlations persisted when only children with SLI were entered into the regression model (Exp. 1: $r = 0.73$, $P < 0.0001$; Exp. 2: $r = 0.86$, $P < 0.001$). These results are presented in Fig. 4.

Given that we used a nonword reading test in Exp. 2, we were also able to check whether the speech intelligibility deficits still predicted the phonological deficit when no spoken input was involved, which the results showed was the case. Speech intelligibility in noise still predicted the phonological deficit in nonword reading ($r = 0.83$, $P < 0.0001$).

General Discussion

The main findings of the present study can be summarized as follows. Under optimal listening conditions (silence), children with SLI showed only subtle speech perception deficits. However, under conditions of stationary noise and fluctuating noise, children with SLI showed substantial speech perception deficits. Note that conditions of fluctuating noise are not artificial; they are actually very representative of the kind of listening conditions that children will encounter in their daily life (in schools, for example). Thus, the present results raise the possibility that children with language learning disabilities have very serious problems with noise exclusion, which will certainly have tremendous consequences for normal phonological development. A similar proposal has recently been made with regard to visual (magnocellular) deficits that seem frequently associated with dyslexia (45). The authors showed that dyslexic children do not have visual (magnocellular) processing problems *per se* but rather problems of noise exclusion that become apparent in visual tasks using noisy displays. Noise exclusion could therefore be a very general problem responsible for poor phonological development of children with language learning problems and dyslexia.

The fact that most previous studies and clinical tests investigated speech perception in optimal listening conditions might also explain why they often failed to find robust deficits (39). We thus suggest that clinical testing in the future must involve speech-perception-

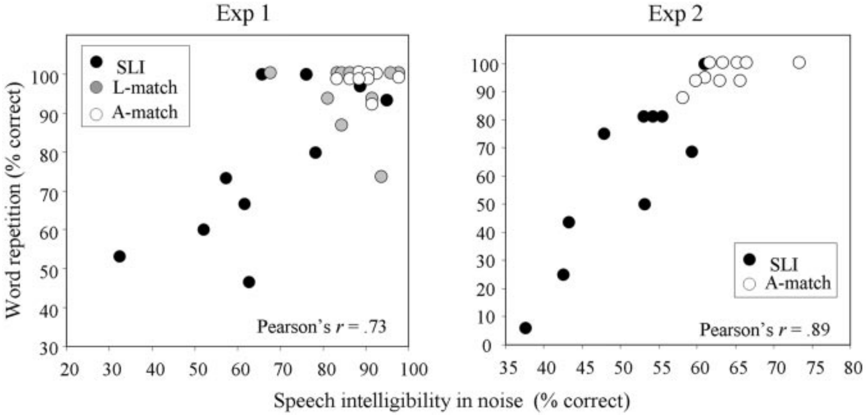

Fig. 4. Correlation between speech intelligibility in noise and performance on a word repetition test for children with SLI and for controls.

in-noise measures to complement the traditional audiometric testing. Such tests are particularly informative if they help assess the reception of specific phonetic features. In this respect, our results showed an interesting pattern. Whereas the voicing feature was most well preserved in noisy conditions in normally developing children, it was the least well preserved in children with SLI. This finding contrasts with the general pattern of phonetic deficits reported in listeners with sensorineural hearing loss in quiet and noise and for whom reception of place of articulation is mostly degraded, whereas reception of voicing and manner are barely affected (40, 41). The comparison with hearing-impaired patients suggests that auditory and speech perception deficits in SLI are central (that is, postcochlear) in origin. The fact that voicing is more affected than other contrasts also suggests that speech perception deficits interfere in a specific way with the development of phonological representations, which in turn might affect other aspects of grammatical development.

An extremely important finding is the presence of intact masking release in children with SLI regardless of the frequency of the noise modulations. This finding suggests that the core deficit of children with SLI is probably not due to poor temporal or spectral resolution, because both of these processes are required to follow the background fluctuations and to access the unmasked parts of the speech spectrum (e.g., ref. 31). If such low-level deficits can be excluded as an explanation of SLI, one might wonder where the speech deficit actually comes from. Overall, our results seem to be consistent with the notion that children with SLI are generally inefficient at processing the information underlying speech identification and that such inefficiency is exacerbated by the adjunction of background noise. More precisely, the present data suggest that the peripheral and central auditory systems of children with SLI encode acoustic information sufficiently well (i.e., envelope, periodicity, fine structure, and spectral cues), but the central auditory system is inefficient at mapping acoustic information onto phonetic features to achieve normal recognition. For unknown reasons, this inefficiency seems to be especially substantial in the case of voicing and particularly susceptible to the interfering effects of the random amplitude fluctuations of masking noise.

Finally, in recent years, a number of criticisms have been made against the importance of auditory deficits in SLI. It has been argued that (*i*) auditory deficits are only associated but not responsible for SLI, (*ii*) that they only affect a small group of children with SLI, and (*iii*) that they do not predict language performance (for review, see ref. 19). None of these criticisms does apply to the present study. First, the present speech perception deficit appears fundamental because it persists even when language level is controlled for; second, it appears general because the great majority of children with SLI show speech-perception-in-noise deficits. Finally, this speech perception deficit is predictive because it correlates very strongly with phonological markers of language impairments.

Conclusion

In conclusion, the present study points to an important connection between SLI and speech perception deficits. Previous studies might have underestimated this link possibly because they investigated speech perception in optimal listening conditions or they focused too narrowly on non-speech deficits. The fact that masking release is intact in children with SLI but not in hearing-impaired patients suggests that basic, temporal, and spectral auditory capacities are relatively spared in SLI. Instead, the deficit must be due to an inefficient mapping of acoustic information onto phonetic features at a central (postcochlear) conversion stage. This result offers new directions for the diagnosis and remediation of SLI.

We thank our brilliant speech therapy students Stephanie Cadet, Marie Gelibert, and Delphine Petinot for help in running the studies; Antoine Giovanni for making this research possible; and Usha Goswami, Patrick Lemaire, John Locke, Marie Montant, Juan Segui, Willy Serniclaes, Liliane Sprenger-Charolles, and an anonymous reviewer for helpful suggestions.

1. Bishop, D. V. M. & Snowling, M. J. (2004) *Psychol. Bull.* **130,** 858–886.
2. Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E. & O'Brien, M. (1997) *J. Speech Lang. Hear. Res.* **40,** 1245–1260.
3. Grela, B. G. & Leonard, L. B. (2000) *J. Speech Lang. Hear. Res.* **43,** 1115–1125.
4. McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M. & Mengler, E. D. (2000) *J. Child Psychol. Psychiatry* **41,** 869–874.
5. Joanisse, M. F. & Seidenberg, M. S. (1998) *Trends Cognit. Sci.* **2,** 240–247.
6. van der Lely, H. K., Rosen, S. & McClelland, A. (1998) *Curr. Biol.* **8,** 1253–1258.
7. Rice, M. L. & Wexler, K. (1996) *J. Speech Lang. Hear. Res.* **39,** 1239–1257.
8. Locke, J. L. (1997) *Brain Lang.* **58,** 265–326.
9. Wright, B. A. & Zecker, S. G. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 9942–9946.
10. Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H. & Merzenich, M. M. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 13367–13372.
11. Ahissar, M., Protopapas, A., Reid, M. & Merzenich, M. M. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6832–6837.
12. Wright, B. A., Bowen, R. W. & Zecker, S. G. (2000) *Curr. Opin. Neurobiol.* **10,** 482–486.
13. Tallal, P. & Piercy, M. (1973) *Nature* **241,** 468–469.
14. Tallal, P. & Piercy, M. (1974) *Neuropsychologia* **12,** 83–93.
15. Tallal, P. & Piercy, M. (1975) *Neuropsychologia* **13,** 69–74.
16. Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L. & Tallal, P. (1996) *Science* **271,** 77–81.
17. Leonard, L. B., Eyer, J. A., Bedore, L. M. & Grela, B. G. (1997) *J. Speech Lang. Hear. Res.* **40,** 741–753.
18. Bailey, P. J. & Snowling, M. J. (2002) *Br. Med. Bull.* **63,** 135–146.
19. Rosen, S. (2003) *J. Phonetics* **31,** 509–527.
20. Marshall, C. M., Snowling, M. J. & Bailey, P. J. (2001) *J. Speech Lang. Hear. Res.* **44,** 925–940.
21. Muneaux, M., Ziegler, J. C., Truc, C., Thomson, J. & Goswami, U. (2004) *NeuroReport* **15,** 1255–1259.
22. Rocheron, I., Lorenzi, C., Fullgrabe, C. & Dumont, A. (2002) *NeuroReport* **13,** 1683–1687.
23. Witton, C., Talcott, J. B., Hansen, P. C., Richardson, A. J., Griffiths, T. D., Rees, A., Stein, J. F. & Green, G. G. R. (1998) *Curr. Biol.* **8,** 791–797.
24. Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S. & Scott, S. K. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 10911–10916.
25. Lorenzi, C., Dumont, A. & Fullgrabe, C. (2000) *J. Speech Lang. Hear. Res.* **43,** 1367–1379.
26. Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J. & Ekelid, M. (1995) *Science* **270,** 303–304.
27. Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M. & Merzenich, M. M. (1997) *Nature* **387,** 176–178.
28. Cornelissen, P. L., Hansen, P. C., Bradley, L. & Stein, J. F. (1996) *Cognition* **59,** 275–306.
29. Mody, M., Studdert-Kennedy, M. & Brady, S. (1997) *J. Exp. Child Psychol.* **64,** 199–231.
30. Miller, G. A. & Licklider, J. C. R. (1950) *J. Acoust. Soc. Am.* **22,** 167–173.
31. Peters, R. W., Moore, B. C. & Baer, T. (1998) *J. Acoust. Soc. Am.* **103,** 577–587.
32. Qin, M. K. & Oxenham, A. J. (2003) *J. Acoust. Soc. Am.* **114,** 446–454.
33. Bacon, S. P., Opie, J. M. & Montoya, D. Y. (1998) *J. Speech Lang. Hear. Res.* **41,** 549–563.
34. Gustafsson, H. A. & Arlinger, S. D. (1994) *J. Acoust. Soc. Am.* **95,** 518–529.
35. Wechsler, D. (1996) *Echelle d'Intelligence pour Enfants* (Centre Psychol. Appl., Paris).
36. Lecocq, P. (1996) *L'ECOSSE*: *Une Épreuve de Compréhension Syntaxico-Sémantique* (Presses Univ. Septentrion, Villeneuve d'Ascq, France).
37. Chevrie-Muller, C., Simon, A. M., Fournier, S. & Brochet, M. O. (1997) *L2MA: Batterie Langage Oral–Langage écrit–Mémoire–Attention* (Centre Psychol. Appl., Paris).
38. Miller, G. A. & Nicely, P. E. (1955) *J. Acoust. Soc. Am.* **27,** 338–352.
39. Blomert, L. & Mitterer, H. (2004) *Brain Lang.* **89,** 21–26.
40. Vickers, D. A., Moore, B. C. & Baer, T. (2001) *J. Acoust. Soc. Am.* **110,** 1164–1175.
41. Baer, T., Moore, B. C. & Kluk, K. (2002) *J. Acoust. Soc. Am.* **112,** 1133–1144.
42. Nelson, P. B. & Jin, S. H. (2004) *J. Acoust. Soc. Am.* **115,** 2286–2294.
43. Summers, V. & Molis, M. R. (2004) *J. Speech Lang. Hear. Res.* **47,** 245–256.
44. McArthur, G. M. & Bishop, D. V. M. (2004) *Cognit. Neuropsychol.* **21,** 79–94.
45. Sperling, A. J., Lu, Z. L., Manis, F. R. & Seidenberg, M. S. (2005) *Nat. Neurosci.* **8**, 862–863.