SPEECH PERCEPTION IN DYSLEXIA: MEASUREMENTS FROM BIRTH ONWARDS

Florien J. Koopmans-van Beinum, Caroline E. Schwippert* & Cecile T.L. Kuijpers***

* University of Amsterdam, Institute of Phonetic Sciences/IFOTT, The Netherlands
** Catholic University of Nijmegen, IWTS, The Netherlands

ABSTRACT

This paper concentrates on a small, but essential part within a large national Dutch research program on developmental dyslexia, namely the development of auditory test material for experiments with children from birth onwards. Since it is likely that the basis of any phoneme awareness in children is laid already in their first year of life, it is of great importance to follow the perceptual development of children at risk of dyslexia from birth onwards. In order to investigate the nature and origin of the perceptual deficit dyslexic people are afflicted with, a number of auditory tests are designed. The various steps in the development of the definite set of auditory tests, to be used in the actual research program, the rationale underlying these various steps, and first results of the pilot tests are described in the present paper.

1. INTRODUCTION

Within the framework of a national Dutch research program ("Identifying the Core Features of Developmental Dyslexia: A Multidisciplinary Approach") 300 Dutch children born with a genetic risk for dyslexia and 120 control children will be tested longitudinally during ten years from birth onwards, with respect to a number of medical, neurophysiological, visual, auditory, and linguistic aspects. Since no auditory perception tests are available that are suitable for very young Dutch children and that can be used (probably in adapted form) during the whole period of ten years for the children as well as for adults, we are faced with the need to develop a set of tests to measure the auditory sensitivity of the subjects to speechlike stimuli.

Literature on the possible deficiency in auditory perception in dyslexic children points into the direction of less distinct phoneme boundaries: classification as well as discrimination tests show lower consistency in dyslexic than in control children [3, 4, 6]. Especially in (synthetic) speech sound continua for place of articulation, where normally categorical perception can be demonstrated, this lower consistency is obvious. Results so far have been interpreted quite differently: at the one hand the phenomenon is explained as a general deficit in auditory temporal perception [5], at the other hand as being phonetic in origin for those contrasts that are acoustically similar [2]. The latter authors showed that syllables that can be discriminated more easily, made differences between good and poor readers disappear.

In order to link the development of our auditory tests with this ongoing discussion, we designed sets of stimuli that had to meet a number of requirements:

- based on natural speech
- spoken by a female speaker
- three continua: 'basic': /bɑ-dɑ/, 'difficult': /bɑ-wɑ/, 'easy': /bɑ-zɑ/
- by means of some kind of acoustical interpolation method
- with equal duration of all stimuli
- with varying interstimulus intervals (ISI)
- to be used in a classification test
- to be used in a discrimination test (AX-paradigm)
- combined with reaction time measurements
- supplemented with ERP measurements
- to be tested on adults first, but to be adapted for infants and young children at a later stage.

2. METHOD

2.1. Stimulus materials

Choice of stimulus material. Natural speech of a female speaker was considered to be the best starting point to construct the stimuli, since the ultimate use of the stimuli will be in tests with infants and children, and the stimuli have to resemble everyday speech as much as possible. Therefore speech material of an adult female speaker was recorded on a DAT-recorder in a anechoic room. She pronounced a series of one-syllable words, among which the target words /bɑk/, /dɑk/, /zɑk/, and /wɑk/, all being normal Dutch words, meaning *'tray'*, *'roof'*, *'bag'*, and *'hole in the ice'*, respectively. Our selection of the three continua that might differ in ease of discrimination was based on results of Dutch consonant confusion matrices. The selection of acceptable interpolation methods was one of the main subjects of our study and is described in detail below. Since we tried to avoid durational differences in the stimuli to be fused with difficulties in temporal auditory processing, we prefered all stimuli to have equal duration. By means of varying the interstimulus interval (ISI) and by measuring reaction times we supposed to get information on the temporal processing.

The /bɑk/-/dɑk/ continuum. The acoustical analysis data of the words /bɑk/ and /dɑk/ was our starting point for further manipulations. The original F_2 onset values of the transitions were about 1100 and 1800 Hz, respectively, whereas F_3 did not show too much difference. The F_0 range was between 125 and 235 Hz, mean F_0 being 165 Hz in both words. The realization of the word /bɑk/ was selected as the starting stimulus signal for the /bɑk-dɑk/ continuum. It was manipulated with the Praat software package [1]. The signal was down-sampled to 11025 Hertz to be able to analyse it with Linear Predictive Coding (LPC). The LPC-analysis was done with 10 linear prediction parameters, window width being 25 ms, time step 5 ms, and pre-emphasis frequency being 50 Hz. An interpolation of the F_2 was done which always resulted in a linear transition. The manipulated part of the signal was a 100 ms' interval at the beginning of the vowel. Number of interpolation steps was 10 and the F_2 onset ranged from 1100 to 1800 Hz. The occlusion period of [k] had to be reduced with 40 ms to 55 ms to obtain equal stimulus duration for all three continua. This did not influence the quality of the signal. The interpolation and additional processing resulted in a 10 point continuum, the total length of each item being 600 ms, consisting of (a) a vocal murmur 170 ms; (b) burst 10 ms; (c) vowel [ɑ] 250 ms, divided into 100 ms transition duration and 150 ms steady state; (d) occlusion period [k] (silence) 55 ms; (e) release [k] 115 ms. The [bɑk]-starting point was characterized by a constant level of the F_2 at 1100 Hz. For each intermediate signal, the F_2 onset was gradually situated at higher frequencies, making the fall of the transition steeper in every step. At the [d]-endpoint of the continuum the transition was falling from 1800 Hz to 1100.

The /bɑk/-/zɑk/ continuum. The perceptual contrast between the Dutch phonemes /b/ and /z/ is very large: they are seldom confused. One might even question whether one can speak of an actual /b/-/z/ continuum. Differences in place and manner of articulation combined with a completely different temporal structure of the phoneme seem to be difficult to overcome. First of all, the same formant manipulation was performed as for the /bɑk/-/dɑk/ continuum since [b] and [z] show the same F_2 contrast as [b] and [d], due to a comparable difference in place of articulation. Secondly interpolation was necessary between the burst of the [b] and the voiced fricative noise of [z]. So we omitted the first part (the vocal murmur, the burst and 2 vocalic periods) and replaced it by the 170 ms fricative part of a /zɑk/ utterance at a very low intensity at the /bɑk/-endpoint and at normal intensity near the /zɑk/-endpoint. The intermediate forms of the fricative part were made by increasing the intensity by 3 dB in every step going from /bɑk/ to /zɑk/. The only stimulus without any fricative noise added was the /bɑk/-endpoint, this was necessary to obtain a high quality realization of /bɑk/.

In summary two parameters were manipulated going from /bɑk/ to /zɑk/ in ten steps: F_2 transition onset increased with approximately 80 Hz and fricative noise intensity doubled in each step. This resulted in signals with a total length of 600 ms, consisting of (a) a fricative noise 170 ms; (b) vowel [ɑ] 240 ms, devided into 90 ms transition duration and 150 ms steady state; (c) occlusion period [k] (silence) 75 ms; (d) release [k] 115 ms.

The /bɑk/-/wɑk/ continuum. It is important to note that the Dutch contrast /b/-/w/ is different from the English one. In Dutch the initial /w/ is a labio-dental approximant, whereas it is a bilabial approximant in English. Therefore an F_1 manipulation can bridge the difference in English between these two phonemes, which in Dutch is not the case. An analysis of our speaker's utterances /bɑk/ and /wɑk/ revealed that changing the F_1 frequencies could not lead to the desired results. To construct a Dutch /b/-/w/ continuum, a different approach was needed. Cutting off increasingly bigger parts of the [√] did produce some of the desired effect, but did not lead to a very acceptable /bɑk/ realization. Furthermore this strategy would result in a continuum where each item would have a different length, which we wanted to avoid. We then adopted the following method: the approximant was cut off from the vowel at point "A" in the waveform and was replaced by the vocal murmur and burst of the /bɑk/ utterance, which resulted in a very natural /bɑk/ realization. The procedure now to create the intermediate signals was to shift point "A" increasingly to the left, first deleting the burst and inserting a period of the original [√] signal, then replacing gradually every period of the vocal murmur of [b] by a period of the original waveform of [√]. This procedure yielded 16 steps. Since we only needed 10 we omitted the four pre-last items at the [√]-endpoint, for they were all clear [√] items, and two items at the [b]-endpoint for the same reason. The 10 resulting items formed a very smooth continuum from [b] to [√]. To make the signals equally long as those of the other continua, some silence was added to the occlusion period, resulting for each signal in a total length of 600 ms, consisting of (a) a voiced labial consonant 180 ms; (b) vowel [ɑ] 250 ms; (c) occlusion period [k] (silence) 55 ms; (d) release [k] 115 ms.

2.2. Perceptual tests

Design. Apart from a screening test for dyslexia, the experimental design consisted of a discrimination test and a classification test. In the discrimination task two ISI's were used (25 and 400 ms). The three continua consisted of 7 stimulus pairs each. The classification test consisted of 10 stimuli for each of the three continua with an intertrial interval of 1500 ms. Details of both tasks are given below.

Subjects. Twelve adult dyslexic subjects (8 male and 4 female) and twelve adult control subjects (4 male and 8 female) participated in the listening tests. Except for 3 dyslexics all subjects were students; they were paid for their participation. None of the subjects reported auditory problems.

All subjects were administered a battery of tests representing a subset of reading- and spelling-relevant skills. Single-word reading was measured by two standardized Dutch reading tests (EMT) and (DMT). These tests consist of cards with real words of increasing difficulty. Both appeal to the lexical decoding skill. Phonological decoding skill was measured by two pseudoword reading tests; a standardized pseudoword reading test which is based on the EMT, and a pseudoword reading test constructed for this purpose, based on the DMT. The measures of single-word spelling were a dictation of 72 isolated real words, and a dictation of 72 isolated pseudowords having the same CV structure as the real words. Phonological awareness was tested by means of a nonword repetition task (40 items). Rapid automatized naming was measured by color, letter, digit, and object naming tests. Eight of the dyslexic subjects performed these tests very poorly, the remainig four performed moderately; the control subjects had no problems.

Experimental procedure. The screening test took three quarters of an hour, the perception test about one hour. Subjects began either with the screening test or with the perception tests. The perception tests were conducted in a room with three sound-attenuated booths. Each booth contained a computer screen, headphones, and a panel of buttons that the subjects had to press to indicate their responses. The buttons were labeled with words and with a corresponding picture or symbol, to avoid lexical confusions. The control panels for the experimental sessions were situated next to the booths. The experimental system NESU was used for real-time stimulus presentation and reaction time registration. The stimuli were presented through headphones. The computer screen in the booth was only used to indicate the beginning of a new block, a time-out, or the end of the experiment. The experimenter could watch the answers given by the subjects, and their reaction times. Each session began with a same-different discrimination task followed by a forced choice classification task. For both discrimination and classification the three continua were presented separately. The order of presentation of the continua was balanced across subjects.

Discrimination. The discrimination task required subjects to discriminate between two stimuli that were always three continuum steps apart (e.g. 1-4, 2-5, 3-6, etc.). The 7 stimulus pairs were presented 12 times. The internal order of a stimulus pair was balanced. In addition to the different trials, ten same trials (1-1, 2-2, etc.) were presented, each pair twice. The stimulus pairs were presented in four blocks of 52 stimuli (6x7 + 1x10), interstimulus intervals (ISI) remaining constant within one block. The first and third block always contained the stimulus pairs separated by a 25 ms ISI, the second and fourth block stimuli were separated by 400 ms ISI. So there was a short-long-short-long ISI-pattern for each continuum. Within blocks stimuli were randomized. The task was preceded by 24 practice stimuli. These were always items of the continuum that was presented first and their ISI was short, like in the first block. No direct feedback was given. Subjects were instructed to listen to the two words presented and to determine whether they sounded "the same" or "different". They were urged to react as adequately and as quickly as possible by pushing the corresponding button. In cases where a subject's reactions were systematically above 1000 ms, the experimenter intervened to make the subject aware of the need to react as quickly as possible. The discrimination task took 45 minutes.

Classification. Subjects were instructed to classify a stimulus as being "bak" or "dak", "bak" or "zak", "bak" or "wak", respectively. Again neither training nor feedback was given, 15 practice trials were presented to accustom the subjects to the task. Each stimulus was presented 12 times, the 120 stimuli of each continuum were randomized not allowing more than 3 identical stimuli in a row. Subjects were instructed to label the words as quickly and as adequately as possible. Again occasionally interventions were made when subjects responded too slowly (>800 ms).

3. RESULTS

3.1 Discrimination

Analysis. Since an ANOVA on all data revealed significant interactions between subject groups and continua, as could be expected because of the large differences in character of the three continua, we conducted an ANOVA on each of the three continua separately, across subject groups (control and dyslexic), ISI (25 and 400 ms), and stimulus pairs (seven 3-step pairs) with the dependent measure being the same-different discrimination scores. Significant main effects ($p<.001$) were found for subject groups (dyslexics performing worse) and for stimulus pairs (as could be expected), whereas the main effect of ISI was moderately significant ($p<.05$).

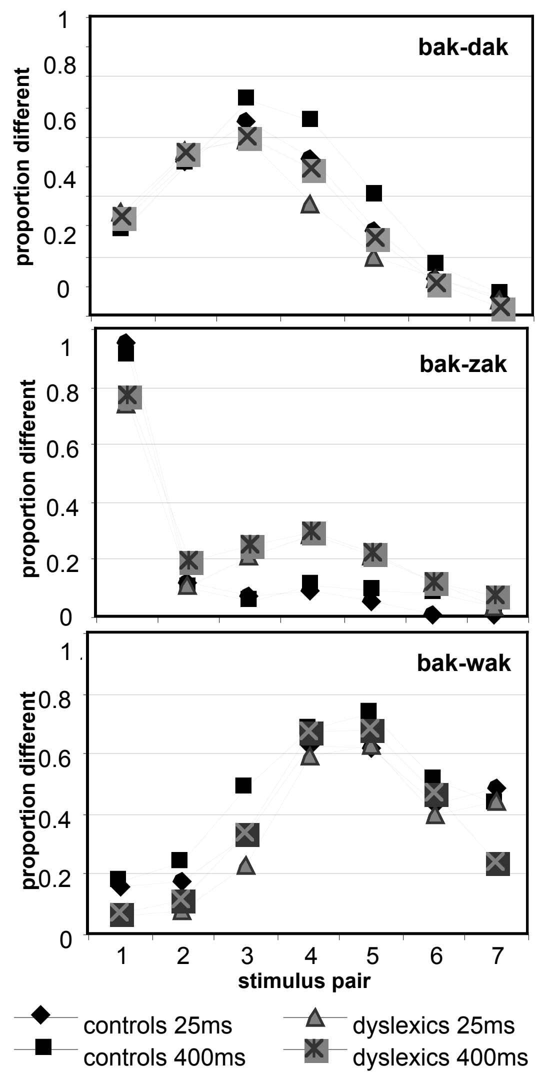

Figure 1: Mean discrimination functions for three continua with short and long ISI for dyslexic and control subjects.

The interaction between ISI and subject groups, however, was not significant, indicating that for both controls and dyslexics a longer ISI caused similarly better discrimination. A preliminary analysis of reaction times confirmed group differences, dyslexics reacting approximately 150 ms slower. In Fig. 1 mean discrimination functions for all continua with short and long ISI are presented for dyslexics and controls.

3.2 Classification

Analysis. ANOVA on each of the three continua separately across subject groups (control and dyslexic) and stimuli (10 in each continuum), with the dependent measure being the forced choice classification scores, revealed significant main effects (p<.001) for both subject groups (again dyslexics performing worse) and stimuli (as could be expected). In Fig. 2 the mean classification functions are presented for each of the three continua and both subject groups. If we compare the functions of the three continua, it is obvious that dyslexics and controls are more similar for the /bɑ k/-/dɑ k/ than for the /bɑ k/-/wɑ k/ continuum, which was to be expected if our assumption concerning degree of difficulty of the continua was correct. As for the /bɑ k/-/zɑ k/ continuum it turned out that the presence or absence of any noise part caused a phoneme boundary effect for the control group as well as for half of the dyslexic group. The other half situated the phoneme boundary in the middle of the continuum. This observation was confirmed by a posthoc analysis.

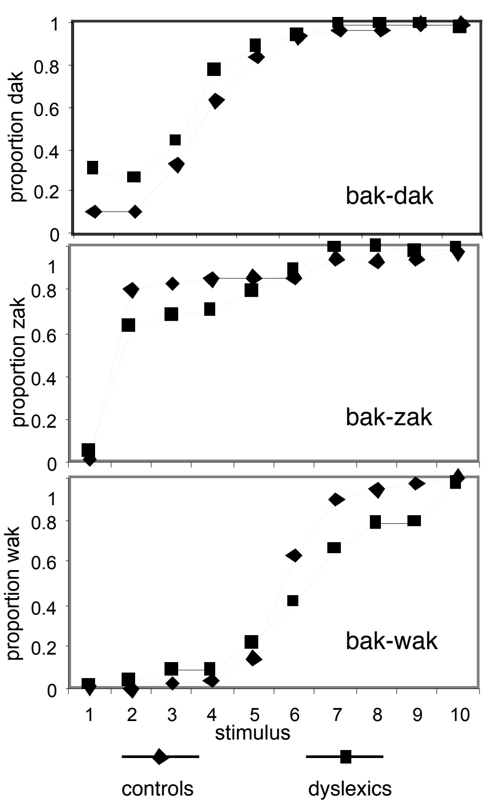

Figure 2: Mean classification functions for three continua for dyslexic and control subjects.

4. CONCLUSIONS

The main conclusions from the present experiment can be summarized as follows. Making various speech continua in Dutch based on natural speech, by interpolating one or two speech parameters turns out to provide quite satisfying stimuli. With some small adaptations they are suitable for further use in the dyslexia project. Although results on the /bɑ k/-/zɑ k/ continuum evoked questions, it could be demonstrated in our pilot experiments that dyslexic and control subjects behave differently on all three continua in discrimination as well as in classification tests. Longer ISIs turn out to be equally profitable for both subject groups. It still has to be studied whether other ISI values probably do differentiate the two subject groups. Since subjects in our pilot were all adults selected roughly from existing subject pools, and the import of the individual behaviour has to be studied in more detail, our results are promising in the scope of the Dutch dyslexia project.

5. REFERENCES

1. Boersma, P. *"Praat: A system for doing phonetics by computer"*, http://fonsg3.hum.uva.nl/praat/praat.html, 1998.

2. Mody, M., Studdert-Kennedy, M. & Brady, S. "Speech perception deficits in poor readers: Auditory processing or phonological coding?" *Journal of Experimental Child Psychology* 64: 199-231, 1997.

3. Reed, M.A. "Speech perception and the discrimination of brief auditory cues in reading disabled children". *Journal of Exp. Child Psychology* 48: 270-292, 1989.

4. Richardson, U. *"Familial dyslexia and sound duration in the quantity distinctions of Finnish infants and adults"*. Ph.D.-thesis, University of Jyväskylä, Finland, 1998.

5. Tallal, P. "Auditory temporal perception, phonics, and reading disabilities in children". *Brain and Language* 9: 182-198, 1980.

6. Werker, J. & Tees, R. "Speech perception in severely disabled and average reading children". *Canadian Journal of Psychology* 41: 48-61, 1987.

Acknowledgement. The authors wish to thank the Max Planck Institute for Psycholinguistics in Nijmegen for making available their experimental facilities.

Authors' correspondence address: Florien J. Koopmans-van Beinum, University of Amsterdam, Institute of Phonetic Sciences/IFOTT, Herengracht 338, 1016 CG Amsterdam, The Netherlands. E-mail: florienk@fon.hum.uva.nl.