Available online at www.sciencedirect.com
SCIENCE DIRECT®
Brain and Language 97 (2006) 64–79
Brain and Language
www.elsevier.com/locate/b&l

Auditory temporal information processing in preschool children at family risk for dyslexia: Relations with phonological abilities and developing literacy skills

Bart Boets [a,*], Jan Wouters [b], Astrid van Wieringen [b], Pol Ghesquière [a]

[a] *Centre for Disability, Special Needs Education and Child Care, University of Leuven, Belgium*
[b] *Laboratory for Experimental ORL, University of Leuven, Belgium*

Accepted 9 July 2005
Available online 19 August 2005

Abstract

In this project, the hypothesis of an auditory temporal processing deficit in dyslexia was tested by examining auditory processing in relation to phonological skills in two contrasting groups of five-year-old preschool children, a familial high risk and a familial low risk group. Participants were individually matched for gender, age, non-verbal IQ, school environment, and parental educational level. Psychophysical thresholds were estimated for gap-detection, frequency modulation detection, and tone-in-noise detection using a three-interval forced-choice adaptive staircase paradigm embedded within a computer game. Phonological skills were measured by tasks assessing phonological awareness, rapid serial naming, and verbal short-term memory. Significant group differences were found for phonological awareness and letter knowledge. In contrast, none of the auditory tasks differentiated significantly between both groups. However, both frequency modulation and tone-in-noise detection were significantly related to phonological awareness. This relation with phonological skills was not present for gap-detection.

Keywords: Auditory temporal processing; Reading development; Dyslexia; Prediction; Preschoolers; Phonological awareness; Verbal short-term memory; Rapid automatic naming

1. Introduction

Developmental dyslexia is characterised by serious reading and spelling difficulties that are persistent and resistant to the usual didactic measures and remedial efforts. At present it is well established that a major cause of these problems lies in the phonological domain (see Snowling, 2000 for a review). One hypothesis maintains that this phonological deficit results from a more fundamental deficit in the basic perceptual mechanisms that are responsible for auditory temporal information processing.

The auditory temporal hypothesis originated from studies on children with specific language impairments (SLI) and was later extended to dyslexia. The empirical evidence started with Tallal's repetition task (Tallal, 1980). In this temporal order judgement (TOJ) task, two complex tones with different fundamentals were presented in pairs at various inter-stimulus intervals (ISI) and the listener responded with two button presses to identify the order of the stimuli presented. Tallal found that children with dyslexia, in comparison to normal readers, were impaired in discriminating and sequencing pairs of short-lived stimuli with short ISI, and concluded that the dyslexic deficit was specific to processing stimuli that are brief and occur in rapid succession. Moreover, she found a high correlation between this basic perceptual processing of non-speech signals and phonological skills ($r = .81$). Following further evidence that dyslexic and SLI children had great difficulty discriminating

* Corresponding author. Fax: +32 0 16 32 59 33.
E-mail address: bart.boets@ped.kuleuven.be (B. Boets).

0093-934X/$ - see front matter
doi:10.1016/j.bandl.2005.07.026

syllables containing stop consonants (such as /ba/ and /da/), the claim of a temporal deficit was extended to apply to both non-linguistic and linguistic auditory stimuli (Tallal & Piercy, 1973; Tallal, Miller, & Fitch, 1993). Since discrimination of such syllables critically depends on accurate detection of the rapid frequency changes in the first milliseconds of voicing, inaccurate detection of these formant transitions would inevitably interfere with the identification of the phonological cues that are typical for spoken language. This hypothesis of a direct association between basic auditory processing and speech or language processing was strengthened by demonstrating that speech stimuli with lengthened transitions were much better discriminated (Tallal & Piercy, 1975). From this association sprang the claim that the temporal auditory problem *caused* the language problem, and subsequently the deficient phonological and reading development. During decades this supposed causal mechanism has been put forward as a plausible explanation of dyslexia.

Since the formulation of this theory there have been multiple studies exploring the auditory temporal abilities of individuals with dyslexia. While the bulk of studies has been done on adults, a minority of recent studies focused on school aged children and some very few on preschoolers. In line with the scope of our study, we will mainly restrict our report to psychophysical studies using specific non-speech stimuli to examine younger subjects.

Probably the most straightforward way to measure temporal processing is a gap-detection task; this task estimates the smallest detectable interruption in an auditory stimulus. Van Ingelghem and colleagues (van Ingelghem et al., 2001, 2005) found a significant gap-detection deficit in 11-year-old dyslexic children compared to normal reading children. Moreover, the results on the task were significantly related to both real word reading and non-word reading ($r = -.57$ and $r = -.60$, respectively). These results were replicated in a broader study in dyslexic and normal reading children matched for sex, age, and intellectual ability (Van Ingelghem, Boets, van Wieringen, Ghesquiere, & Wouters, 2004). The observed results are in line with McCroskey and Kidder (1980), but are not consistent with observations reported by McAnally and Stein (1996), Schulte Korne, Deimel, Bartling, and Remschmidt (1998), and Adlard and Hazan (1998). Hautus, Setchell, Waldie, and Kirk (2003) also observed higher gap-detection thresholds in dyslexic subjects, but these thresholds were only significantly higher for the young reading-impaired subjects (aged 6–9 years) and not for the older ones (aged 10 years up to adulthood). The authors interpreted these results as suggestive for a passing maturational lag in temporal acuity in children with dyslexia. In an interesting study of Fischer and Hartnegg (2004), investigating a large group of subjects covering an age-range of 7–22 years, a higher proportion of subjects with dyslexia were unable to perform a gap-detection task even at its easiest level. However, within the group of participants for whom a threshold value could be assigned, there was no significant difference between dyslexics and normal readers.

Studdert Kennedy and Mody (1995) challenged Tallal's auditory theory and argued that the observed phonological impairments in dyslexics are in origin speech-specific and cannot be attributed to a more general lower-level auditory deficit. Besides this fundamental criticism they also postulated that stimulus processing should only be regarded as temporal when the defining features of the stimuli are changing in time and not merely because of their rapid and brief presentation. This new temporal concept resulted in a new series of studies that investigated auditory temporal processing in dyslexia using "dynamic stimuli" (see Talcott et al., 2000). Most of these studies were carried out on adult samples and demonstrated a relative impairment in sensitivity to amplitude modulation (AM) (McAnally & Stein, 1997; Menell, McAnally, & Stein, 1999; Rocheron, Lorenzi, Fullgrabe, & Dumont, 2002) and frequency modulation (FM) (e.g., Stein & McAnally, 1995). In addition, Witton et al. (1998) found that sensitivity to 2 and 40 Hz FM, for both dyslexics and controls, significantly correlated with phonological decoding skills. This relationship between FM sensitivity and phonological ability has also been demonstrated by Talcott et al. (1999) in a random group of children. More recently, Van Ingelghem et al. (2005) demonstrated a significant difference in FM sensitivity in a group of 11-year-old dyslexic children compared to normal reading children. However, in a similar but broader well-controlled study with IQ-matched control subjects, this difference could not be replicated (Van Ingelghem et al., 2004).

These studies with 'dynamic' stimuli again point to an auditory temporal processing deficit as a possible cause of dyslexics' phonological problems. Accurate tracking of amplitude and frequency changes is exactly what is needed for the perception of speech, which is characterised by temporal and spectral variations. Since speech perception is the basis for developing phonological skills, it is likely that impairments in AM and FM detection affect phonological skill development via speech perception (McBride Chang, 1996).

With respect to preschool subjects, as far as we know, there have only been a few longitudinal studies applying psychophysical measures. Heath and Hogben (2004) and Share, Jorm, Maclean, and Matthews (2002) administered Tallal's repetition test to a large unselected group of kindergarten children and followed them up until, respectively, second and third grade. However, neither of both research groups was able to predict grade two or three literacy scores based on the auditory data collected in preschool. Conversely, Benasich and Tallal (2002) administered an operantly conditioned head-turn version of the repetition test to infants 7.5 months of age born into families who were either positive or negative for family history of language impairment (SLI). Not only did these authors observe significantly poorer thresholds for children born into risk families, but they also demonstrated that rapid auditory processing thresholds at 7.5 months of age were the single best predictor of language development at two years of age and together with gender predicted up to 40% of variance in language outcome at three years of age. Unfortunately, information about literacy development and its relation with rapid processing thresholds is currently not yet available for these children. In contrast with the sparse psychophysical studies, there is a growing number of neurophysiologic studies focusing on the temporal characteristics of speech processing in very young subjects that already demonstrated promising results comparing genetically high risk versus low risk children (e.g., Jyväskylä Longitudinal Study of Dyslexia, see Lyytinen et al., 2001; Dutch Dyslexia Research Programme; Molfese, 2000).

Notwithstanding the large number of studies demonstrating an auditory deficit in dyslexics, the explicit causality of the auditory hypothesis has never been established directly by means of a longitudinal study. Here, we report data from a longitudinal study that explores (i) the development of basic auditory skills, speech perception, phonological abilities, and reading skills over a two-year time period from the beginning of the last year of kindergarten[1] up to the end of the first year of primary school; (ii) the mutual relations between these abilities and the way they influence each other over time. In this paper, we will discuss the first results about the relation between auditory temporal processing skills and phonological skills in two contrasting groups of preschool children, i.e., a genetically high risk and a genetically low risk group.

Auditory processing was assessed by means of three psychophysical threshold tests: one for gap-detection in noise (GAP), one for 2 Hz FM-detection (FM) and one for tone-in-noise detection (TN). With the GAP-detection task, we tested the hypothesis of a deficit in 'rapid and brief' temporal processing. With the FM-detection task, we verified the hypothesis of a deficit in the processing of 'dynamic stimuli.' The TN-task was included as a non-temporal control task to verify the specificity of any observed temporal deficit, i.e., we wanted to examine whether a deficit might be the result of failing performance on auditory psychophysical tasks in general. Phonological processing was assessed by administering a broad test battery comprising tasks for rapid serial naming, verbal short-term memory and phonological awareness. Developing literacy skills were measured using a letter knowledge task.

In this study, we aimed to answer the following questions. First, is it possible to obtain reliable results while administering such complex psychophysical tasks to very young subjects? Second, do genetically high risk children, in comparison to low risk children, perform significantly worse on phonological tasks? Third, do genetically high risk children, in comparison to low risk children, perform significantly worse on psychophysical tasks for auditory temporal processing? Fourth, are these auditory processing abilities related to phonological and developing literacy skills?

2. Methods

2.1. Participants

Sixty-two five-year-old children were included in the study. Half of the participants were children of 'dyslexic families,' the so-called high-risk group (HR); the other half were control children of 'normal reading families,' the so-called low-risk group (LR). Since dyslexia tends to run strongly in families, preschoolers with dyslexic relatives are more likely than other children to develop reading problems. Gilger, Pennington, and DeFries (1991) estimate that roughly between 30 and 50% of such children will become reading disabled.

The *HR children* were recruited by means of referrals and public announcements to encourage families with a child entering the last year of nursery school and having at least one member with a formal diagnosis of dyslexia to engage in the study. Following this recruitment, we received over 300 registrations. Based on a short semi-structured telephone interview we selected 162 potential candidate families who were sent three questionnaires. One questionnaire investigated in detail the reading and spelling (dis)abilities of all family members up to third degree and investigated the general development of the preschooler. The two other questionnaires were translations and adaptations of the Adult Reading History Questionnaire (Lefly & Pennington, 2000) and investigated the reading experiences and educational level of each parent. We assessed educational level using the ISCED-scale (International Standard Classification of Education by UNESCO, 1997), by converting classifications on the original seven-point scale to a three-point scale. Out of these 162 potential candidates, we selected 31 preschoolers based on the following criteria: having at least one first-degree relative being diagnosed as reading disabled by an authorized educational psychology service; being native Dutch speaker; born in 1998 and entering last year of nursery school at the beginning of the study; no history of brain damage, long-term hearing loss or visual problems.

[1] In Belgium school system formal instruction starts in Grade 1 at six years. This means in kindergarten no reading instruction is offered. This is in contrast to the kindergarten group studied by Share et al. (2002), who already received formal reading instruction.

Table 1
Characteristics of participants: descriptive statistics, paired t tests and Fisher's Exact Test

	HR ($n = 31$)			LR ($n = 31$)			p
Non-verbal IQ (mean/SD)[a]	107 (14)			111 (13)			.07
Age in months (mean/SD)	64 (3)			64 (3)			.83
Frequency distribution	Low (1)	Middle (2)	High (3)	Low (1)	Middle (2)	High (3)	
Maternal educational level	3	8	20	1	9	21	.71
Paternal educational level	6	14	11	2	16	13	.43

[a] Transformed RCPM scores with $M = 100$, $SD = 15$.

Of the 31 selected HR children, 3 reported with reading problems in only one (first-degree) relative, 5 reported with reading problems in two (first-degree) relatives, 12 with reading problems in three (first-degree or extended) relatives, and 11 with reading problems in four or more (first-degree or extended) relatives.

To further increase the number of future dyslexic children in our test population, we selected relatively more male than female preschoolers (ratio: 18M/13F). Finally, to exclude the "garden variety" poor readers whose literacy is poor due to a low IQ, we selected proportionally more children out of gifted and higher educated families (see Snowling, 2000).

Children of the LR group met the same selection criteria, with the restriction that they were not allowed to show any history of speech or language problems and that none of their family members might have suffered any learning or language deficiencies. For every individual HR child, we searched for the best matching LR control child based on five criteria: (1) educational environment, i.e., same nursery school, (2) gender, (3) age, (4) non-verbal intelligence, and (5) parental educational level. In effect, we selected the particular control child out of the group of same-sex classmates of the HR child. All these children—including the HR child—were administered an adapted version of the Raven Coloured Progressive Matrices (RCPM) (Raven, Court, & Raven, 1984), a collective non-verbal intelligence test measuring spatial reasoning. To assure the child's motivation and attention during testing, we integrated the test procedure within a game-like fairy tale.

From this group, we selected the best overall matching LR control child. For the first two criteria (educational environment and gender) matching was perfect; for the remaining three criteria matching was as good as possible within the restrictions of having to choose within a concrete class group. For the age criterion, we were able to select all control children within an age-difference range of maximal five months. For the IQ criterion only four children differed more than one standard deviation with their matched counterpart. Since the RCPM is very age-sensitive in young children, we used age corrected norms to calculate this non-verbal IQ score. Concerning parental educational level we gave relatively more importance on getting a good fit for maternal than for paternal educational level. In this way, we were able to match 21 children in a perfect way for maternal educational level and 15 for paternal educational level.

Table 1 gives descriptive characteristics of both groups and test statistics. The mean age for both the HR and LR group was 5 years and 4 months, not being statistically different ($p = .83$). The non-verbal IQ scores were slightly above population average (107 for HR group and 111 for LR group) and neither differed significantly ($p = .07$). Fisher's Exact Test also confirms that both groups did not differ in frequency distribution of the different educational categories ($p = .71$ for maternal and $p = .43$ for paternal educational level).

2.2. Apparatus

2.2.1. Phonological tests

Tests were selected to reflect the three traditional domains of phonological skills (Wagner & Torgesen, 1987). Phonological awareness was measured by three sound identity tasks and a rhyme fluency task. Verbal short-term memory was measured by a digit span test and a non-word repetition task. Rapid automatic naming was assessed by administering a colour and an object rapid naming task.

2.2.1.1. Sound identity tasks. The child was required to choose from four alternatives the word that had the same (a) first sound, (b) end sound or (c) end rhyme as a given word (de Jong, Seveke, & van Veen, 2000, adapted by van Otterloo & Regtvoort). The distracter alternatives were systematically constructed to prevent guessing. All words were high frequent one-syllabic Dutch words. Each item consisted of a row of five pictures. The first picture represented the given word and was separated from the other pictures by a vertical line. All items were named for the child. The first-sound and end-sound identity tasks both consisted of 10 items, preceded by two practice items, and had a maximum score of 10. The rhyme identity task consisted of 12 items, preceded by two practice items, and had a maximum score of 12.

2.2.1.2. Rhyme fluency test. Participants were presented a one-syllable word and were required to produce as many as possible rhyming (non-)words within a 20 s time period. To familiarize them with the task, the experimenter already offered an example rhyme word for every target item. Since we were interested in measuring rhyming abilities irrespective of vocabulary knowledge, the test score was the total number of phonologically correct responses, regardless of whether it was a real Dutch word or not. The test consisted of eight items, gradually getting more difficult, and was preceded by two practice items.

Based on the factor analysis of the phonological data (see Section 3), we recalculated the results of the rhyme fluency test to create a purer measure of rhyming ability, uncontaminated by fluency. For this *Simple rhyme test* each item was scored in binary fashion, and treated as correct if at least one rhyming response (word or non-word) was produced. The maximum score on the test was eight.

2.2.1.3. Non-word repetition test. A non-word repetition test (NRT) is frequently used as a pure measure of verbal short-term memory. Since neither the non-words nor the constituent syllables of the non-words used in the NRT correspond to existing words, the use of long-term memory representations to support recall of the non-words is prevented. The test was developed after a Dutch adaptation (Scheltinga, 2003) of the non-word repetition test reported by Gathercole, Willis, Baddeley, and Emslie (1994). This Dutch version of the NRT was again adapted for the use with Flemish children. The Flemish NRT consisted of four categories of non-words, varying in word length from two to five syllables. Each category contained 12 non-words. All 48 non-words and two test items were recorded on a CD and were presented once. In contrast to Gathercole et al. (1994), the presentation order of the words was determined by word length; starting with all the two-syllabic words and climbing gradually up to the five-syllabic words. The test consisted of 48 non-words, preceded by two practice items, and had a maximum score of 48.

2.2.1.4. Digit span forward. The test assessed the immediate serial recall of spoken lists of digits between 1 and 9. Prior to testing, children were asked to count from 1 to 10 to familiarize them with the counting string and to reduce the influence of possible differences in digit knowledge. After a practice session, three trials of each list length were presented, starting at a sequence of two digits. Testing continued with increasing list length until the child failed on two of three trials of the same list length. The test score was the total number of correctly recalled lists. To standardize assessment, all lists were recorded and presented on CD at a rate of one digit per second. For each list length, the stimuli of the first two trials were taken from the WISC-III (Wechsler, 1992). The third trial was selected from the Working Memory Battery for Children (see Gathercole & Pickering, 2000).

2.2.1.5. Rapid automatic naming. The test assessed the rapid serial naming for two types of familiar symbols: colours and objects (van den Bos, Zijlstra, & Spelberg, 2002). The objects represented five high-frequent, one-syllabic words: boom (‘*tree*’), eend (‘*duck*’), stoel (‘*chair*’), schaar (‘*scissors*’), and fiets (‘*bicycle*’). The colours were represented by small rectangles in black, blue, red, yellow, and green. For each type of symbol one test card was given, consisting of 50 symbols in a random order (5 columns of 10 symbols). The child was instructed to name the symbols on a card as fast and accurately as possible. Prior to testing, the child was required to name the symbols in the last column of a card to determine whether he/she was familiar with all the presented symbols. For each card, the number of errors and the time to completion were recorded. Subsequently, the time to completion was transformed to the number of symbols named per second. As such, a higher speed score on the test corresponded to a higher naming speed.

2.2.2. Letter knowledge

To get a preliminary idea about the stage of reading development, we administered a letter knowledge task, since many studies have consistently proven this task to be the best predictor of the later development of literacy skills (see, e.g., Elbro & Scarborough, 2003, for a recent overview). To test for the receptive and productive letter knowledge, the 16 most frequently used letters in Dutch books for children were selected (Rolf & Van Rijnsoever, 1984).

2.2.2.1. Productive letter knowledge. Sixteen printed letters were presented on a card. The child had to name each of these letters. Both the sound and the name of a letter were considered correct. The maximum score was 16.

2.2.2.2. Receptive letter knowledge. Sixteen printed letters were presented on a card. The experimenter named all letter sounds in random order. After each sound, the child had to indicate the printed letter that matched the sound. The maximum score was 16.

2.2.3. Auditory tests

2.2.3.1. Audiometric pure-tone detection. Prior to administering any auditory psychophysical test we assessed all children on an audiometric pulsed pure-tone detection task to check for any hearing loss. All but one child obtained a PTA-score below the 25 dB HL criterion. For this child showing mild hearing loss, all further auditory testing was administered with increased stimulus amplitude proportionally to the hearing loss. Since detailed inspection of all her test results did not show any anomalies, her data were not excluded from further analyses.

2.2.3.2. GAP-detection test. In the GAP-detection test, white noise stimuli were used. The target stimulus was a white noise stimulus containing a silent gap. The reference stimulus was an uninterrupted white noise. Stimuli were cosine gated on and off with 50 ms rise and fall times. Gap rise and fall times were 0.5 ms. Sixty-four target stimuli were constructed, comprising 32 gap sizes, varying between 100 and 0.1 ms. Gap length decreased with a factor 1.2 from 100 ms towards 6.5 ms. From here on gap length decreased with a fixed step size of 0.4 ms. To prevent participants from using overall duration as a cue for detection, the length of both the target and the reference stimulus was varied randomly from presentation to presentation (van Wieringen & Wouters, 1999). In the target stimulus, the length of the markers (i.e., noise components surrounding the gap) varied between 250 and 650 ms including on and off set (i.e., 250, 400, 500, and 650 ms). The length of the reference stimulus was 750, 900 or 1050 ms including on and off set. Stimuli were presented monaurally at 70 dB SPL with an inter stimulus interval (ISI) of 400 ms.

2.2.3.3. FM-detection test. In the FM-detection test, stimuli can be defined as $x(t) = A\sin[2\pi f_c t + \beta \sin(2\pi f_m t)]$ in which β is the modulation index ($\beta = \Delta f / f_m$), f_c the carrier frequency, f_m the modulation frequency and Δf the frequency deviation. The target stimulus was a 2 Hz frequency modulation (f_m) of a 1 kHz carrier tone (f_c) with varying modulation depth Δf. Modulation depth decreased with a factor 1.2 from 100 Hz towards 11 Hz. From a Δf of 11 Hz, a step size of 1 Hz was used. Frequency modulation in the target stimulus was sinusoidal and the modulation envelope was always in sine phase. The reference stimulus was a pure tone of 1 kHz ($\beta = 0$). The length of both the reference and the target stimulus was 1000 ms including 50 ms cosine-gated onset and offset. Stimuli were presented monaurally at 70 dB SPL with an ISI of 350 ms.

2.2.3.4. Tone-in-noise detection task. In the tone-in-noise (TN) detection task participants had to detect pure tone pulses within a one-octave noise signal, centered around 1 kHz (from 707 to 1414 Hz). Noise was presented monaurally at 55 dB SPL with an ISI of 300 ms. The length of both the target and the reference stimulus was 1620 ms, including 20 ms linearly gated rise and fall times. For the target stimulus noise contained two pure 1 kHz pulses of 440 ms, including 20 ms linearly gated on and offset. The first pulse started 320 ms after stimulus onset, the second one 960 ms after stimulus onset. The amplitude of the pulses varied with a signal-to-noise ratio (SNR) between +25 and −20 dB. From +25 to −3 dB SNR amplitude decreased with a 4 dB step size. From here on amplitude decreased with a 1 dB step size.

All stimuli for GAP, FM, and TN were generated in MATLAB 5.1 and saved as 16-bit wav-files (sample frequency 44,100 Hz) on the hard disc of a Dell Latitude C800 and Toshiba Satellite 1400-103 portable computer. They were presented using an integrated audio PC-card and routed to an audiometer (Madsen OB622) in order to control the level of presentation. The stimuli were presented monaurally over calibrated TDH-39 headphones.

2.2.4. Psychophysical procedure

For all GAP, FM, and TN tests a similar psychophysical procedure was used. Thresholds were estimated using a three-interval forced-choice oddity paradigm. The subject's task was to identify the 'odd' stimulus, the one that sounded different from the other two. The length of the gap, the depth of modulation and the amplitude of the sinusoidal pulses were adjusted adaptively using a two-down, one-up rule, which targeted the threshold corresponding to 70.7% correct responses (Levitt, 1971). In all tests, a threshold run was terminated after eight reversals. Thresholds for an individual run were calculated by the geometric mean of the values of the last four reversals. For each participant three reliable threshold estimates were determined for every experiment. Prior to auditory data collection, participants were given a short period of practice, comprising supra-threshold trials, to familiarise them with the stimuli and the task.

The forced-choice oddity paradigm was controlled by APEX, a software module developed for psycho-acoustical and psycho-electrical auditory testing (Laneau, Boets, Moonen, van Wieringen, & Wouters, 2005). To make the rather boring psychophysical tests more interesting and child friendly we integrated them in an interactive video game with intro and outro animation movies, aimed to transform the abstract meaningless acoustical signal into a concrete and well known 'daily life signal.' This concept of testing was based upon earlier work of Soderquist and Shilling (1992) and Wightman and co-workers (Allen, Wightman, Kistler, & Dolan, 1989; Wightman, Allen, Dolan, Kistler, & Jamieson, 1989). During testing, the three intervals of every trial were visually represented on the screen by an identical character. The characters were animated synchronously with the presentation of the corresponding sound interval in order to create a psychological link between the moving object on the screen and the presented sound signal. The child's task was to choose, by pointing to the touch screen, which of the three characters corresponded to the target sound (i.e., sounded different from the other two). Immediately after the child touching the screen, visual feedback was given in the form of a moving cartoon. After correct selection, feedback was much more spectacular—and as such much more reinforcing—than after incorrect selection. During the sequence of trials constituting a single run, APEX also provided a more global reinforcement by adding little smiley faces to a rising ladder structure for every correct response and removing them again after every incorrect response. At the end of the run the number of smiley faces was evaluated and the child was rewarded proportionally.

For the GAP-detection experiment the video game displayed a teasing mouse who is trying to wake up three sleeping snakes because she wants to play with them. The child had to predict which one is going to wake up by listening to the specific interrupted noise sound 'ssss-ssss.' For the FM-detection experiment the introductory movie showed a mother dragon sitting in front of three eggs. The child's task was to predict which egg is ready to hatch, by listening to the baby dragon crying inside ('wouwouwouw'). For the TN-detection task we presented an introductory animation movie about a little monkey waiting desperately at the school gate for his father to pick him up by car. Three cars arrive at the school (making noise sound) and the child had to identify the father's car by listening to him honking his horn (= the pure tone pulses).

2.3. Data collection

All auditory and phonological data were collected within a 70-day period between the second and the fourth month of the last year in kindergarten. Intelligence testing (RCPM) took place one month earlier. Data collection was carried out by qualified psychologists and audiologists. Testing took place in a quiet room at the children's school. Since the LR child was selected out of the HR child's classmates, we could always test both children in exactly the same circumstances.

All phonological tests and the letter knowledge task were administered individually in one day during three sessions. After every subtest children were rewarded by receiving little stickers or stamps.

Auditory data were collected during two consecutive days. Testing always started with the pure-tone audiogram. Then we administered one run of the TN test since this is conceptually the easiest psychophysical task. Consequently, we administered one run of the GAP and FM test in a contra balanced way. This sequence was continued until we had three threshold estimates for every experiment.

2.4. Statistical analysis

Prior to analysis all data were individually checked for unexpected outliers. This resulted in the removal of only two unreliable phonological test scores for one subject of the dyslexic group.

All results were analysed in a paired wise manner, comparing HR versus LR group at the level of the matched individuals. Although both groups did not show a significant difference on any of the matching criteria, we decided to rule out any possible influence of age, non-verbal intelligence or parental educational level by controlling for these variables in our analyses. As such, we analysed the data using Mixed Model Analysis (MMA) with school as a random variable (1–31) and participant group (HR versus LR) as the fixed between-subject variable (Littell, Milliken, Stroup, & Wolfinger, 1996). Age, non-verbal IQ and educational level of both mother and father were added as fixed (co)variables. Additionally, for the auditory data, we also computed a series of Repeated Measures MMA with threshold run (1–3) as the within-subject variable, participant group as the between-subject variable and with the same covariates as mentioned above. MMA was chosen not merely to allow a paired wise comparison, but also because of its robustness in analysing semi-normally distributed data (Verbeke & Lesaffre, 1997). To approach a normal distribution for more variables, the GAP and FM thresholds and the results on the letter knowledge task were log-transformed prior to MMA. To explore the internal consistency of the different phonological tasks, Cronbach's α-coefficients were calculated. To explore the structure and the mutual relationships of the phonological tests a principal component factor analysis was done with varimax rotation. Relationships between variables were analysed using Spearman correlation coefficients.

3. Results

3.1. Phonological skills and letter knowledge

Descriptive statistics, MMA results and reliabilities (Cronbach's α) for all measures are displayed in Table 2. The internal consistency of the simple rhyme task, the non-word repetition test and the letter knowledge task was good. The reliability of the rhyme, first phoneme and end phoneme identity tasks was somewhat lower, probably because these tasks appeared to be rather difficult.[2]

For the colour and picture rapid naming tasks, we examined whether there might have been a speed-accuracy trade-off. Since there was no significant correlation between naming speed and error number and since the quantity of errors did not differ between both groups, we did not correct the speed scores for error rate.

Although the results on almost any phonological task were in the expected direction with the HR-group scoring less well than the LR-group, not all these group differences turned out to be statistically significant. Considering the tasks meant to measure phonological awareness, the scores of the HR-group were significantly lower for the rhyme fluency test and for the end phoneme identity test, with the results on the simple rhyme task being

[2] It is worth mentioning that for both HR and LR-group, the mean scores on the sound identity tasks are well above chance level. This is in contrast to a Dutch study where an odd-one-out categorization version of these tests was administered and where children of the same age were not able to exceed chance level on the last-sound and first-sound categorization test (de Jong & van der Leij, 1999).

Table 2
Phonological abilities and letter knowledge: descriptive statistics and *p* values for paired wise MMA, controlling for non-verbal IQ, age, and parental educational level

Measures	Cronbach's α	Maximum	HR		LR		*p*
			M	*SD*	*M*	*SD*	
Phonological awareness							
Rhyme fluency	—	—	14.3	9.6	18.9	8.2	.02
Simple rhyme	.88	8	6.2	2.4	7.3	1.8	.05
Rhyme identity	.69	12	8.7	2.7	9.8	1.9	.23
First-sound identity	.59	10	4.6	2.1	5.6	2.4	.20
End-sound identity	.63	10	4.5	2.2	5.9	2.4	.02
Rapid serial naming							
Colour naming	—	—	0.6	0.1	0.7	0.2	.12
Picture naming	—	—	0.6	0.1	0.7	0.2	.11
Verbal short-term memory							
Digit span	—	21	7.0	1.6	6.9	1.5	.66
Non-word repetition test	.84	48	16.8	5.3	20.3	7.1	.08
Letter knowledge	.90	32	5.6	6.6	8.2	6.9	.03
Factor phonological awareness	—	—	−0.73	1.2	0.00	1.0	.04
Factor rapid naming	—	—	−0.19	0.9	0.00	1.0	.30
Factor verbal STM	—	—	−0.14	0.9	0.00	1.0	.72

marginally significantly lower. For the rapid serial naming tests neither the picture naming nor the colour naming task differed significantly between both groups. On the tests of verbal short-term memory only the non-word repetition test showed a slight tendency towards differing between both groups.

It is noteworthy that some of the group differences have been tempered by applying the strict controlling MMA design. For example, by comparing both groups without controlling for non-verbal IQ, age, and parental educational level, we also found significant group differences for the non-word repetition test and simple rhyme test, and marginally significant differences on the first phoneme identity task.[3]

With respect to developing literacy skills, we found a significant group difference on the log-transformed letter knowledge scores.

Because the number of participants was not big enough to perform a reliable confirmatory factor analysis, an exploratory principal component factor analysis with varimax rotation was carried out to examine the data structure. Since we wanted to explore the *unique* relation between the individual phonological sub skills and the auditory processing skills, unrelated (i.c. orthogonal) phonological factors were calculated. This analysis revealed that the rhyme fluency task disturbed the assumed threefold phonological structure. A post hoc explanation for this phenomenon might be that this fluency task depended only marginally on rhyming skills, while depending mostly on skills as flexible and creative thinking. Replacement of this test by the described simple rhyme test, which was a pure rhyme measure, resulted in an excellent three-factor structure (based on the eigenvalue criterion). The first factor had heavy loadings of the three sound identity tasks and the simple rhyme task, and a more modest loading of the non-word repetition test (see Table 3). As a consequence, this factor could be reliably labelled as the Phonological Awareness Factor (PhAw). The second factor was completely determined by heavy loadings of both the colour and the picture naming task, and as such could be regarded as the Rapid Automatic Naming Factor (RAN). Finally, the non-word repetition test and the digit span loaded heavily on the last factor, namely the Verbal Short-Term Memory Factor (VSTM).

Table 2 displays descriptive statistics and MMA results for these three latent phonological factors. To assist in the interpretation of these results, factor values

Table 3
Principal component factor analysis with varimax rotation: factor loadings of the phonological measures

	Factor 1 Phonological awareness	Factor 2 Rapid automatized naming	Factor 3 Verbal short-term memory
Simple rhyme	.68		
Rhyme identity	.83		
First-sound identity	.82		
End-sound identity	.80		
Colour naming		.91	
Picture naming		.92	
Digit span			.83
Non-word repetition test	.36		.74

Note. Only factor loadings above .35 have been depicted.

[3] Interestingly, this change in significance was caused almost completely by the influence of maternal educational level.

were transformed to effect sizes relatively to the mean and standard deviation of the LR-group. As can be seen, both groups did not differ on Verbal Short-Term Memory and Rapid Automatic Naming, but they differed significantly on Phonological Awareness ($p = .04$).

3.2. Auditory measures

For every auditory experiment, a paired wise Repeated Measures MMA was computed with group as between-subject variable (HR versus LR) and threshold run as within-subject variable (run 1–3). Results can be summarised as follows: (a) neither for GAP, nor for FM, nor for TN-detection there was a significant group effect ($p = .08$, $p = .33$, and $p = .66$, respectively); (b) the three auditory tests showed a significant effect of threshold run ($p < .001$, $p = .01$, and $p < .001$, respectively); and (c) for none of the tests the group by run interaction was significant ($p = .67$, $p = .21$, and $p = .47$, respectively). For every auditory experiment, post hoc analysis revealed that none of the three threshold measures differentiated significantly between HR and LR group. Furthermore, for every experiment there was only a significant learning effect from the first to the second run; the second and third run did not differ significantly from each other.

Importantly, in contrast to the phonological data, the results were not influenced by applying the conservative MMA design. Even while analysing the auditory data without any covariates added, the null-results were virtually identical.

Although the Repeated Measures MMA revealed a general learning effect from the first to the second threshold run, this tendency did certainly not apply to all subjects. For many of them, the first threshold was better than the second or third, or the second threshold was better than the third. Moreover, since we are interested in threshold estimations as an indicator of a subject's sensory capability, average threshold (or the average of the last two threshold runs) might not be the most appropriate measure; especially not in this age group that traditionally shows a high intrasubject variability (Wightman & Allen, 1992). Because our interest is in the best level of performance a subject is able to reach, a more reasonable estimator of threshold is each subject's "best" performance, or the lowest threshold of the three estimates for every experiment (see Wightman et al., 1989). While using a three-interval oddity paradigm, the probability of progressing to the next more difficult level just by chance is only 11.1% (1–9). This means that the probability of progressing two consecutive levels by guessing is very small (only 1.23%). Hence, on an average of 35 trials for every run and every experiment, it turns out to be very implausible that this "best threshold" would just be the result of lucky guessing instead of reflecting the real sensory capability limit. Threshold estimates and test statistics for the best and second best threshold and for the mean of the two best thresholds are given in Table 4. For every auditory experiment, Mixed Model Analysis showed there was no significant group effect.

Table 4
Auditory measures: descriptive statistics and p values for paired wise MMA, controlling for non-verbal IQ, age, and parental educational level

Measures	HR		LR		p
	M	SD	M	SD	
PTA (dB HL)	13.3	5.1	13.4	6.0	.94
Best GAP1 (ms)	4.2	2.4	4.2	3.1	.88
Best GAP2 (ms)	7.4	6.7	5.8	4.3	.30
AV GAP1/2 (ms)	5.8	4.3	5.0	3.6	.40
Best FM1 (Hz)	6.0	3.8	5.4	2.2	.91
Best FM2 (Hz)	10.4	6.2	9.0	6.3	.28
AV FM 1/2 (Hz)	8.2	4.6	7.2	4.0	.41
Best TN1 (dB SNR)	−8.0	2.3	−9.0	1.8	.71
Best TN2 (dB SNR)	−6.9	2.4	−7.5	1.4	.94
AV TN1/2 (dB SNR)	−7.5	2.3	−8.2	1.5	.81

Note. PTA, Pure Tone Average; BestGAP1-2, best and second best GAP threshold; AVGAP1/2, average of the two best GAP thresholds; BestFM1-2, best and second best FM threshold; AVFM1/2, average of the two best FM thresholds; BestTN1-2, best and second best TN threshold; and AVTN1/2, average of the two best TN thresholds.

Table 5
Spearman correlation coefficients: age, non-verbal intelligence, and auditory measures

	AVGAP1/2	AVFM1/2	AVTN1/2
Age	0.1	0.15	−0.13
Non-verbal IQ	−0.32*	−0.16	−0.18
AVGAP1/2		0.27*	0.36**
AVFM1/2			0.43***

Note. AVGAP1/2, average of the two best GAP thresholds in ms; AVFM1/2, average of the two best FM thresholds in Hz; AVTN1/2, average of the two best TN thresholds in SNR.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Spearman rank correlations between the best and the second best threshold estimate for every experiment appeared to be satisfactory and were $r_s = .79$ for GAP, $r_s = .75$ for FM, and $r_s = .83$ for TN, all being significant at the $p < .0001$ level.

Table 5 shows the Spearman rank interrelations between the different auditory psychophysical measures, and their relation to age and non-verbal IQ. None of the auditory tasks seemed to be related to age, and only GAP-detection showed a significant relation to non-verbal intelligence. All auditory tasks appeared to be significantly correlated with each other, with the relation between FM and TN being the most substantial one.

3.3. Relations between phonological and auditory skills

To analyse the relationship between participants' auditory processing skills and their phonological

Table 6
Spearman partial correlation coefficients, partialed out for age and non-verbal IQ

	Simple rhyme	Rhyme identity	First-sound identity	End-sound identity	Colour naming	Picture naming	Digit span	Non-word repetition test	Factor PhAW	Factor RAN	Factor VSTM	Letter knowledge
Total group (N = 61)												
AV GAP1/2												
AV FM 1/2	−0.33*	−0.34**	−0.43***	−0.45***		−0.24			−0.48****			−0.29*
AV TN1/2		−0.29*	−0.36**	−0.28*		−0.24			−0.35**			
HR (N = 30)												
AV GAP1/2	−0.36											
AV FM 1/2	−0.43*		−0.34	−0.33	−0.32	−0.41*			−0.34	−0.31		
AV TN1/2		−0.39*	−0.55**	−0.50**		−0.41*			−0.52**		0.36	
LR (N = 31)												
AV GAP1/2												
AV FM 1/2		−0.42*	−0.50**	−0.59***					−0.60***			
AV TN1/2												

Note. AVGAP1/2, average of the two best GAP thresholds in ms; AVFM1/2, average of the two best FM thresholds in Hz; AVTN1/2, average of the two best TN thresholds in SNR. Only correlations with a p value above .10 have been depicted.
* $p < .05$.
** $p < .01$.
*** $p < .001$.
**** $p < .0001$.

abilities and developing literacy skills, Spearman correlation coefficients were calculated between the participants' GAP, FM, and TN thresholds on the one hand, and the raw and combined phonological and letter knowledge scores on the other hand. Table 6 offers an overview of these correlations for total group (TG), HR-group and LR-group. Correlations have been partialed out for the possible influence of age and non-verbal IQ.

For the total group both FM- and TN-detection were significantly related to all variables measuring phonological awareness skills, and consequently they were also significantly related to the composite Phonological Awareness factor ($r_s = -.48$ and $-.35$, respectively). In contrast, GAP-detection was not related to any of the phonological variables. FM-detection was the only auditory variable being significantly related to letter knowledge.

For both groups separately, slightly different relationships could be observed. In the HR-group only TN-detection was significantly related to Phonological Awareness and its constituent subtasks, whereas in the LR-group FM-detection was exclusively related to this Phonological Awareness factor. Remarkably, in the HR-group both FM and TN-detection appeared to be significantly related to Picture Naming ($r_s = -.41$). The GAP-detection task again appeared to be unrelated to any of the phonological variables. In the subgroups, the power of the Spearman test was too weak to reveal a relation between letter knowledge and any of the auditory variables.

3.4. Individual deviance analysis

Since one of the goals of this study was to explore early indicators of dyslexia and in view of the fact that group comparisons might mask significant individual differences, we also carried out analyses on the subject level. To decide which individual did and did not show abnormal performance, we adopted the two-step criterion as suggested by Ramus et al. (2003). Applying this procedure, the criterion for deviance has been placed on 1.65 standard deviations of the mean of the LR-group. In a normal distribution, this corresponds to the fifth percentile and as such it is a fairly strict criterion. However, if a LR subject may occasionally show abnormal performance, this would make the criterion much more stringent by excessively influencing the LR mean and standard deviations. Moreover, the occurrence of these low scoring LR subjects might be especially probable in this preschool population since there is still a chance that even children of the LR group would become dyslexic. For this reason, the criterion has been applied in two steps: (1) compute the control mean and standard deviation and identify LR subjects who qualify for abnormal performance according to the 1.65 *SD* criterion (typically, this applied to 1 or 2 LR subjects for each measure); (2) recompute the LR mean and standard deviation excluding these deviant LR subjects, and identify HR subjects who are outside ±1.65 *SD*.

Individual scores for the psychophysical measures (average of best and second best threshold) and for the phonological factors are plotted in Fig. 1. A distribution analysis on the data of the 'restricted' LR-group confirmed the normality of these variables. As such, the 1.65 *SD* deviance criterion corresponded indeed to the postulated fifth percentile. A deviance analysis on Phonological Awareness revealed that 12 HR subjects out of 30 showed abnormal performance, this corresponds to 40% of the HR group. In the LR group 5 subjects out of 31 (=16%) had abnormal performance. For Rapid

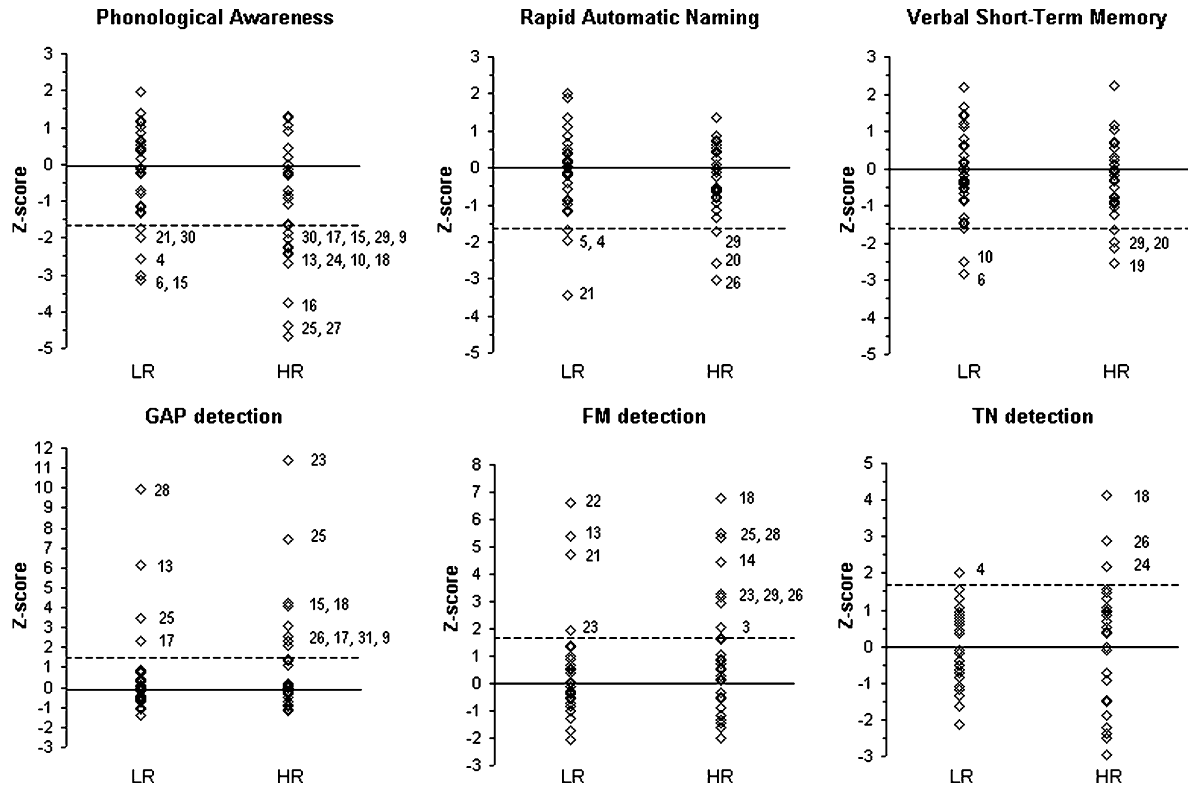

Fig. 1. Individual Z scores on phonological and auditory measures. The solid line indicates the mean for all LR subjects above Pc 5; the dashed line indicates the chosen deviance criterion (1.65 *SD* deviating of the LR mean after excluding deviant LR subjects). Deviant individuals are identified by their pair number.

Automatic Naming 3 subjects out of each group had abnormal performance (corresponding to 10%) and for Verbal STM 3 HR versus 2 LR subjects showed deviant scores (10% versus 6%). Considering the auditory measures, for both FM and GAP detection there were 8 HR subjects versus 4 LR subjects who showed abnormal performance (26% versus 13%). For TN there were 3 HR subjects (10%) versus 1 LR subject (3%) having deviant results.

To summarize, again Phonological Awareness turned out to discriminate best between both groups by having a significant higher proportion of deviant subjects in the HR group (Fisher Exact Test, $p = .04$). For both GAP detection and FM detection the proportion of subjects showing abnormal performance was twice as high in the HR group compared to the LR group. However, this tendency was not significant (Fisher Exact Test, $p = .17$ for both tests).

By looking at the individual scores for all subjects showing abnormal performance in at least one measure, it becomes clear that there is no straightforward regularity or tendency between these measures. Auditory deficits appeared to be largely mutually unrelated with some subjects differing only on one task, others on two or three tasks, but without any consistency. The same applied for deficits on the phonological factors. Moreover, the relation between auditory deficiencies on the one hand and phonological deficiencies (more specifically in phonological awareness) on the other hand was even more ambiguous. Some subjects suffered really serious deficits in auditory processing without showing any phonological problems. Conversely, other subjects obtained deviant phonological results while demonstrating perfectly intact auditory processing. Finally, in some subjects phonological and auditory processing deficits appeared to be partially related.

4. Discussion

4.1. Feasibility of psychophysical testing in preschoolers

One of the main objectives of this study was to explore the feasibility of administering complex psychophysical tests to very young subjects. Based on our results, this research question can be answered entirely confirmative. Not only did the children perform surprisingly accurately, but they also really enjoyed the auditory tasks.

While comparing the auditory thresholds of our LR-subjects with results on identical tasks administered to 11-year-old normal reading children (Van Ingelghem et al., 2004), the preschoolers showed an overall weaker discrimination capacity (GAP: 4.2 versus 2.1 ms, FM: 5.4

versus 3.3 Hz, and TN: −9.0 versus −12.7 dB SNR). The higher thresholds of preschoolers compared to older children and adults is a general observation that has been demonstrated in numerous auditory studies across all kind of discrimination tasks (e.g., Allen et al., 1989; Irwin, Ball, Kay, Stillman, & Rosser, 1985; Jensen & Neff, 1993; Morrongiello, Kulig, & Clifton, 1984; Trehub, Schneider, & Henderson, 1995; Schneider & Trehub, 1992). However, it remains unclear whether this weaker discrimination reflects an underlying sensory immaturity or has a cognitive origin (e.g., non-optimal listening strategies, fluctuations in attention, etc.). Another general observation in psychophysical testing with preschoolers concerns the higher inter- and intrasubject variability. As Wightman and colleagues demonstrated, the intrasubject variability might mainly result from central non-auditory attentional factors (e.g., Oh, Wightman, & Lutfi, 2001; Wightman et al., 1989; Wightman, Callahan, Lutfi, Kistler, & Oh, 2003). For this reason, the way we reduced this excessive variability by opting for a 'best threshold analysis,' can be justified. In contrast, the intersubject variability is not a result of error variance, but reflects reliable differences in the speed of neural development (e.g., Allen & Wightman, 1994). Because of this substantial intersubject variability caution is required while interpreting averages and group performances (Wightman & Allen, 1992).

4.2. Phonological abilities and letter knowledge

With respect to the phonological data, the exploratory factor analysis convincingly revealed the three-dimensional phonological structure as postulated by Wagner and Torgesen (1987). Moreover, both the group analysis and the individual deviance analysis clearly demonstrated the robustness of the phonological deficit hypothesis in dyslexia-prone children. Even at a preschool age, the HR-children already showed a significant deficit in phonological awareness, not only at the rhyme-level but also at the level of the phonemes. These results are consistent with other longitudinal prospective studies that revealed similar deficits in genetically at risk children (e.g., Elbro, Borstrom, & Petersen, 1998; Gallagher, Frith, & Snowling, 2000; Pennington & Lefly, 2001; Scarborough, 1989, 1990, 1998). The group differences on the factors rapid automatic naming and verbal short-term memory were insignificant but in the expected direction with the HR-group scoring less well than the LR-group. These results are in line with findings by Elbro et al. (1998). However, some researchers did find significant differences on verbal short-term memory (Pennington & Lefly, 2001—significant group difference in first grade, but not in kindergarten) and rapid automatic naming (de Jong & van der Leij, 2003; Pennington & Lefly, 2001), but only while comparing HR *dyslexic* subjects versus LR *normal* readers. This means they retrospectively reanalysed the kindergarten data after having diagnosed their subjects in second or third grade. Since currently the children in our study are still attending kindergarten, we do not yet know who will finally become dyslexic, and consequently we obviously cannot carry out this analysis yet.

Although most children hardly knew any letters at the beginning of the last year in kindergarten (on average about three or four letters), the group difference was already significant. Again this result is in line with any of the previously mentioned longitudinal studies. Since both letter knowledge and phonological awareness have consistently been proven to be among the best single preschool predictors of literacy development (see, e.g., the impressive meta-analysis of Scarborough, 1998; based on 61 studies), and since the HR versus LR group differed especially on these measures, it is likely that the familial high risk group will contain a disproportionally high number of future cases of dyslexia.

4.3. Auditory processing skills

We studied GAP- and FM-detection in order to investigate auditory temporal ('rapid and brief' versus 'dynamic') processing. In line with the temporal hypothesis we expected these tasks to differentiate between both risk groups and to be related to specific pre-reading skills like phonological processing and letter knowledge. However, we did not find any significant differences in auditory processing between the HR and the LR group, neither at a group level nor in individual deviance analyses. Although there were twice as many subjects showing abnormal performance for GAP- and FM-detection in the HR-group, this tendency did not reach significance. Assuming the correctness of the causal auditory hypothesis, this lack of significance might be attributed either to the typically greater interindividual variability in children (cfr. supra) or to the fact that we did not study a well-defined clinical group but only a risk group that still might show substantial overlap with the non-affected control group. Moreover, Bishop et al. (1999) demonstrated in a twin study on SLI children that in contrast to the highly heritable phonological skills, auditory skills depend less on genetic and more on environmental influences. As such, our finding of a phonological deficit in combination with relatively intact auditory skills in this genetic high risk group corresponds well with the results of Bishop and colleagues.

Recently, an alternative explanation has been put forward to explain the variably observed auditory deficits in subjects with specific language impairment and dyslexia. Bishop and colleagues (Bishop & McArthur, 2004, 2005; McArthur & Bishop, 2004) and Wright and Zecker (2004) suggested that these subjects might not suffer from a specific chronic auditory deficit, but rather from a more general and passing auditory maturational delay (estimated to encompass about three or four years). In that perspective the differential sensitivity of a task will be at best when administered in the age period that the measured skill is sharply improving in typically developing subjects. However, in spite of having administered our auditory tasks during a sensitive developmental period—it has been demonstrated that normally developing five-year-old subjects undergo a rapid maturation of both spectral and temporal auditory abilities[4] (see, e.g., Irwin et al., 1985; Jensen & Neff, 1993; Thompson et al., 1999; Wightman et al., 1989)—we did not observe the hypothesized developmental delay in the HR group. This is markedly in contrast with results obtained by Hautus et al. (2003), who demonstrated that younger reading-impaired children (aged 6–9 years) had significantly higher gap-detection thresholds than age-matched controls while older dyslexic subjects (aged 10–13 and adults) did no longer differ from controls.

4.4. Relations between auditory and phonological abilities

With regard to relations between auditory processing and phonological processing, both FM- and TN-detection thresholds were significantly related to phonological awareness and—to some extent—to rapid automatic naming and letter knowledge. In contrast, the GAP-detection task was completely unrelated to any phonological measure. These results suggest that it is not the specific temporal aspect of auditory processing that is related to developing phonological abilities, since the temporal GAP-detection task was not related whereas the non-temporal TN-detection task turned out to be significantly related to phonological processing. Instead it appears as if the common spectral or frequency sensitive characteristic of TN- and FM-detection causes the relation with phonological ability. At a neurophysiologic level this might imply that a more accurate phase-locking system or smaller and more sharply tuned auditory filters are somehow related to better phonological processing (see, e.g., Carney, Heinz, Evilsizer, Gilkey, & Colburn, 2002; Moore, 1997). This is in line with many studies demonstrating an impairment in frequency discrimination in dyslexic subjects (e.g., Ahissar, Protopapas, Reid, & Merzenich, 2000; Amitay, Ben-Yehudah, Banai, & Ahissar, 2002; Fischer & Hartnegg, 2004; Cacace, McFarland, Ouimet, Schrieber, & Marro, 2000). Talcott et al. (2002) reported similar observations from a large-scale primary school study in which auditory frequency resolution differed between groups of children with different literacy skills. In the same way, Hulslander et al. (2004) observed a significant correlation between FM-thresholds and scores on a phoneme awareness composite, even while controlling for individual differences for full-scale IQ. Evidence supporting the neurophysiologic explanation for a deficit in frequency sensitivity has been put forward by McAnally and Stein (1996) who demonstrated that dyslexics were less able to generate neural discharges phase-locked to the temporal fine structure of the acoustic stimuli. Data consistent with these findings were also reported by Baldeweg, Richardson, Watkins, Foale, and Gurzelier (1999); Dougherty, Cynader, Bjornson, Edgell, and Giaschi (1998) and Schulte Korne et al. (1998). Recently, Amitay, Ahissar, and Nelken (2002) also found evidence for a deficit in tone-in-noise detection in adult dyslexic subjects. However, their results were conflicting with McAnally's findings of a phase-locking deficit in dyslexics (for a similar conclusion see also Hill, Bailey, Griffiths, & Snowling, 1999).

The significant relation we observed between the auditory measures and phonological awareness is in line with the results of Share et al. (2002) who also found a reliable *concurrent* correlation between a non-linguistic TOJ task and phoneme segmentation at school entry. Unfortunately, this relation could not be interpreted in a causal way since the auditory measures were not able to *predict* any later phonological or reading skills. Instead, Share and colleagues speculated that the association between early temporal processing and phonological awareness might be the result of a higher-order common cause of an unspecified metalinguistic nature. Similarly, the substantial concurrent correlation we found between phonological and spectral auditory measures should not be interpreted in a directional or causal way. After all, while inspecting individual data in the deviance analysis, it is clear that there is no obvious straightforward relation between deficits in auditory measures and deficits in phonological skills. Although deviant auditory processing tends to be a risk indicator for a deficit in phonological development, intact auditory processing is certainly not a sufficient prerequisite for developing normal phonological skills. More generally, we have to conclude that we are not able to demonstrate a consistent and convincing pattern in individual deficiencies, neither within the auditory skills, nor within the phonological sub-skills, nor within the relation between auditory and phonological skills.

5. Conclusion

To conclude, phonological awareness and letter knowledge turn out to be the best indicators to differen-

[4] For instance, regarding frequency discrimination, an almost linear improvement has been observed from 4 to 9 years, after which adult levels of performance are obtained (Jensen & Neff, 1993; Thompson, Cranford, & Hoyer, 1999). Similarly, with respect to GAP-detection thresholds, Wightman et al. (1989) and Irwin et al. (1985) observed a sharp maturational improvement from 3 years on up to 7 years and extending slightly up to 11 years before reaching adult levels.

tiate between preschool children with low versus high familial risk of developing dyslexia. In contrast, none of the auditory processing tasks is able to differentiate significantly between both groups. However, auditory spectral tasks (FM and TN-detection thresholds) are highly significantly related to phonological awareness. This relation is not present for a specific temporal GAP-detection task. Nevertheless, identifying deviant subjects in auditory spectral processing in order to predict deficiencies in phonological skills and subsequent reading development does not yet seem to be a viable option, since at the level of individual subjects the relation between auditory and phonological skills seems to be much less straightforward.

Acknowledgments

This research was financed by the Queen Fabiola Fund and the Fund for Scientific Research—Flanders (Belgium), Grant G.0216.02. We thank the children and school staff who were involved in this study, Caroline Van Eccelpoel, Sandra Schoonbrood, and Leen Vanrie for assistance with testing, and we especially show gratitude to Erik Vandenbussche who was one of the inspirators of this project but unfortunately deceased in December 2002. Furthermore, we thank Mieke Van Ingelghem for stimulating discussion, Geert Verbeke for advice about statistical modeling, Willy Leung for creating the animation movies, Johan Laneau for setting up the software for the test platform, and Frederic Wightman and two anonymous reviewers for helpful comments on previous drafts of this paper.

References

Adlard, A., & Hazan, V. (1998). Speech perception in children with specific reading difficulties (dyslexia). *Quarterly Journal of Experimental Psychology, 51*(1), 153–177.

Ahissar, M., Protopapas, A., Reid, M., & Merzenich, M. (2000). Auditory processing parallels reading abilities in adults. *Proceedings of the National Academy of Science of the United States of America, 97*(12), 6832–6837.

Allen, P., Wightman, F., Kistler, D., & Dolan, T. (1989). Frequency resolution in children. *Journal of Speech and Hearing Research, 32*(2), 317–322.

Allen, P., & Wightman, F. (1994). Psychometric functions for children's detection of tones in noise. *Journal of Speech and Hearing Research, 37*(1), 205–215.

Amitay, S., Ahissar, M., & Nelken, I. (2002). Auditory processing deficits in reading disabled adults. *Journal of the Association for Research in Otolaryngology, 3*, 302–320.

Amitay, S., Ben-Yehudah, G., Banai, K., & Ahissar, M. (2002). Disabled readers suffer from visual and auditory impairments but not from a specific magnocellular deficit. *Brain, 125*(10), 2272–2285.

Baldeweg, T., Richardson, A., Watkins, S., Foale, C., & Gurzelier, J. (1999). Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials. *Annals of Neurology, 45*(4), 495–503.

Benasich, A. A., & Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural Brain Research, 136*, 31–49.

Bishop, D. V., Bishop, S. J., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: Evidence from a twin study. *Journal of Speech, Language and Hearing Research, 42*(1), 155–168.

Bishop, D. V. M., & McArthur, G. M. (2004). Immature cortical responses to auditory stimuli in specific language impairment: Evidence from ERPs to rapid tone sequences. *Developmental Science, 7*(4), F11–F18.

Bishop, D. V. M., & McArthur, G. M. (2005). Individual differences in auditory processing in specific language impairment: A follow-up study using event-related potentials and behavioural thresholds. *Cortex, 41*(3), 327–341.

Cacace, A. T., McFarland, D. J., Ouimet, J. R., Schrieber, E. J., & Marro, P. (2000). Temporal processing deficits in remediation-resistant reading-impaired children. *Audiology and Neurootology, 5*(2), 83–97.

Carney, L. H., Heinz, M. G., Evilsizer, M. E., Gilkey, R. H., & Colburn, H. S. (2002). Auditory phase opponency: A temporal model for masked detection at low frequencies. *Acta Acustica, 88*, 334–347.

de Jong, P. F., & van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91*(3), 450–476.

de Jong, P. F., Seveke, M. J., & van Veen, M. (2000). Phonological sensitivity and the acquisition of new words in children. *Journal of Experimental Child Psychology, 76*(4), 275–301.

de Jong, P. F., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology, 95*(1), 22–40.

Dougherty, R. F., Cynader, M. S., Bjornson, B. H., Edgell, D., & Giaschi, D. E. (1998). Dichotic pitch: A new stimulus distinguishes normal and dyslexic auditory function. *Neuroreport, 9*(13), 3001–3005.

Elbro, C., Borstrom, I., & Petersen, D. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33*(1), 36–60.

Elbro, C., & Scarborough, H. S. (2003). Early identification. In P. Bryant & T. Nunes (Eds.), *Handbook of children's reading* (pp. 339–359). Dordrecht: Kluwer Academic Publishers.

Fischer, B., & Hartnegg, K. (2004). On the development of low-level auditory discrimination and deficits in dyslexia. *Dyslexia, 10*(2), 105–118.

Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry, 41*(2), 203–213.

Gathercole, S., & Pickering, S. (2000). Assessment of working memory in six- and seven-year-old children. *Journal of Educational Psychology, 92*, 377–390.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory, 2*(2), 103–127.

Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1991). Risk for reading disability as a function of parental history in three family studies. *Reading and Writing, 3*(3-4), 205–217.

Hautus, M. J., Setchell, G. J., Waldie, K. E., & Kirk, I. J. (2003). Age-related improvements in auditory temporal resolution in reading-impaired children. *Dyslexia, 9*, 37–45.

Heath, S. M., & Hogben, J. H. (2004). Cost-effective prediction of reading difficulties. *Journal of Speech, Language and Hearing Research, 47*, 751–765.

Hill, N. I., Bailey, P. J., Griffiths, Y. M., & Snowling, M. J. (1999). Frequency acuity and binaural masking release in dyslexic listeners. *Journal of the Acoustical Society of America, 106*(6), 53–58.

Hulslander, J., Talcott, J., Witton, C., DeFries, J., Pennington, B., Wadsworth, S., et al. (2004). Sensory processing, reading, IQ and attention. *Journal of Experimental Child Psychology, 88*, 274–295.

Irwin, R. J., Ball, A. K., Kay, N., Stillman, J. A., & Rosser, J. (1985). The development of auditory temporal acuity in children. *Child Development, 56*(3), 614–620.

Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science, 4*(2), 104–107.

Laneau, J., Boets, B., Moonen, M., van Wieringen, A., & Wouters, J. (2005). A flexible auditory research platform using acoustic or electric stimuli for adults and young children. *Journal of Neuroscience Methods, 142*(1), 131–136.

Lefly, D. L., & Pennington, B. F. (2000). Reliability and validity of the adult reading history questionnaire. *Journal of Learning Disabilities, 33*(3), 286–296.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467–477.

Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: Sas Institute Inc.

Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T. K., Laakso, M. L., Leinonen, S., et al. (2001). Developmental pathways of children with and without familial risk for dyslexia during the first years of life. *Developmental Neuropsychology, 20*(2), 535–554.

McAnally, K. I., & Stein, J. F. (1996). Auditory temporal coding in dyslexia. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 263*, 961–965.

McAnally, K. I., & Stein, J. F. (1997). Scalp potentials evoked by amplitude-modulated tones in dyslexia. *Journal of Speech Language and Hearing Research, 40*(4), 939–945.

McArthur, G. M., & Bishop, D. V. M. (2004). Which people with specific language impairment have auditory deficits. *Cognitive Neuropsychology, 21*(1), 79–94.

McBride Chang, C. (1996). Models of speech perception and phonological processing in reading. *Child Development, 67*(4), 1836–1856.

McCroskey, R. L., & Kidder, H. C. (1980). Auditory fusion among learning disabled, reading disabled, and normal children. *Journal of Learning Disabilities, 13*(2), 69–76.

Menell, P., McAnally, K. I., & Stein, J. F. (1999). Psychophysical sensitivity and physiological response to amplitude modulation in adult dyslexic listeners. *Journal of Speech, Language and Hearing Research, 42*(4), 797–803.

Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and Language, 72*(3), 238–245.

Moore, B. C. J. (1997). *An introduction to the psychology of hearing*. London: Academic Press.

Morrongiello, B. A., Kulig, J. W., & Clifton, R. K. (1984). Developmental changes in auditory temporal perception. *Child Development, 55*(2), 461–471.

Oh, E. L., Wightman, F. L., & Lutfi, R. A. (2001). Children's detection of pure-tone signals with random multitone maskers. *Journal of the Acoustical Society of America, 109*, 2888–2895.

Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development, 72*(3), 816–833.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., et al. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*(4), 841–865.

Raven, J. C., Court, J. H., & Raven, J. (1984). *Manual for Raven's progressive matrices and vocabulary scales*. London: Lewis.

Rocheron, I., Lorenzi, C., Fullgrabe, C., & Dumont, A. (2002). Temporal envelope perception in dyslexic children. *Neuroreport, 13*(13), 1683–1687.

Rolf, P. C., & Van Rijnsoever, R. J. (1984). *Positionele letterfrekwenties van het Nederlands*. Lisse: Swets and Zeitlinger.

Scarborough, H. S. (1989). Prediction of reading disability from familial and individual differences. *Journal of Educational Psychology, 81*(1), 101–108.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development, 61*(6), 1728–1743.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In P. J. Accardo, A. J. Capute, & B. K. Shapiro (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–119). Timonium, MD: York Press.

Scheltinga, F. (2003). *The Dutch Nonword Repetition Test*. Unpublished manuscript, University of Amsterdam.

Schneider, B. A., & Trehub, S. E. (1992). Sources of developmental change in auditory sensitivity. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 3–46). Washington: American Psychological Association.

Schulte Korne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1998). Role of auditory temporal processing for reading and spelling disability. *Perceptual and Motor Skills, 86*(3), 1043–1047.

Share, D. L., Jorm, A. F., Maclean, R., & Matthews, R. (2002). Temporal processing and reading disability. *Reading and Writing, 15*(1–2), 151–178.

Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Malden, MA: Blackwell Publishers.

Soderquist, D. R., & Shilling, R. D. (1992). Forward masking in young children—Rocket ship psychophysics. *Current Psychology: Research and Reviews, 11*(2), 145–156.

Stein, J. F., & McAnally, K. (1995). Auditory temporal processing in developmental dyslexics. *Irish Journal of Psychology, 16*(3), 220–228.

Studdert Kennedy, M., & Mody, M. (1995). Auditory temporal perception deficits in the reading-impaired: A critical review of the evidence. *Psychonomic Bulletin & Review, 2*(4), 508–514.

Talcott, J. B., Witton, C., McClean, M., Hansen, P. C., Rees, A., Green, G., et al. (1999). Can sensitivity to auditory frequency modulation predict children's phonological and reading skills?. *Neuroreport, 10*(10), 2045–2050.

Talcott, J. B., Witton, C., McLean, M. F., Hansen, P. C., Rees, A., Green, G. G., et al. (2000). Dynamic sensory sensitivity and children's word decoding skills. *Proceedings of the National Academy of Science of the United States of America, 97*(6), 2952–2957.

Talcott, J. B., Witton, C., Hebb, G. S., Stoodley, C. J., Westwood, E. A., France, S. J., et al. (2002). On the relationship between dynamic visual and auditory processing and literacy skills; results from a large primary-school study. *Dyslexia, 8*(4), 204–225.

Tallal, P., & Piercy, M. (1973). Defects of nonverbal auditory perception in children with developmental aphasia. *Nature, 241*, 468–469.

Tallal, P., & Piercy, M. (1975). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia, 13*, 69–74.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language, 9*(2), 182–198.

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences, 682*, 27–47.

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

Trehub, S. E., Schneider, B. A., & Henderson, J. L. (1995). Gap detection in infants, children, and adults. *Journal of the Acoustic Society of America, 98*(5 Pt. 1), 2532–2541.

van den Bos, K. P., Zijlstra, B. J. H., & Spelberg, H. C. (2002). Life-span data on continuous-naming speeds of numbers, letters, colors, and pictured objects, and word-reading speed. *Scientific Studies of Reading, 6*(1), 25–49.

Van Ingelghem, M., van Wieringen, A., Wouters, J., Vandenbussche, E., Onghena, P., & Ghesquiere, P. (2001). Psychophysical evidence for a general temporal processing deficit in children with dyslexia. *Neuroreport, 12*(16), 3603–3607.

Van Ingelghem, M., Boets, B., van Wieringen, A., Ghesquiere, P., & Wouters, J. (2004). Auditory and visual temporal processing in dyslexic and normal reading children matched for sex, age and intellectual ability. In Paper presented at the 6th BDA International conference, 27–30 March 2004, University of Warwick, UK.

Van Ingelghem, M., Boets, B., van Wieringen, A., Onghena, P., Ghesquiere, P., & Wouters, J. (2005). An auditory temporal processing deficit in children with dyslexia. In P. Ghesquiere & A. J. J. M. Ruijssenaars (Eds.), *Children with learning disabilities: A challenge to teaching and instruction Series: Studia Paedagogica* (pp. 47–63). Leuven: University Press.

van Wieringen, A., & Wouters, J. (1999). Gap detection in single- and multiple- channel stimuli by LAURA cochlear implantees. *Journal of the Acoustical Society of America, 106*, 1925–1939.

Verbeke, G., & Lesaffre, E. (1997). The effect of misspecifying the random-effects distribution in linear mixed models for longitudinal data. *Computational Statistics and Data Analysis, 23*, 541–556.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*(2), 192–212.

Wechsler, D. (1992). *Wechsler intelligence scale for children* (3rd ed.). London: The Psychological Corporation.

Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child Development, 60*(3), 611–624.

Wightman, F., & Allen, P. (1992). Individual differences in auditory capability among preschool children. In L. A. Werner & E. W. Rubel (Eds.), *Developmental Psychoacoustics* (pp. 113–134). Washington DC: American Psychological Association.

Wightman, F. L., Callahan, M. R., Lutfi, R. A., Kistler, D. J., & Oh, E. (2003). Children's detection of pure-tone signals: Informational masking with contralateral maskers. *Journal of the Acoustical Society of America, 113*(6), 3297–3305.

Witton, C., Talcott, J. B., Hansen, P. C., Richardson, A. J., Griffiths, T. D., Rees, A., et al. (1998). Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Current Biology, 8*(14), 791–797.

Wright, B. A., & Zecker, S. G. (2004). Learning problems, delayed development and puberty. *Proceedings of the National Academy of Science of the United States of America, 101*(26), 9942–9946.