Cognition

The impact of alphabetic literacy on the perception of speech sounds

--Manuscript Draft--

Manuscript Number:	
Article Type:	SI: Jacques Mehler SI
Keywords:	phoneme perception; categorical perception; categorical precision; literacy effects
Corresponding Author:	Régine Kolinsky, Ph.D. Université Libre de Bruxelles: Universite Libre de Bruxelles Bruxelles, BELGIUM
First Author:	Régine Kolinsky, Ph.D.
Order of Authors:	Régine Kolinsky, Ph.D.
	Ana Luiza Navas, Ph.D.
	Fraulein Vidigal de Paula, Ph.D.
	Nathalia Ribeiro de Brito
	Larissa de Medeiros Botecchia
	Sophie Bouton, Ph.D.
	Willy Serniclaes, Ph.D.
Abstract:	The aim of the present study was to reevaluate the impact of literacy on phoneme perception by using more controlled materials and by examining the impact of literacy more finely than in former studies. We therefore presented three groups of adults of varying literacy levels (Experiments 1 and 2) as well as Grade 2 children (Experiment 2) with identification and discrimination tasks, using a place-of-articulation continuum (Experiment 1) and a voicing continuum (Experiment 2). In addition to examining their categorical perception of speech sounds and the precision of phonemic categories, we also carefully evaluated the participants' literacy level. Literacy did not modulate categorical perception, neither with the place-of-articulation continuum used in Experiment 1, nor with the voicing continuum used in Experiment 2. Yet, with the latter material literacy strongly impacted the precision of phonemic categories, an effect that was found to be independent of age. The observation that literacy impacts categorical precision only with the voicing continuum, not with the place-of-articulation one, is discussed in relation to the difference in the phonemic status of these two features, with voicing perception being intrinsically dependent on the relationships with other features. The present data thus offer new perspectives to investigate the possible consequences of literacy on the binding between different features that are at the root of phoneme perception.
Suggested Reviewers:	Arthur Samuel, Ph.D. Professor, Stony Brook University arthur.samuel@stonybrook.edu specialized in speech perception
	Sven Mattys, Ph.D. Professor, University of York sven.mattys@york.ac.uk specialized in speech perception

Brussels, September 13, 2019

Dear Editors,

With my colleagues Ana Luiza Navas, Fraulein Vidigal de Paula, Nathalia Ribeiro de Brito, Larissa de Medeiros Botecchi, Sophie Bouton, and Willy Serniclaes, we would like to submit a manuscript titled "The impact of alphabetic literacy on the perception of speech sounds" for publication as an original empirical research article in the special issue of *Cognition* in honor of our dear friend Jacques Mehler.

The aim of the study was to reevaluate the impact of literacy on phoneme perception by using more controlled materials and by examining the impact of literacy more finely than in former studies. We therefore presented three groups of adults of varying literacy levels (Experiments 1 and 2) as well as Grade 2 children (Experiment 2) with identification and discrimination tasks, using a place-of-articulation continuum (Experiment 1) and a voicing continuum (Experiment 2). In addition to examining their categorical perception of speech sounds and the precision of phonemic categories, we also carefully evaluated the participants' literacy level. We observed that literacy did not modulate categorical perception, neither with the place-of-articulation continuum used in Experiment 1, nor with the voicing continuum used in Experiment 2. Yet, with the latter material literacy strongly impacted the precision of phonemic categories, an effect that was found to be independent of age. The observation that literacy impacts categorical precision only with the voicing continuum, not with the place-of-articulation one, is discussed in relation to the difference in the phonemic status of these two features, with voicing perception being intrinsically dependent on the relationships with other features. The data thus offer new perspectives to investigate the possible consequences of literacy on the binding between different features that are at the root of phoneme perception.

We hope you will find it suitable for publication in *Cognition* and look forward to hearing from you.

On behalf of all authors, we state that there is no conflict of interest.

Sincerely,

Régine Kolinsky

The impact of alphabetic literacy on the perception of speech sounds

Régine Kolinsky[1,2], Ana Luiza Navas[3], Fraulein Vidigal de Paula[4], Nathalia Ribeiro de Brito[3], Larissa de Medeiros Botecchia[3], Sophie Bouton[5], & Willy Serniclaes[2,6]

1. Fonds de la Recherche Scientifique-FNRS (FRS-FNRS), Belgium
2. Unité de Recherche en Neurosciences Cognitives (Unescog), Center for Research in Cognition & Neurosciences (CRCN), Université Libre de Bruxelles (ULB), Belgium.
3. Curso de Fonoaudiologia, Faculdade de Ciências Médicas da Santa Casa de São Paulo, School of Medical Sciences of Santa Casa de São Paulo, Brazil
4. Fraulein Vidigal de Paula, Instituto de Psicologia, USP, Brazil
5. Laboratoire Dynamique du Langage, CNRS, and Université Lyon 2 UMR 5596, France
6. CNRS and Université de Paris, France

Declarations of interest: none

Corresponding author:

Régine Kolinsky

Unescog, Université Libre de Bruxelles, CP 191

50 Ave. F. Roosevelt

B- 1050 Bruxelles, Belgium

Email: Kolinsky.Regine@ulb.be

Abstract

The aim of the present study was to reevaluate the impact of literacy on phoneme perception by using more controlled materials and by examining the impact of literacy more finely than in former studies. We therefore presented three groups of adults of varying literacy levels (Experiments 1 and 2) as well as Grade 2 children (Experiment 2) with identification and discrimination tasks, using a place-of-articulation continuum (Experiment 1) and a voicing continuum (Experiment 2). In addition to examining their categorical perception of speech sounds and the precision of phonemic categories, we also carefully evaluated the participants' literacy level. Literacy did not modulate categorical perception, neither with the place-of-articulation continuum used in Experiment 1, nor with the voicing continuum used in Experiment 2. Yet, with the latter material literacy strongly impacted the precision of phonemic categories, an effect that was found to be independent of age. The observation that literacy impacts categorical precision only with the voicing continuum, not with the place-of-articulation one, is discussed in relation to the difference in the phonemic status of these two features, with voicing perception being intrinsically dependent on the relationships with other features. The present data thus offer new perspectives to investigate the possible consequences of literacy on the binding between different features that are at the root of phoneme perception.

Keywords: phoneme perception; categorical perception; categorical precision; literacy effects.

1. Introduction

Several behavioral results have been interpreted as showing that reading instruction in an alphabetic script constitutes an encouragement to process speech in terms of phonemes. This holds certainly true as regards the conceptualization of speech. As a matter of fact, phoneme awareness develops in interaction with the acquisition of an alphabetic script (Morais, Alegria, et al., 1987). Consistently, neither preliterate children (e.g., Liberman et al., 1974) nor adults who have never learned an alphabet either because they remained illiterate for socioeconomic reasons as they never attended school (e.g., Morais et al., 1979; 1986) or because they only learned a nonalphabetic system (e.g., Read et al., 1986), are able to perform phonemic awareness tasks. For instance, they remain quite poor at deleting the first phoneme of an expression (e.g., /kat/→/at/; in all those studies, around 20% average correct responses in illiterates or nonalphabetic readers, vs. more than 70% in alphabetic literates).

Notably, the representations involved in such tasks that require attention to and/or explicit manipulation of phonemes differ from perceptual representations. The same illiterate people who perform poorly on those tasks discriminate almost perfectly between pairs like /ta–sa/ or /pa–ba/ (Adrián et al., 1995; Scliar-Cabral et al., 1997). Yet, whether speech perceptual representations might be more finely tuned by literacy[1] acquisition remains controversial. Although several observations seem to support this idea, most of them cannot be unambiguously interpreted. For instance, illiterate adults recognize spoken words less well than literates in difficult listening conditions, namely when the task is made difficult by adding noise (Ventura et al., 2007), or with two words presented dichotically (i.e., simultaneously, one to each ear, Morais, Castro et al., 1987). This literacy effect may not reflect differences in perceptual processes or representations, but rather an attentional strategy based on the explicit awareness of phonemes, which literate but not illiterate people can develop. This idea accounts

[1] Henceforth, *literacy* will refer only to alphabetic literacy.

for the fact that, in word dichotic listening, the ratio between segmental word identification errors (involving only one phoneme) and global word identification errors (involving all the segments of a syllable) was found to be lower in illiterate than in literate adults (Morais, Castro et al., 1987). In addition, this ratio is increased in undergraduate students by instructions to pay attention to the phonemic structure of the items (Castro, 1988, 1993). Although under strategical control, this effect illustrates the fact that phoneme awareness might improve performance even in word-level recognition tasks in which this ability is not compelling.

Another way to examine whether literacy finely tunes speech perceptual representations is to use simple identification (e.g., labeling) or discrimination (e.g., same-different judgment) tasks in which stimuli differ along physical continua such as formant transitions or voice onset time (*VOT*, responsible for voicing), and/or to examine the developmental course of attenuation for non-native speech contrasts (cf. e.g., Werker & Tees, 1983) compared to native ones. A number of studies using these approaches have reported that children present an improvement of speech identification and discrimination performance that occurs maximally at the beginning of reading acquisition. As a matter of fact, comparing native and non-native speech contrasts, three separate experiments on children (Burnham, 2003; Burnham et al., 1991) reported that for native speech contrasts, identification function is steeper and discrimination enhanced at the beginning of reading acquisition, these changes being significantly associated with literacy-related abilities (reading and phoneme awareness).

However, none of these studies independently manipulated age and instruction related to school experience, including reading instruction. Thus, definitive conclusions as regards the role of literacy per se cannot be drawn from those data. To overcome this limitation, Horlyck et al. (2012) adopted a *cut-off design* (e.g., Cahan & Davis, 1987; Morrison et al., 1995), using the relatively broad temporal window over which children start school as a means of independently manipulating age and school experience. Indeed, in some countries, children of the same age may start school up to one year apart. Horlyck et al. observed that the discrimination of a native speech contrast was not related to age, but rather to school experience and phonological awareness, as assessed by tests of syllable counting, rhyme, phoneme identification, segmentation, blending, phoneme deletion, and grapheme-to-phoneme correspondence. Although no reading test was included, these researchers speculated that the crucial element of school experience is reading acquisition. Yet, as acknowledged by Horlyck et al., a limitation in the implications of their results resides in a selection bias associated with the parents' choice of school entry age.

Crucially, all those studies examined identification and discrimination performance on independent groups of participants. When considered separately, these scores reflect *categorical precision*, which indeed can be evaluated either by measuring the steepness of the identification function around the perceptual boundary (the steeper the slope of the identification function, the greater the precision, Simon & Fourcin, 1978) or by measuring the size of the discrimination peak around the phoneme boundary (the huger the peak, the greater the precision, Wood, 1976). However, categorical precision is relatively independent from *categorical perception* (for a review, see Damper & Harnad, 2000; see also discussion in Schouten et al., 2003), a phenomenon in which stimuli differing along physical continua such as formant transitions or VOT are perceived as belonging to distinct phonological categories. More precisely, according to Liberman et al. (1957), categorical perception means that only differences between identified phonemic categories (e.g., between phonemes identified as /b/ vs. as /d/) can be distinguished, not the within-category variants (e.g., between two physically different sounds, both identified as /b/). Categorical perception is thus estimated through the relation between performance in identification and discrimination.

To our knowledge, only two studies so far aimed at examining the impact of literacy acquisition on both categorical precision and categorical perception. One was run on adults and on children from the last year of kindergarten to the second grade of primary school (Hoonhorst et al., 2011), the other compared illiterate adults to schooled literates (Serniclaes et al., 2005). In both studies, there was a significant relation between literacy (and correlated schooling level) and categorical precision, whereas categorical perception was found to be the same whatever the level of literacy/schooling. These two sets of data thus suggest that instruction (either specifically literacy, and/or schooling) helps in finely tuning phonemic boundaries and hence in increasing the precision of phoneme identification, but does not change categorical perception per se.

Nonetheless, Serniclaes et al. (2005) acknowledged that their adult data could be attributed to a lexical bias, as one of the end points of the /ba-da/ speech continuum that they used was a word. Indeed, in addition to display a shallower slope of the labeling function than literates, illiterate adults differed from literates in the intercept of the labeling function. This intercept difference reflected a bias towards /da/ responding in the illiterate group, which may be due to a lexical bias: whereas in Portuguese (the language of the study) /ba/ is nonsense, /da/ is an extremely frequent form of the verb “dar”, which means “give”. A lexical bias has repeatedly been documented in the illiterate population. For instance, in pseudoword immediate repetition, lexicalization errors (i.e., word responses to phonologically related pseudowords) were fewer than 2% of the errors in the literate adults examined by Castro-Caldas et al. (1998) but reached 11% in the illiterate group. The strong illiterate adults’ bias for identifying the first stimuli of the continuum as /da/ instead of /ba/ is likely to be a further instance of this lexical bias and may have affected categorical precision. In agreement with this idea, Serniclaes et al.

found the intercept of the labeling function to be highly correlated to the slope of this function.[2] It is thus still unknown whether illiterate adults actually differ from literates in categorical precision, or whether this effect was only related to the lexical bias induced by the material used by Serniclaes et al.

A further question raised by previous evidence on the effect of literacy on speech perception concerns the nature of the stimuli in both the adult study ran by Serniclaes et al. (2005) and the developmental study ran by Hoonhorst et al. (2011). In the latter study, responses were collected with a VOT continuum with two lexical endpoints (the French words "de" and "te"), excluding an explanation in terms of lexical bias. However, the effect of reading experience on categorical precision was probably due to another stimulus factor. Although VOT is the primary acoustic cue to voicing, the perception of voicing (and other phonological features) depends on the integration of multiple acoustic cues (burst loudness: Repp, 1979; first formant characteristics: Summerfield, 1982; formant transitions: Stevens & Klatt, 1974; fundamental frequency – F0 – characteristics: Haggard, Summerfield, & Roberts, 1981). In many studies using a synthesized voicing continuum, including the Hoonhorst et al. one, only the primary acoustic cue to voicing (VOT) was modified. Thus, some of the secondary cues to voicing did not covary in the stimuli, giving rise to conflicting phonological information at the continua endpoints. With conflicting-cue stimuli, kindergarteners do not fully recognize the phonological categories at the continua endpoints, a deficit that Hoonhorst et al. reported as fading with age, in relation with reading acquisition. Note that such deficit does not really impact the precision of the categorical boundary, but rather the precision of the categorical

[2] However, and in agreement with the assumption of independence between categorical perception and precision, the intercept was uncorrelated to the degree of categorical perception.

prototypes inside each category. The same reasoning holds true for the continuum used by Serniclaes et al., in which only the primary acoustic cues to place of articulation (the second and third formant transitions, F2 and F3, respectively) were modified. The use of stimuli with conflicting acoustic cues was thus probably at the origin of an effect of reading experience on the precision of the categorical prototypes in both studies, in addition to the lexical bias that Serniclaes et al. reported. In fact, both types of effects concern the identification responses at the continua endpoints (the asymptotes of the response function), not those around the boundary. These effects correspond to a generalization of the categorical representations, for either non-words (Serniclaes et al.) or extra-prototypical stimuli (both Hoonhorst et al. and Serniclaes et al.), rather than to actual improvements in the precision of the boundary between categories. The question that remains is whether literacy really affects boundary precision, irrespective of the generalization effects that were evidenced in previous studies.

The aim of the present study was to reevaluate the impact of literacy on categorical precision and categorical perception by using more controlled materials, without lexical bias and without conflicts between acoustic cues, using both a place of articulation continuum (as did Serniclaes et al., 2005) and a voicing continuum (as did Hoonhorst et al., 2011). As in the two former studies, for each continuum we presented the participants with both an identification (labelling) and a discrimination (same/different) task. In order to avoid lexical biases, we selected words of similar frequency, actually letter names, as continua endpoints. We used a /be/–/de/ place-of-articulation continuum with the <be> and <de> letter names as endpoints (Experiment 1), and a /de/–/te/ voicing continuum with the <de> and <te> letter names as endpoints (Experiment 2). Both continua were generated by morphing (interpolation) between natural speech exemplars pronounced by a native (Brazilian) male speaker. We chose this procedure to minimize possible differences in categorical precision arising from a lack of perceptual integration between the different acoustic cues that contribute to the perception of the same feature. With morphing, possible conflicts between these different acoustic cues are minimized because all the relevant acoustic cues for a given feature are modified simultaneously. Thus, in the present study, the use of stimuli generated by morphing allows to focus on the possible effects of reading on boundary precision.

We also aimed at examining more thoroughly the impact of literacy than the former studies. As a matter of fact, as Serniclaes et al. (2005) contrasted two extreme groups, namely fully illiterate adults and schooled literates, they could not examine the correlations between literacy level and categorical precision or perception. In the present study, we examined the categorical perception of speech sounds and the precision of phonemic categories in groups of adults (Experiment 1) and of adults and Grade 2 children (Experiment 2) varying by their literacy level, which was carefully evaluated. The adult participants either had attended school and hence had learned there to read and write in childhood or had not attended school in childhood for socio-economic reasons (or only for a short time), but were attending adult literacy classes at the time of testing and thus were beginning readers. All adult participants being of modest socioeconomic status (SES), we expected even the schooled literates to present some variability in their literacy level, as they had attended (usually low-quality) school for up to five years.

2. Experiment 1

2.1 Method

2.1.1 Participants

The initial sample included 55 adult participants, all fully functional in their daily lives and socially integrated. They were all volunteers and gave their informed consent. All were asked about their native language and screened for auditory difficulties. The auditory screening consisted of acoustic immittance tests with broadband stimulus, search of acoustic reflex thresholds and capture of distortion product otoacoustic emissions, using the TITAN -

Interacoustic Equipment (IMP440 Module - Broadband Immittance / Tympanometry Module - and DPOAE440 Module - Distortion Product Otoacoustic Emissions Module).

Five participants presenting hearing impairments were excluded, as well as two participants who were not Brazilian Portuguese natives, two who were quite different in age from the overall age range (one was 16 years old, the other 67 years old; see Table 1 for the age range of the final sample) and five who correctly identified only one of the endpoints of the speech continuum (2 very poorly literates and 2 schooled literates, see below). On the remaining 42 adult participants (11 very poorly literates, 17 poorly literates, 14 schooled literates, see below), we checked for cognitive impairment by means of the Mini-Mental State Examination (MMSE; Folstein et al., 1975; for the Portuguese version: Guerreiro et al., 1994). Although the MMSE scores were quite low (see Table 1 for the characteristics of the final samples), in particular in the very poorly literates, they were in the range of those usually observed on unschooled or poorly schooled adults (Caramelli et al., 1999).[3]

The final sample of adults was subdivided into three groups according to two criteria. We first considered whether they had learned to read and write at school in childhood or not. Those who did were included in the *schooled literate* group. These participants had been recruited at the School of Medical Sciences Santa Casa de São Paulo, where they were working in the cleaning, security or maintenance staff. They had attended school for four years on the average (maximum: 5 years). The other participants had not attended school in childhood for socio-economic reasons, or only for a very short time (maximum 2 years) and had been recruited in the first and second terms of the adult literacy classes, which are

[3] Consistently, a majority of the very poorly literates (about 57%) and a minority of the poorly literates (about 35%) were unable to answer to the two items that required to read or write.

organized by reading level. These unschooled adults were assigned either to the group of *poorly literates* or to the group of *very poorly literates* as a function of their reading proficiency: the former had obtained at least 70% correct on average in two reading tests[4] (see Table 1), one including 36 words (12 simple, 12 complex and 12 irregular ones), the other 16 pseudowords (8 simple and 8 complex ones); the others either had obtained a score lower than 70% or had refused to participate to those tests, feeling unable to succeed them. Word reading fluency was also assessed in all the participants: presented with a list of 108 upper-case words ordered by increasing level of difficulty (depending on length, complexity, and regularity), they were required to read as many words as they could in one minute (Paula, 2007).

In addition, all participants were presented with a test of letter knowledge (oral identification of 23 upper-case letters) as well as with three metaphonological tests of varying difficulty: phonological sensitivity, initial syllable deletion, and the more difficult test of initial phoneme deletion (e.g., Liberman et al., 1974). In the phonological sensitivity test, participants were presented with six panels, each with six drawings of common objects. On each panel, they were asked to point to the drawings corresponding to names of objects that started with a target phoneme. In the demo trial, for the phoneme /f/, they heard six words uttered by the experimenter: “fita”, “fada”, “ferro”, “fumo”, “fato”. Their attention was directed to the fact that all words started with the /f/ “sound”. They were next presented with six drawings and asked to point to those corresponding to a name starting with /f/. On each panel there were from two to four images to be pointed to (total: 18), half with a name starting with a simple (C) onset (e.g., “faca”), the others with a name starting with a complex (CC)

[4] Except when specified otherwise, the literacy-related screening tests were adapted to Brazilian Portuguese from those developed for a study commissioned by the National Reading Plan of the Portuguese Ministry of Education (Morais et al., 2010).

onset (e.g., “flor”, “frasco”). The foils names had quite different onsets compared to those of the target names (e.g., “escada”, “osso”, “índio”). In both deletion tests, participants had to repeat part of a spoken pseudoword uttered by the experimenter after having deleted its initial part, either the first syllable (e.g., /kɔbu/ → /bu/) or the first phoneme (e.g., /tɔbu/ → /ɔbu/). All expected responses were also pseudowords. Each deletion test included 10 disyllabic items (all CVCV in the phoneme deletion test), plus two training trials with corrective feedback. Two tests of rapid automatized naming (RAN) were also presented, one on pictures (of a boat, key, chair, pencil, fish and star) and one on digits (from 2 to 8), as this measure of phonological processing seems to be related to reading abilities (e.g., Araújo et al., 2019) in a way that is independent from other abilities such as phonological awareness (e.g., Powell et al, 2007; for a metanalysis, see Araújo et al, 2015). For the RAN tests, the participants were instructed to name the items, repeated in random order for a total of 36 stimuli. The time in seconds that participants took to name the 36 stimuli was recorded. The average performance of the final samples on all the ancillary tests are presented in Table 1, as well as the sociodemographic characteristics of the samples.

All participants were from relatively modest SES, estimated by applying the Brazilian Criteria of Economic Classification (BCEC, Critério de Classificação Econômica Brasil – CCEB, Associação Brasileira de Empresas de Pesquisa, 2015; http://www.abep.org/criterio-brasil). These classify the population based on the ownership of assets with a score for all possessions, as estimated through a questionnaire about home facilities and transportation (e.g., number of bathrooms, of domestic servants, of refrigerators, of cars, etc.), formal education, and access to public utility services (e.g., paved street or not). Each SES class (from the highest to the lowest, A1 to D/E) is defined by the sum of those scores; according to the BCEC, these classes correspond to a monthly family income of R$ 12926 for class A1, R$ 8418 for A2, R$ 4418 for B1, R$ 2565 for B2, R$ 1541 for C1, R$ 1024 for C2 and from R$

714 to R$ 477 for classes D and E (R$ 477). Using the 2015 revised Brazilian criteria, no difference in SES was observed between the three groups, $\chi^2(8) = 10.75$, $p = .22$. Participants varied between the lowest (D/E) SES category and the second highest one (B1), with a majority of participants (28) presenting a C1 or C2 level. As can be seen in Table 1, a similar conclusion comes from the analysis ran on the sum of the scores. Yet, as can also be seen in Table 1, participants differed on age, with both very poor literates and poor literates being older than schooled literates, but not differing from each other.

2.1.2 Material

A six-steps place-of-articulation continuum, /be/–/de/ (S1–S6) was presented. High-quality speech stimuli were synthesized using TANDEM-STRAIGHT (Kawahara & Morise, 2011; Kawahara et al., 2009), a MATLAB tool that provides a temporally stable power spectral representation of the initial (prototypical) two syllables (i.e., /be/ and /de/). On the basis of the spectrographic representations of these two periodic signals, we implemented a morphing algorithm (Kawahara & Matsui, 2003) that first consisted in manually assigning anchor points. We located 20 anchors that were regularly spaced and similarly placed on each prototypical syllable. Then, the morphing algorithm followed a four-stage procedure, which consisted in: (i) aligning the time frequency coordinates of the two prototypical syllables, (ii) interpolating parameters represented on the aligned time-frequency coordinates according to the given morphing rate (i.e., a six step-rate), (iii) deforming the time-frequency coordinates according to the given morphing rate, and (iv) resynthesizing six sounds using the morphed parameters in the morphed time-frequency coordinates. The morphing procedure was applied to all parameters (F0, periodicity and spectrogram).

Each stimulus was presented 10 times in the identification test, in pseudo-random order, leading to a total of 60 identification trials. For the discrimination test, 10 “different” pairs of stimuli were created, each stimulus being paired with the adjacent one on the continuum (e.g., S1–S2; S2–S1; S2–S3, etc.). There were also six “same” pairs in which each stimulus was paired with itself. Each of these 16 pairs was presented five times, in random order, leading to a total of 80 discrimination trials.

2.1.3 Procedure

The study was approved by Santa Casa Medical Sciences ethics committee (protocol no. 64381417.5.0000.5479).

Participants were first presented with the questionnaires on schooling history, SES level, MMSE, and the auditory screening. Then they were presented with the phonological processing tasks, namely the metaphonological (phonological sensitivity, phoneme and syllable deletion) and RAN (pictures and digits) tests. On a second session participants were presented with the speech identification and discrimination tests and performed the reading tasks (letter, word and pseudoword recognition, and reading fluency). As the same participants were presented with another (voicing) continuum as well (see Experiment 2), order of the materials was counterbalanced across participants.

For the speech identification and discrimination tests, stimuli were presented through headphones at about 40 dB, using the PRAAT software (Boersma, 2001; Boersma & Weenink, 2019) running on a Macbook Air 13" computer. Participants first performed the identification test, then the discrimination test. As some participants were illiterates or (very) poorly literates, the experimenter asked them to respond orally in both tests and entered the response herself into the computer by pressing the corresponding response (keyboard) keys.

In the identification test, one syllable was presented on each trial. Participants were told that they would hear either /be/ or /de/ and were instructed to report which of the two sounds they heard. In the discrimination test, syllables were presented in pairs with a 300 ms between-syllables interval. Participants had to say whether the stimuli within each pair were the same (either two /be/, or two /de/ syllables) or different (/be/–/de/ or /de/–/be/), according to their impression to hear the same thing or not.

2.1.4 Data analysis

All the analyses were performed using *JASP* 0.13.1 (JASP Team, 2020). For each participant, identification scores were calculated as the mean of /de/ responses collected for each stimulus. Correct discrimination scores were calculated for each participant and for each pair (e.g., S1S2). In order to assess categorical perception following the classical procedure (Liberman et al., 1957), such "observed" discrimination scores were compared to "predicted" discrimination scores that were derived from the identification scores using elementary probability formulas (see Gottfried & Strange, 1980; adapted from Pollack & Pisoni, 1971).

These discrimination responses were then converted into *d'* values by taking the difference between the normal deviate (z-score) corresponding to the proportion of hits (correct difference detections i.e., proportion of "different" responses to different pairs) and the proportion of false alarms (i.e., proportion of "different" responses to same pairs). Since 0% and 100% scores correspond to infinite z scores, response scores were adjusted following the classical procedure described by Macmillan and Creelman (2005).

These *d'* values were then entered into an omnibus ANOVA with Pair (5 values: S1S2 to S5S6), Task (identification, discrimination) and Group as factors.[5] A group difference in categorical precision would lead to a significant Pair X Group interaction and a group difference in categorical perception to a significant Pair X Task X Group interaction.

For the computation of correlations, several indexes were calculated for each participant. Categorical precision was estimated by means of the *discrimination peak* and the

[5] We first had checked that order of presentation of the materials did not interact with these variables of interest.

slope of the identification function, calculated separately for each participant. The discrimination peak was taken as the difference between the largest *d'* score (averaged over observed and expected data), irrespective of the location of the phoneme boundary, and the mean of the other *d'* scores (also averaged over observed and expected data). Categorical perception was estimated by means of a *categorical perception index*. The latter was obtained by taking the absolute difference between the expected and observed discrimination peaks, after having individually adjusted those scores for the location of the boundary.

As regards literacy, we considered five scores in the computation of correlations: correct performance on letter knowledge, average word and pseudoword reading, average phonological awareness (three tests), word reading fluency, and average RAN performance. All the correlations with literacy scores were Bonferroni-corrected one-sided tests, as the alternative hypothesis was that literacy would improve categorical perception and precision.

2.2 Results

2.2.1 Ancillary tests

Performance on ancillary tests were analyzed in separate ANOVAs, each with group (very poorly literates vs. poorly literates vs. schooled literates) as a between-subjects factor. As shown in Table 1, the group effect was highly significant in all these analyses, except for letter knowledge. In the other tests, except for phonological sensitivity, both very poorly literates and poorly literates presented lower scores than schooled literates, and, as expected given the group inclusion criterium, very poorly literates presented lower scores than poorly literates in both pseudoword reading and word reading fluency.[6]

[6] The very poorly literates' performance on word and pseudoword reading presented in Table 1 is inflated by the fact that most did not accept to take part to these tests.

Table 1

Experiment 1: characteristics of the samples and average values on the ancillary tests. Standard deviation in brackets. The shaded cells signal when only a minority of participants of one or two groups have been tested (see valid values). Scores in bold were those used in the correlation analyses.

Group	TEST	AGE (yrs,months)	SES	MMSE	LETTER KNOWLEDGE % correct	READING Words % correct	READING Pseudowords % correct	READING Mean	READING FLUENCY correct/min	PHONOLOGICAL AWARENESS Syllable deletion % correct	PHONOLOGICAL AWARENESS Phoneme deletion % correct	PHONOLOGICAL AWARENESS Phonological sensitivity % correct	PHONOLOGICAL AWARENESS Mean	RAN Pictures sec	RAN Digits sec	RAN Mean
Very poorly literates	**Valid**	**11**	**11**	**11**	**11**	7	4	**4**	**11**	11	11	11	**11**	11	11	**11**
	Mean	**51.55**	**21.82**	**21.27**	**96.05**	78.57	42.19	**54.07**	**12.55**	42.72	6.36	43.94	**31.01**	37.43	27.65	**32.55**
	SD	**[10.84]**	**[6.21]**	**[3.52]**	**[4.94]**	[24.57]	[20.65]	**[21.15]**	**[9.53]**	[33.19]	[15.67]	[25.03]	**[18.77]**	[4.92]	[4.76]	**[3.32]**
	Range	**28 – 63**	**16– 36**	**17 – 27**	**86.96 – 100**	27.78 – 100	18.75 – 62.5	**23.26 – 68.4**	**0 – 36**	0 – 100	0 – 50	0 – 83.33	**5.56 – 66.11**	31.75 – 47.34	22.27 – 38.43	**27.41 – 37.16**
Poorly literates	**Valid**	**17**	**17**	**17**	**17**	17	17	**17**	**17**	17	17	17	**17**	17	17	**17**
	Mean	**46.53**	**23.41**	**24.82**	**97.7**	85.78	76.10	**80.94**	**27.71**	64.12	15.88	65.69	**48.56**	35.02	26.31	**30.66**
	SD	**[4.69]**	**[5.61]**	**[2.72]**	**[4.64]**	[10.57]	[11.32]	**[9.08]**	**[14.72]**	[40.94]	[30.01]	[37.49]	**[29.09]**	[5.63]	[4.39]	**[4.68]**
	Range	**36 – 55**	**17 – 39**	**19–30**	**82.61 – 100**	66.67 – 100	56.25 – 100	**70.14– 94.44**	**6 – 54**	0 – 100	0 – 100	0 – 100	**0 – 100**	23.73 – 46	18.26 – 32.88	**21 – 39**
Schooled literates	**Valid**	**14**	**14**	**14**	**14**	14	14	**14**	**14**	14	14	14	**14**	14	14	**14**
	Mean	**39.93**	**26.21**	**26.93**	**99.07**	98.41	92.86	**95.64**	**64.64**	94.29	83.57	69.05	**82.302**	30.04	18.51	**24.28**
	SD	**[7.58]**	**[6.36]**	**[1.77]**	**[2.52]**	[2.61]	[7.3]	**[4.27]**	**[17.72]**	[16.04]	[21.34]	[18.32]	**[14.87]**	[8.09]	[2.61]	**[4.67]**
	Range	**25 – 52**	**16 – 38**	**24 – 29**	**91.3 – 100**	91.67 – 100	75 – 100	**86.11– 100**	**28 – 107**	40 – 100	20 – 100	33.33 – 100	**53.33 – 100**	18 – 48.7	13.5 – 23.28	**15.75 – 33.8**
ANOVA[a]		**7.321*******	**3.927**	**12.522********	**1.634**	12.678****	19.907***		**43.818* *****	12.957****	41.267****	3.87*		4.48*	20.632****	
	η^2_p	**.273**		**.412**	**.077**	.284	.672		**.692**	.292	.679	.119		.187	.514	
	$t(29)$[b] Poorly literates vs. schooled literates	**2.395***		**-2.163***		-2.792*	-4.141****		**-7.139********	-2.573*	-7.759****	-0.32		2.156(*)	5.407****	
	$t(26)$ Very poorly literates vs. poorly literates	**-1.698**		**3.403******		1.281	5.444****		**2.485***	1.701	1.018	1.929		-0.975	-0.867	
	$t(23)$ Very poorly literates vs. schooled literates	**-3.776******		**5.206********		3.42***	8.973****		**8.782********	3.938****	7.927****	2.139		-2.868*	-5.676****	

* p < .05,** p < .01 ,*** < .005, **** < .001

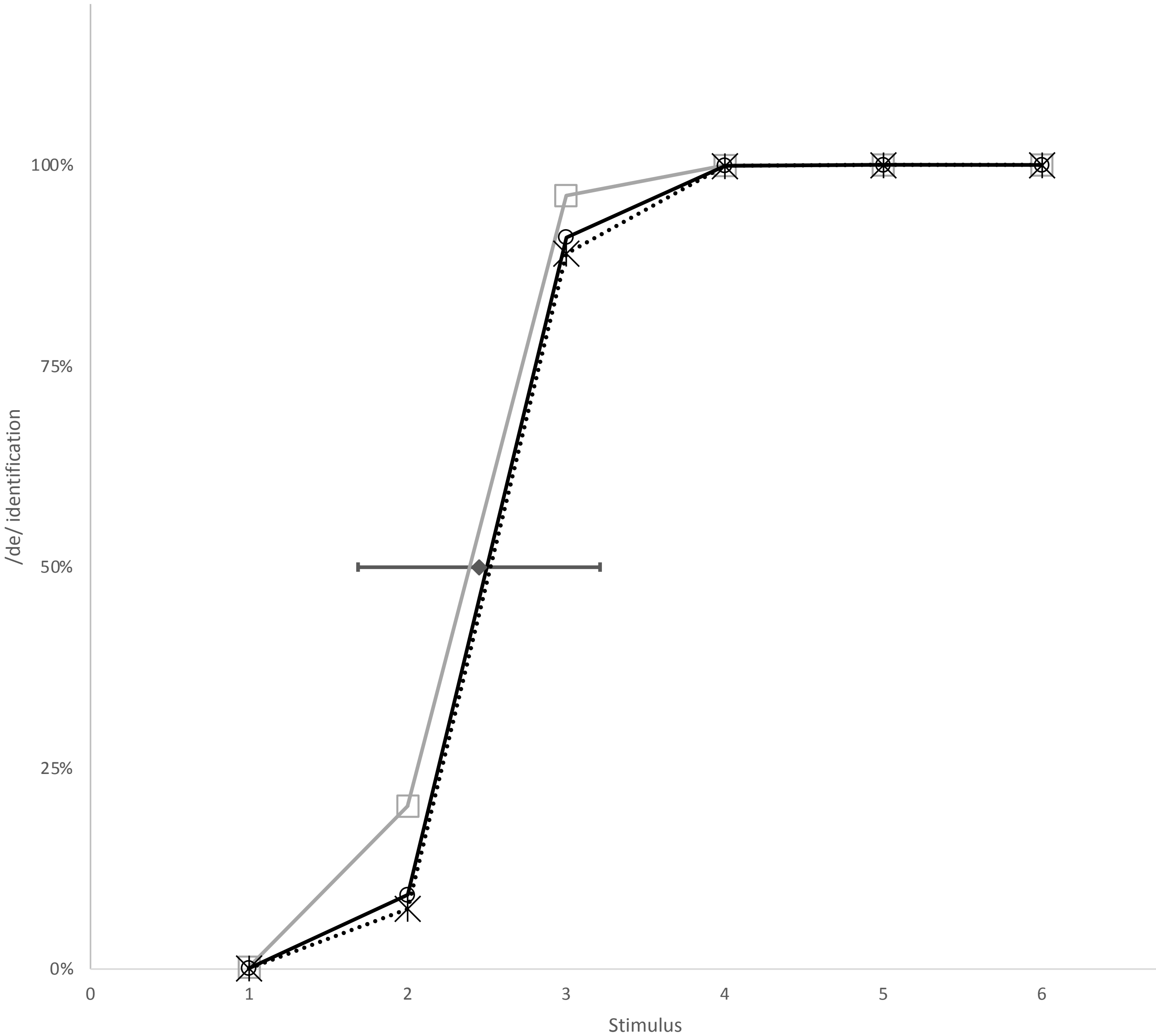
a: all values are $F(2, 39)$, or $F(2, 35)$ and $F(2, 32)$ for word and pseudoword reading, respectively, except for SES, for which we ran a one-way analysis of variance on ranks (Kruskall-Wallis test). Welch correction was applied when necessary in the parametric ANOVAs.

b: All post-hoc t tests were Holm-corrected

2.2.2 Categorical perception and precision

The labeling functions, as a function of stimulus, are presented in Figure 1, separately for each group; discrimination performance, as a function of pair, is displayed in Figure 2, separately for each group.

Figure 1



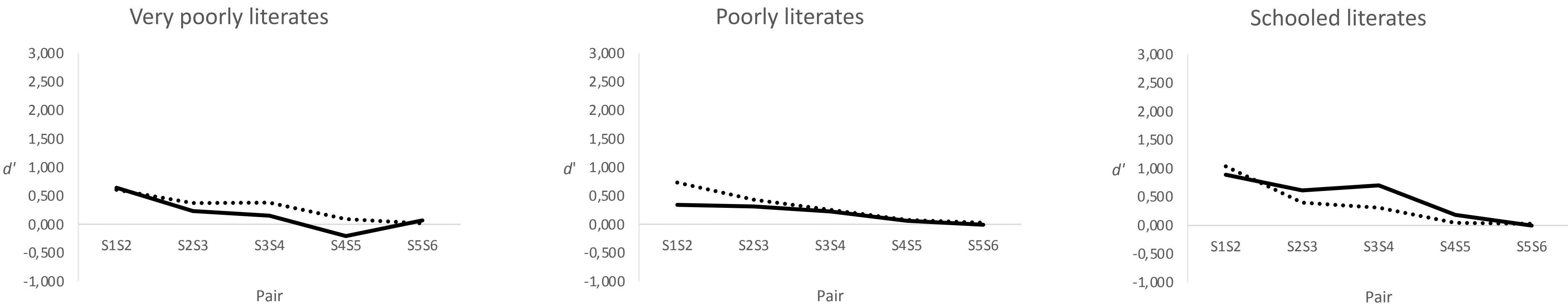

Very poorly literates
Poorly literates
Schooled literates
3,000
2,500
2,000
1,500
1,000
0,500
0,000
-0,500
-1,000
d'
S1S2
S2S3
S3S4
S4S5
S5S6
Pair

Figure 2

The Pair X Task X Group (very poorly literates, poorly literates, schooled literates) repeated measure ANOVA conducted on *d'* showed a significant main effect of Pair, $F(2.14, 83.55) = 10.457$, $p < .001$, $\eta^2_p = 0.21$, with S2S3 leading to better performance than S3S4, S4S5 and S5S6, $t(41) = 3.006$, $p < .025$, $= 5.383$, $p < .001$, and $= 5.542$, $p < .001$, respectively, and with S2S3 leading to the better performance compared to S4S5 and S5S6, $t(41) = 2.839$, $p < .05$ and $= 2.998$, $p = .025$. The main effect of Group was also significant, $F(2, 39) = 4.282$, $p < .025$, $\eta^2_p = .18$, with both poorly literates and very poorly literates displaying overall lower performance than schooled literates, $t(29) = -2.589$, and $t(23) = 2.461$, both *p*s < .05, without differing from each other, $t < 1$. Neither the Pair X Group nor the Pair X Task X Group interactions were significant, both $Fs < 1$.

A similar Bayesian repeated measure ANOVA supported the idea that groups differ neither in categorical precision, Bayesian factor $BF_{excl} = 59.669$ for the Pair X Group interaction[7], nor in categorical perception, $BF_{excl} = 24.44$ for the Pair X Task X Group interaction. In those analyses, the Bayes factors indicate how likely the data are under the null hypothesis (H_0: no group difference) compared with the alternative hypothesis (H_1: there is a group difference) and are directly interpretable as odds ratios. Thus, the Bayes factors reported here indicate that the data are about 60 times more likely under the hypothesis that there is no group difference in categorical precision than under H_1, which is considered as

[7] We estimated the contribution of effects and interactions using the JASP matched-model comparison procedure suggested by Sebastiaan Mathôt. This procedure compares models that contain the effect to equivalent models stripped of the effect. Concerning interactions, this allows evidence for the interaction to be evaluated on its own by comparing the *BF* of a model with the interaction against the *BF* of a model with only the main effects (i.e., without the interaction).

strong (Raftery, 1995) or very strong (Jeffreys, 1961) evidence for H_0 over H_1. They also indicate that the data are about 24 times more likely under the hypothesis that groups do not differ on categorical perception than under H_1, which is considered as strong evidence for H_0 over H_1 (Jeffreys, 1961; Raftery, 1995).

Consistently, no significant correlation was observed between the literacy scores and either the slope of the identification function or the discrimination peak (all *p*s > .35; for average values of those indexes, see Table 2). Similarly, no significant correlation between the literacy scores and the categorical perception index reached significance, all *p*s > .62 (see also Table 2 for average values of this index). Neither did age correlate significantly with any of the speech perception indexes, all *p*s > .48 on bilateral tests. Yet the evidence in favor of H_0 (no correlation) was anecdotal according to the rather conservative Bayesian correlation tests, all $BFs_{01} \leq 4.52$, except that there was strong evidence in favor of absence of correlation between the categorical perception index and RAN, $\tau(40) = -.13$, $BF_{0+} = 10.31$, as well as, to a lesser extent, between the categorical perception index and reading fluency, $\tau(40) = .08$ $BF_{0+} = 8.31$.

Table 2
Experiment 1: average values of the speech perception indexes used in the computation of correlations. Standard deviation in brackets.

		Very poorly literates	Poorly literates	Schooled literates
Slope	**Mean**	**4.6**	**4.6**	**4.59**
	SD	[1.68]	[1.72]	[2.02]
	Range	2.33 – 6	1.88 – 6	1.18 – 6
Categorical precision index	**Mean**	**1.13**	**1.05**	**1.46**
	SD	[0.75]	[0.63]	[0.85]
	Range	0.44 – 2.46	0.31 – 2.22	0.29 – 3
Categorical perception index	**Mean**	**0.8**	**1.1**	**1.09**
	SD	[0.7]	[0.86]	[0.74]
	Range	0.06 – 2.23	0.14 – 3.86	0.21 – 2.75

2.3 Discussion

In the present experiment, as in Serniclaes et al. (2005) we did not observe a significant effect of literacy on categorical perception of place of articulation. Yet, contrary to Serniclaes et al. we did not observe a significant effect of literacy on categorical precision, either. Hence, the results of that former study were probably due to the lexical bias the authors referred to, and/or to the use of synthetic stimuli giving rise to a conflict between various acoustic cues and hence leading to an effect of literacy on the precision of the categorical prototypes. In the next experiment, we examined whether the same result pattern would hold true for a voicing contrast.

3. Experiment 2

3.1 Method

3.1.1 Participants

The initial sample included the same adult participants as those selected in Experiment 1 on the basis of the auditory screening, language proficiency, and age. Among these, two (one poorly literate and one very poorly literate) were excluded because they correctly identified only one of the endpoints of the speech continuum. As these were not the same individuals as those who had been excluded in Experiment 1, and hence as only 40 among the 45 participants of the final sample were included in both experiments, Table 3 present the average performance of Experiment 2 final samples on the ancillary tests, as well as their sociodemographic characteristics.

The initial sample of 57 Grade 2 children were all volunteers and gave their informed consent, as did their parents. The same screening and ancillary tests were applied as for adults, except that there were presented neither with the MMSE nor with the SES questionnaire. None of the children presented a hearing impairment, but five were discarded because they presented serious language troubles.

3.1.2 Material, procedure and data analysis.

Material, procedure and data analysis were the same as in Experiment 1, except that a six-steps /de/–/te/ voicing continuum was presented.

3.2 Results

In order to get a close comparison with the data of former studies as well as with those of Experiment 1, we first compared the three groups of adults with each other, and then compared the unschooled adults to the Grade 2 children.

3.2.1 Comparisons between the three adult groups

3.2.1.1 Ancillary tests

Performance on ancillary tests were analyzed in separate ANOVAs, each with group (very poorly literates vs. poorly literates vs. schooled literates) as a between-subjects factor. As presented in Table 3, the group effect was highly significant in all these analyses. Both very poorly literates and poorly literates presented lower scores than schooled literates in all tests, and, as expected given the group inclusion criterium, very poorly literates presented lower scores than poorly literates in both pseudoword reading and word reading fluency.[8]

[8] As in Experiment 1, the very poorly literates' performance on word and pseudoword reading (here presented in Table 3) is inflated by the fact that most did not accept to take part to these tests.

Table 3

Experiment 2: characteristics of the samples and average values on the ancillary tests. Standard deviation in brackets. The shaded cells signal when only a minority of participants of one or two groups have been tested (see valid values). Scores in bold wer

			AGE	SES	MMSE	LETTER KNOWLEDGE	READING			READING FLUENCY	PHONOLOGICAL AWARENESS				RAN		
Group	TEST						Words	Pseudowords	**Mean**		Syllable deletion	Phoneme deletion	Phonological sensitivity	**Mean**	Pictures	Digits	**Mean**
			(yrs,months)			% correct	% correct			correct/min	% correct				sec		
Very poorly literates		**Valid**	13	13	13	**13**	6	4	**4**	**13**	13	13	13	**13**	13	13	**13**
		Mean	**50.54**	**21.77**	**21.31**	**96.32**	77.31	42.19	**54.07**	**11.23**	46.15	5.39	42.31	**31.28**	36.72	28.86	**32.79**
		SD	[10.66]	[7.42]	[3.38]	**[4.64]**	[26.67]	[20.65]	**[21.15]**	**[9.22]**	[38.63]	[14.5]	[24.17]	**[18.76]**	[4.22]	[6.5]	**[4.5]**
		Range	28 – 63	13 – 36	17 – 27	**86.96 – 100**	27.78 – 100	18.75 – 62.5	**23.26 – 68.4**	**0 – 36**	0 – 100	0 – 50	0 – 83.33	**5.56 – 66.11**	31.75 – 44.4	22.27 – 44.8	**27.41 – 44.6**
Poorly literates		**Valid**	16	16	16	**16**	16	16	**16**	**16**	16	16	16	**16**	16	16	**16**
		Mean	**46.63**	**23.75**	**24.81**	**97.55**	85.76	77.34	**81.55**	**27.63**	66.25	16.88	64.58	**49.24**	35.18	26.53	**30.85**
		SD	[4.83]	[5.62]	[2.81]	**[4.76]**	[10.92]	[10.43]	**[9.01]**	**[14.69]**	[41.29]	[30.71]	[38.43]	**[29.91]**	[5.77]	[4.44]	**[4.77]**
		Range	36 – 55	17 – 39	19–30	**82.61 – 100**	66.67 – 100	62.5 – 100	**70.14– 94.44**	**6 – 54**	0 – 100	0 – 100	0 – 100	**0 – 100**	23.73 – 46	18.26 – 32.88	**21 – 39**
Schooled literates		**Valid**	16	16	16	**16**	16	16	**16**	**16**	16	16	16	**16**	16	16	**16**
		Mean	**39.75**	**26.88**	**26.88**	**99.19**	98.44	92.97	**95.7**	**65.69**	95.00	84.38	68.75	**82.71**	29.15	18.24	**23.7**
		SD	[7.13]	[6.93]	[1.67]	**[2.37]**	[2.48]	[6.8]	**[3.99]**	**[16.84]**	[15.06]	[19.99]	[18.13]	**[14.04]**	[7.91]	[2.54]	**[4.62]**
		Range	25 – 52	16 – 40	24 – 29	**91.3 – 100**	91.67 – 100	75 – 100	**86.11– 100**	**28 – 107**	40 – 100	20 – 100	33.33 – 100	**53.33 – 100**	18 – 48.7	13.5 – 23.28	**15.75 – 33.8**
Grade 2 children		**Valid**	52			**52**	52	52	**52**	**52**	52	52	52	**52**	51	49	**48**
		Mean	**7.94**			**91.67**	56.76	54.96	**55.86**	**21.73**	87.31	55.39	37.82	**60.17**	55.04	36.49	**46**
		SD	[0.37]			**[23.85]**	[35.45]	[35.31]	**[34.88]**	**[14.65]**	[16.46]	[41.65]	[34.63]	**[23.08]**	[17.17]	[15.05]	**[13.85]**
		Range	7.25 – 8.75			**0 – 100**	0 – 97.44	0 – 100	**0 – 96**	**0 – 55**	0 – 100	0 – 100	0 – 100	**0 – 100**	29 – 110	20 – 112	**28 – 90**
ANOVA on adults[a]			6.802***	5.177(*)	15.341***	**2.351**	11.288***	20.542****		**57.233******	11.198****	76.731****	5.281**		6.141***	30.409****	
	η^2_p		.262		.43	**.081**	0.305	0.703		**.732**	.273	.71	.142		.226	.51	
	t (30)[b]	Poorly literates vs. schooled literates	2.533*		-2.194*		-2.876*	-4.229****		**-7.563******	-2.434*	-8.219****	-0.414		2.722*	5.062****	
	t (27)	Very poorly literates vs. poorly literates	-1.365		3.53***		1.416	6.018****		**3.085*****	1.611	1.325	2.094(*)		-0.657	-1.351	
	t (27)	Very poorly literates vs. schooled literates	-3.763***		5.607****		3.541***	8.693****		**10.246******	3.915****	9.107****	2.486*		-3.233**	-6.144****	
ANOVA on adults vs. Grade 2 children						**< 1**	12.685****	11.318***		**5.046****	8.431***	24.941****	3.799*		16.978****	4.793**	
	η^2_p						.14	.10		**.115**	.256	.257	.089		.306	.11	
	t (66)	Poorly literates vs. Grade 2 children					3.244***	2.522*		**-1.477**	2.717*	3.669****	2.751*		4.893****	2.771*	
	t (63)	Very poorly literates vs. Grade 2 children					1.525	< 1		**2.425***	4.896***	4.391****	< 1		4.164****	1.96	

(*) p < .10, * p < .05,** p < .01 ,*** < .005, **** < .001

a: all values are $F(2, 42)$ in the adult analyses (or $F(2, 35)$ and $F(2, 33)$ for word and pseudoword reading, respectively), and $F(2, 78)$ for the children analyses (or $F(2, 71)$ and $F(2, 69)$ for word and pseudoword reading, respectively), except for SES, for which we ran a one-way analysis of variance on ranks (Kruskall-Wallis test). Welch correction was applied when necessary in the parametric ANOVAs.

b: All post-hoc t tests were Holm-corrected.

3.2.1.2 Categorical perception and precision

The labeling functions, as a function of stimulus, are presented in Figure 3, separately for each group; discrimination performance, as a function of pair, is displayed in Figure 4, separately for each group.

The Pair X Task X Group (very poorly literates, poorly literates, schooled literates) repeated measure ANOVA conducted on *d'* showed significant main effects of Pair, $F(1.3, 54.6) = 303.181$, $p < .001$, $\eta^2_p = .878$, with S3S4 leading to the best performance compared to all other pairs, all $ts(44) > 27$, all ps < .001 (ps = 1 for all other comparisons). There was also a significant main effect of Group, $F(2, 43) = 9.862$, $p < .001$, $\eta^2_p = .32$, with very poorly literates displaying overall lower performance than both schooled literates and poorly literates, $t(27) = 4.428$, $p < .001$ and = 2.732 , < .025, respectively, and poorly literates only tending to differ from schooled literates, $t(30) = 1.79$, $p < .10$. In addition, there was a significant Pair X Group interaction, $F\ (2.6, 54.6) = 7.054$, $p < .001$, $\eta^2_p = .251$, which reflects the fact that groups differed significantly only on the S3S4 pair, $F(2, 168) = 8.914$, $p < .001$ (all other $ps \geq .10$, except on S2S3, $F(2, 168) = 2.584$, $p = .09$). On the S3S4 pair, very poorly literates displayed lower performance than both schooled literates and poorly literates, $t(27) = 8.608$ and = 4.525, both ps < .001, with the latter also differing from schooled literates, $t(30) = 4.313$, $p < .005$. However, the Pair X Task X Group interaction was not significant, $F(2.7, 57.5) = 1.16$, $p = .33$, $\eta^2_p = .052$.[9]

[9] Results were similar with age as covariate, with a significant Pair X Group interaction, $F\ (2.6, 53.2) = 5.377$, $p < .005$, $\eta^2_p = .208$ and a nonsignificant Pair X Task X Group interaction, $F < 1$.

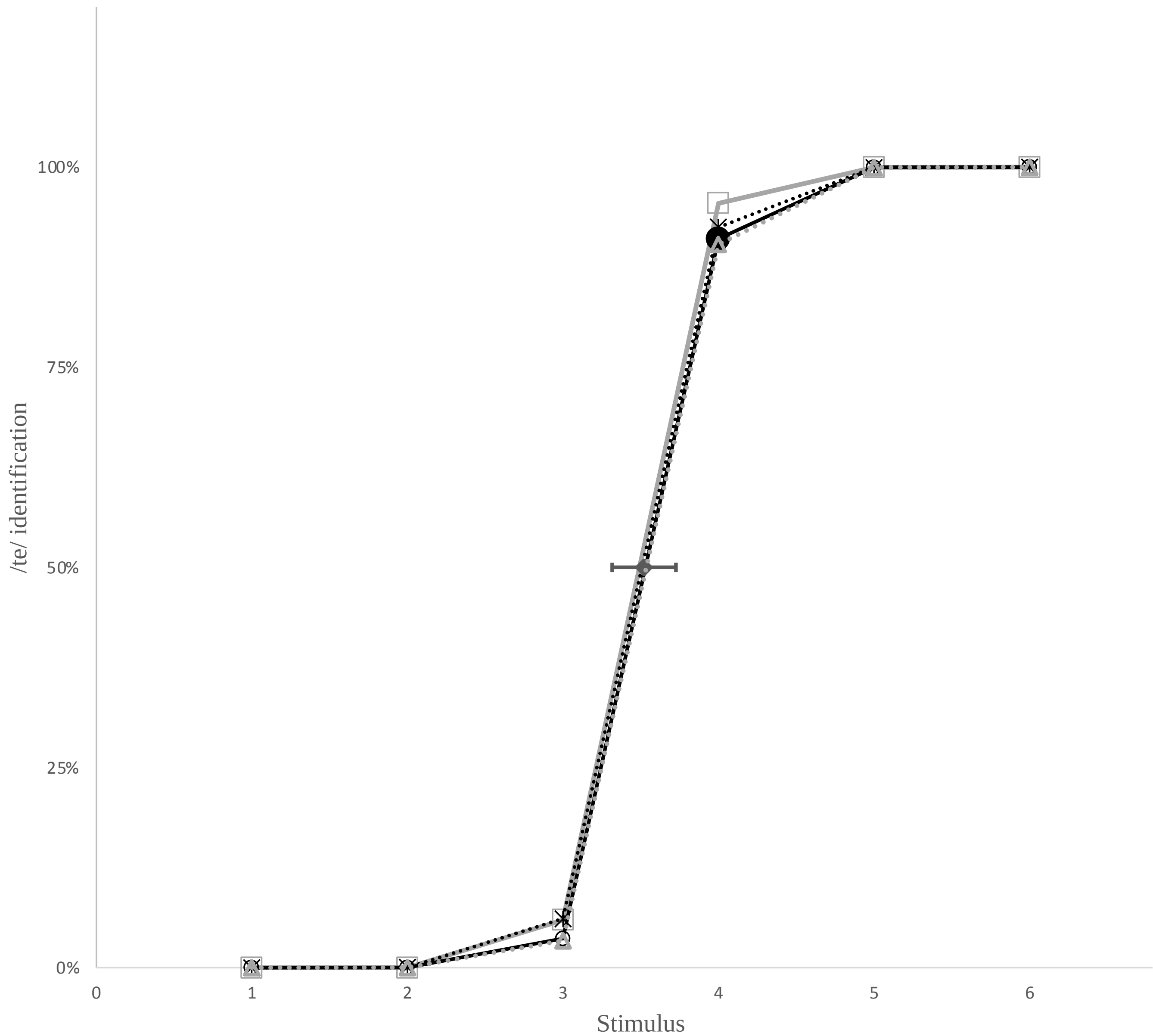
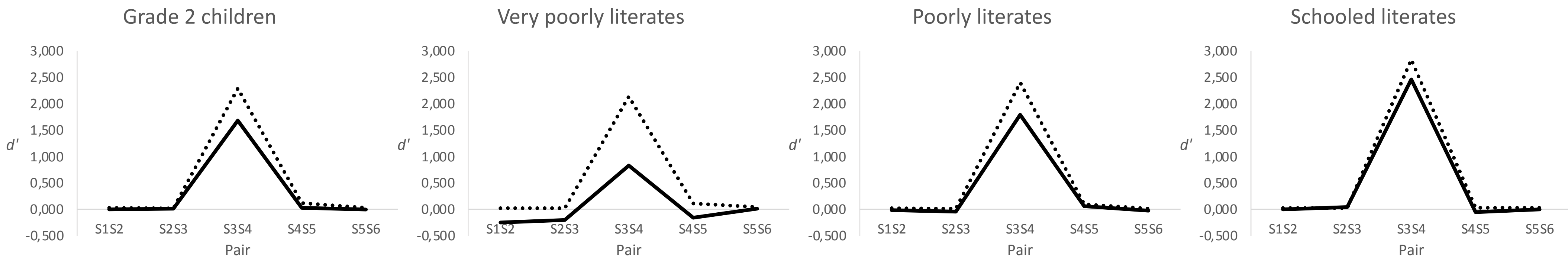


Figure 3



Grade 2 children
Very poorly literates
Poorly literates
Schooled literates
3,000
2,500
2,000
1,500
1,000
0,500
0,000
-0,500
d'
S1S2
S2S3
S3S4
S4S5
S5S6
Pair

Figure 4

A similar Bayesian repeated measure ANOVA supported the idea that there is evidence in favor of a group difference in categorical precision, with a Bayesian factor BF_{incl} = 1.583E+7 for the Pair X Group interaction[7], as well as in favor of no group difference in categorical perception, with a Bayesian factor BF_{excl} = 11.378 for the Pair X Task X Group interaction. In other words, the data are about 15 million times more likely under the hypothesis that there is a group difference in categorical precision than under H_0, which is considered as very strong (Raftery, 1995) or decisive (Jeffreys, 1961) evidence for H_1 over H_0. Yet they are about 11 times more likely under the hypothesis that group do not differ on categorical perception than under H_1, which is considered as positive (Raftery, 1995) or strong (Jeffreys, 1961) evidence for H_0 over H_1.

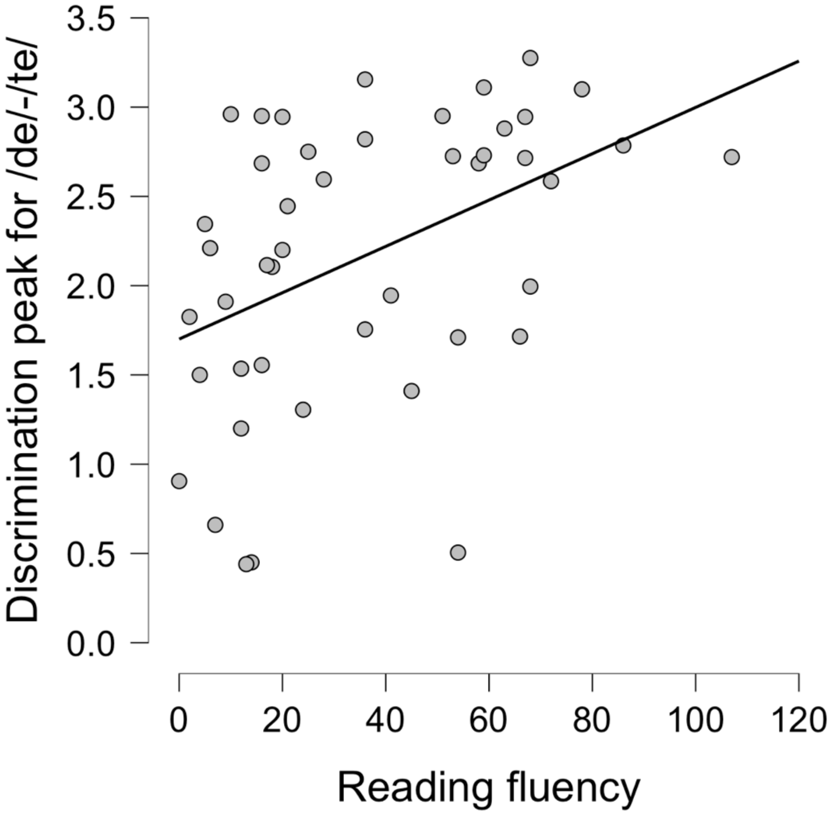
Table 4 presents the average values of the speech perception indexes that were used in the computation of correlations. No significant correlation was observed between the literacy scores and the slope of the identification function, all $ps \geq 1$, probably due to a ceiling effect on this parameter (Figure 3). On the contrary, the discrimination peak does correlate with the literacy scores, with highly significant correlations with reading fluency and phonological awareness, $\tau(43) = .32$, = .35, respectively, both ps ≤ .005 (see Figure 5), and marginally significant correlations with average reading and with letter knowledge, $\tau(34) = .30$ and = .28, respectively, both *ps* < .05 (correlation with RAN: $\tau(43) = -.22$, $p < .10$). More conservative Bayesian correlation tests showed strong evidence in favor of H_1 as regards the correlations between the discrimination peak and literacy scores (with the exception of RAN, BF_{0-} = 3.377), with BF_{0+} = 39.927, = 98.26, = 10.502 and = 14.683 for reading fluency, phonological awareness, average reading, and letter knowledge, respectively. On the contrary, in agreement with the lack of three-way significant interaction in the ANOVA, no correlation between the literacy scores and the categorical perception index reached significance, all *p*s ≥ .11. Nonetheless, Bayesian correlation tests showed no evidence in favor of H_0, either, all BFs $_{0+}$ or BFs $_{0-}$ ≤ 2.436. In fact, there was a trend toward a correlation between the categorical perception index and reading fluency only in the poorly literates, $\tau(14) = -.41$, $p = .08$ (very poorly literates and schooled literates: both *p*s > 1), although the evidence in favor of H_1 was anecdotal according to a Bayesian correlation test, $BF_{+0} = .10$. Besides, age did not correlate significantly with any of the speech perception indexes, all *p*s ≥ .44 (bilateral tests). Yet, here too the evidence in favor of H_0 was anecdotal according to Bayesian tests of correlation, BFs $_{01}$ ≤ 3.126.

Table 4

Experiment 2: average values of the speech perception indexes used in the computation of correlations. Standard deviation in brackets.

		Grade 2 children	Very poorly literates	Poorly literates	Schooled literates
Slope	**Mean**	**-5.59**	**-5.25**	**-5.6**	**-5.82**
	SD	[1.07]	[1.84]	[0.9]	[0.73]
	Range	-6 – -0.92	-6 – -1.07	-6 – -3.14	-6 – -3.08
Categorical precision index	**Mean**	**1.97**	**1.64**	**2.12**	**2.66**
	SD	[0.82]	[0.85]	[0.75]	[0.47]
	Range	0.28 – 3.28	0.44 – 2.96	0.505 – 3.16	1.71 – 3.28
Categorical perception index	**Mean**	**1**	**1.67**	**0.94**	**0.97**
	SD	[0.68]	[0.92]	[0.76]	[0.77]
	Range	0.01 – 2.61	0.12 – 3.13	0.01 – 2.35	0.04 – 3.12

A


Discrimination peak for /de/-/te/
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
20
40
60
80
100
120
Reading fluency

B

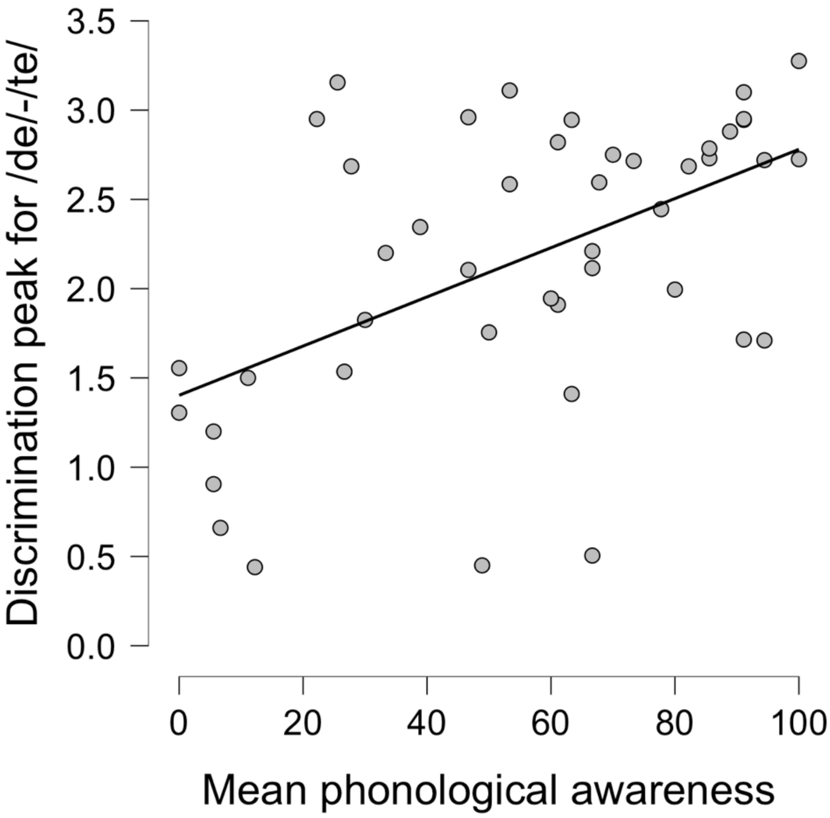

Discrimination peak for /de/-/te/
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
20
40
60
80
100
Mean phonological awareness

Figure 5

3.2.2 Comparisons between the Grade 2 children and the unschooled adults

3.2.2.1 Ancillary tests

Performance on the ancillary tests were analyzed in separate ANOVAs, each with group (Grade 2 children vs. poorly literates vs. very poorly literates) as a between-subjects factor. As shown in Table 3, the group effect was significant in all these analyses, except for letter knowledge. Poorly literates presented better reading fluency, word reading and pseudoword reading scores than Grade 2 children, which was not the case of the very poorly literates, who even presented somewhat lower reading fluency scores than the children. In addition, both adult groups struggled on phoneme (and even syllable) deletion, presenting poorer performance than the children, which is a common observation in adults with low levels of literacy (e.g., Eme et al., 2014; Greenberg et al., 1997; Thompkins & Binder, 2003). They were however slightly more rapid than children on the RAN tests.

3.2.2.1 Categorical perception and precision

The children's mean identification (labeling) function are presented in Figure 3, and detailed discrimination performance, as a function of pair, is displayed in Figure 4.

The repeated measure ANOVA on the *d'* response scores, with Pair, Task and Group (Grade 2 children, very poorly literates, poorly literates) as factors, showed significant main effects of Pair, $F(1.27, 99.24) = 235.398$, $p < .001$, $\eta^2_p = .751$, with S3S4 leading to the best performance compared to all other pairs, all $ts(80) > 23$, all $ps < .001$ ($ps = 1$ for all other comparisons), and Group, $F(2, 78) = 4.243$, $p <.025$, $\eta^2_p = .098$, with very poorly literates displaying overall lower performance than both poorly literates, $t(27) = 2.476$, $p < .05$ and Grade 2 children, $t(63) = 2.792$, $p < .025$. The interactions Pair X Group and Pair X Task X Group were both non-significant, $F(2.55, 99.24) = 1.56$, $p = .21$, $\eta^2_p = .038$ and $F = 1$, respectively.

A similar Bayesian repeated measure ANOVA supported the idea that there is positive (Raftery, 1995) or substantial (Jeffreys, 1961) evidence in favor of the hypothesis that there is no group difference in categorical precision, with BF_{excl} = 4.214 for the Pair X Group interaction, and strong (Raftery, 1995) to very strong (Jeffreys, 1961) evidence in favor of the hypothesis that there is no group difference in categorical perception, with BF_{excl}= 49.631 for the Pair X Task X Group interaction.

Consistently, age did not correlate[10] with any of the speech perception indexes, all *p*s ≥ .55 (see Table 4 for average values of those indexes). Bayesian correlations showed that there is only anecdotal evidence in favor of H_0 as regards the correlation between age and categorical precision, BF_{0+} = 7.602, but strong evidence in favor of H_0 as regards the correlation between age and categorical perception, BF_{0-} = 17.17.

No significant correlation between literacy scores and slope was observed, all *p*s ≥ .41 (see above). On the contrary, literacy scores did correlate with the discrimination peak, with $\tau(79)$ = .31, $p < .005$ for letter knowledge, = .24, $p = .01$ for average reading, = .19, $p < .05$ for metaphonology, and = .18, $p < .05$ for reading fluency (on RAN, $\tau(78)$ = -.05, $p > .20$). The more conservative Bayesian correlation test values showed very strong evidence in favor of H_1 as regards the correlation between the discrimination peak and letter knowledge, BF_{+0} = 1278.48, and strong evidence in favor of H_1 as regards the correlation between the discrimination peak and average reading, BF_{+0} = 21.67 (metaphonology: BF_{-0} = 7.05). No correlation between the literacy scores and the categorical perception index reached significance, all *p*s ≥ .37, except a marginal correlation with reading fluency, $\tau(79)$ = -.183, $p < .05$. Yet, this correlation was not significant in children, $\tau(50)$ = .09, $p = .85$, and was only due to the poorly literates, as already commented on. In addition, on all participants, the Bayesian correlation test only showed anecdotal evidence supporting a positive correlation,

[10] Here, one-sided correlations were used.

BF_{-0} = 5.2. Nonetheless, as regards literacy scores and categorical perception, there was no evidence in favor of H_0, either, all BFs $_{01} \leq 7.606$.

4. General Discussion

The aim of the present study was to reevaluate the impact of literacy on categorical precision and categorical perception by using more controlled materials than in previous studies. Indeed, one study on adults using a place-of-articulation continuum (Serniclaes et al., 2005) and one study on children and adults using a voicing continuum (Hoonhorst et al., 2011) had reported a literacy effect on categorical precision, although no effect was observed on categorical perception. Yet, these previous studies had used synthesized stimuli in which only the primary acoustic cues to either place of articulation (F2 and F3 transitions, Serniclaes et al., 2005) or voicing (VOT, Hoonhorst et al., 2011) were modified, a procedure that produces conflict with secondary cues that remain unmodified. This conflict may itself hinder the precision of the categorical prototypes inside each category. In these studies, the use of stimuli with conflicting acoustic cues was thus probably at the origin of an effect of literacy on the precision of the categorical prototypes rather than on the precision of the boundary. In addition, in the adult study, the authors reported a lexical bias that may also have affected the results.

We thus presented adults (Experiment 1) and both children and adults (Experiment 2) with a material without lexical bias and without conflict between acoustic cues, as it was generated by morphing between natural speech exemplars pronounced by a native speaker. We used both a place-of-articulation continuum, as did Serniclaes et al. (2005) and a voicing continuum, as did Hoonhorst et al. (2011).

We further aimed at examining more thoroughly the impact of literacy than the former studies by examining one group of children (Experiment 2) and three groups of adults (Experiments 1 and 2) of varying literacy levels, which were carefully evaluated. Indeed, in addition to Grade 2 children, we examined three groups of adult participants: adults who had attended school and hence had learned there to read and write in childhood, and adults who had not (or almost not) attended school in childhood but were attending adult literacy classes at the time of testing and thus were beginning readers. The latter, unschooled, participants were assigned either to a group of poorly literates or to a group of very poorly literates as a function of their reading proficiency.

With the place-of-articulation continuum used in Experiment 1, we observed that the three adult groups differed neither in categorical precision, nor in categorical perception. Bayesian analyses even showed strong evidence in favor of the hypothesis that there is no group difference in categorical precision or categorical perception. Consistently, no significant correlation was observed between the literacy scores and either the categorical perception or the categorical perception indexes. This suggests that the effect reported by Serniclaes et al. (2005) was either due to the lexical bias they mentioned, or to the use of synthetic stimuli with conflicting acoustical cues, leading to an effect of literacy on the precision of the categorical prototypes rather than on the precision of the boundary.

With the voicing continuum used in Experiment 2, the adults' results were similar to those of Experiment 1 as regards categorical perception, but quite different as regards categorical precision. Indeed, as in Experiment 1, the three adult groups did not differ significantly in categorical perception, and Bayesian analyses even showed strong evidence supporting the hypothesis that there is actually no group difference. Consistently, there was no significant correlation between the literacy scores and the categorical perception index. Much on the opposite, the three adult groups differed strongly on categorical precision, with significant correlations between the discrimination peak and the literacy scores (with the exception of RAN).

In Experiment 2, the comparison of the two groups of unschooled adults to Grade 2 children reinforced the idea that literacy impacts only categorical precision, not categorical perception, and further showed that this effect is independent of age. Indeed, no significant group difference was observed in either categorical precision or categorical perception, and Bayesian analyses even showed positive evidence in favor of the hypothesis that there is no group difference in categorical precision, and strong evidence in favor of the hypothesis that there is no group difference in categorical perception. Consistently, age did not correlate with any of the speech perception indexes. On the contrary, literacy scores did correlate with the discrimination peak, but not with the categorical perception index.

The observation that literacy impacts categorical precision only with the voicing continuum, not with the place-of-articulation one, was unexpected. Indeed, in typical development of Brazilian Portuguese natives' speech production, voicing contrasts are mastered before place of articulation contrasts (Azevedo, 1995; Jardim-Azambuja, 2004), and hence, on the basis of age of acquisition, we would have expected the reverse result pattern. Yet, one major difference between voicing and place of articulation is that voicing perception depends on the temporal relationship with other phonological features. The presence/absence of "voice" (i.e., of periodic vibrations) that characterizes voicing is not phonologically relevant unless it occurs in the vicinity of an obstruent consonant, a temporal dependency that is captured by the VOT concept (Lisker & Abramson, 1964). On the contrary, place of articulation is characterized by spectral cues, mainly the burst spectrum and the direction of formant transitions (Stevens & Blumstein, 1978), that are phonologically exploitable without referring to other features. Such difference between place of articulation and voicing probably has implications for their relationships with phoneme perception. Remembering that phonemes are classically defined as "bundles" of distinctive features (Jakobson, Fant & Halle, 1952; Chomsky & Halle, 1968), voicing perception should be more dependent than place of articulation on the phonemic framework. Factors that consolidate phoneme percepts, including literacy, should thus mainly impact voicing perception. This might explain why the effect of literacy on perceptual precision was only evidenced for voicing, not for place of articulation.

The present results clearly confirm that literacy has no effect on the categorical perception of phonological features. However, the results open a new perspective for conceiving a possible effect of literacy on phoneme perception. One main achievement of the present study was to evidence an effect of literacy on the precision of the perceptual boundary. Such effect is quite different of the differences in the precision of category prototypes that were evidenced in previous studies (Hoonhorst et al., 2011; Serniclaes et al., 2005). The fact that the effect of literacy on boundary precision was only evidenced here for voicing, and not for place of articulation, might be related to the difference in the phonemic status of these two features. Voicing perception is intrinsically dependent on the relationships with other features and should therefore be more influenced than other features by the factors that contribute to improve phoneme perception, including literacy. The effect of literacy on voicing perception would thus participate to a more general effect of literacy on phoneme perception. This conjecture offers new perspectives to investigate the possible consequences of literacy on the binding between different features that are at the root of phoneme perception.

Studying such effects more thoroughly is important to grasp how exactly literacy impacts phoneme perception and may thus also contribute to understand why literacy strongly enhances the activation of regions involved in phonological processing. Indeed, studies using functional magnetic resonance imaging (fMRI) reported that activation of the planum temporale (PT) is much stronger in literate compared to non-literate individuals listening to spoken sentences, this effect being observed in both adults (Dehaene et al., 2010) and children

(Monzalvo & Dehaene-Lambertz, 2013). This enhancement of PT activation is most probably not due to direct on-line feedback from either orthographic or conscious representations of spoken inputs. As a matter of fact, on the one hand the PT activates only in response to auditory input, not to written words (Dehaene et al., 2010) or isolated letters (van Atteveldt et al., 2004); on the other hand, phonological awareness tasks typically involve inferior frontal areas (e.g., Burton et al., 2000; Zatorre et al., 1996). Rather, the PT, as well as the surrounding superior temporal cortex, probably houses relatively abstract phonemic representations, as it encodes acoustic changes that are crucial for the categorical perception of speech (e.g., Chang et al., 2010; Mesgarani et al., 2014). The increase in PT activation found in literates compared to non-literates may therefore indicate that reading acquisition refines that kind of abstract phonological coding. What the present results suggest, however, is that not all phonemic representations would be equally susceptible to the effect of literacy.

5. Acknowledgments

R. Kolinsky is Research Director of the Fonds de la Recherche Scientifique–FNRS (FRS–FNRS), Belgium. This research was supported by a grant attributed by the São Paulo Research Foundation to Ana Luiza Navas (FAPESP 2016/10754-0), and preparation of this paper was also supported by a Concerted Research Action grant of the Belgian French community attributed to R. Kolinsky and O. Klein (*The Socio-Cognitive Impact of Literacy*).

6. References

Adrián, J. A., Alegria, J., & Morais, J. (1995). Metaphonological abilities of Spanish illiterate adults. *International Journal of Psychology, 30,* 329–353.

Araújo, S., Fernandes, T., & Huettig, F. (2019). Learning to read facilitates the retrieval of phonological representations in rapid automatized naming: Evidence from unschooled illiterate, ex- illiterate, and schooled literate adults. *Developmental Science, 22*:e12783. https://doi.org/10.1111/desc.12783.

Araújo, S., Reis, A., Petersson, K. M., & Faísca, L. (2015). Rapid automatized naming and reading performance: A meta- analysis. *Journal of Educational Psychology, 107*, 868–883. https://doi.org/10.1037/edu0000006.

Azevedo, C. (1995). Acquisição normal e com desvios da fonologia do Português – contrastes de sonoridade e de ponto de articulação. [Normal and deviated acquisition of the phonology of Portuguese – voicing and lace of articulation contrasts]. *Letras de Hoje, 30*(4), 285-294.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. Glot International 5:9/10, 341-345.

Boersma, P. & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.56, retrieved 20 June 2019 from http://www.praat.org/

Burnham, D. K. (2003). Language specific speech perception and the onset of reading. *Reading and Writing, 16*, 573–609.

Burnham, D.K., Earnshaw, L.J. & Clark, J.E. (1991). Development of categorical identification of native and non-native bilabial stops: infants, children and adults. *Journal of Child Language, 18,* 231-260.

Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience, 12*, 679–690. https://doi.org/10.1162/089892900562309.

Cahan, S., & Davis, D., (1987). A between-grades-level approach to the investigation of the absolute effects of schooling on achievement. *American Educational Research Journal*, *24*, 1-13.

Caramelli, P., Herrera, J.R. E., Nitrini, R. (1999). O mini-exame do estado mental no diagnóstico de demência em idosos analfabetos. [The Mini Mental State

Examination in the diagnosis of dementia in illiterate elderly people]. *Arquivos de Neuro-Psiquiatria*, *57*(11),7-12.

Castro, S. L. (1988). *Alfabetização e percepção da fala: contribuição experimental para o estudo dos efeitos do conhecimento da escrita em aspectos do processamento da linguagem falada*. [Literacy and speech perception: experimental contribution to the study of the effects of writing knowledge on aspects of spoken language processing]. PhD dissertation. Universidade do Porto. https://sigarra.up.pt/flup/pt/pub_geral.pub_view?pi_pub_base_id=29881&pi_pub_r1_id=.

Castro, S. L. (1993). *Alfabetização e percepção da fala*. Porto, Portugal: Instituto Nacional de Investigação Científica.

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain,121*, 1053–1063. https://doi.org/10.1093/brain/121.6.1053.

Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience, 13*, 1428–1432. https://doi.org/10.1038/nn.2641.

Chomsky, N., & Halle, M. (1968). *The sound pattern of Engli*sh. New York: Harper and Row.

Damper, R. I., & Harnad, S. R. (2000). Neural network models of categorical perception. *Perception & Psychophysics*, *62*, 843–867. https://doi.org/10.3758/bf03206927.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330*, 1359–1364. https://doi.org/10.1126/science.1194140.

Eme, E., Lambert, E., & Alamargot, D. (2014). Word reading and word spelling in French adult literacy students: the relationship with oral language skills. *Journal of Research in Reading*, *37*(3), 268–296. https://doi.org/10.1111/j.1467-9817.2011.01508.x.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-Mental State”: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189-198.

Gottfried, T.L., & Strange, W. (1980). Identification of coarticulated vowels. *The Journal of the Acoustical Society of America*, *68*, 1626-1635.

Greenberg, D., Ehri, L. C. & Perin, D. (1997). Are word-reading processes the same or different in adult literacy students and third–fifth graders matched for reading level? *Journal of Educational Psychology*, *89*(2), 262–275. https://doi.org/10.1037//0022-0663.89.2.262.

Guerreiro, M., Silva, A. P., Botelho, M., Leitão, O., Castro-Caldas, A., & Garcia, C. (1994). Adaptação à população portuguesa da tradução do Mini Mental State Examination (MMSE) [Adaptation to the Portuguese population of the translation of the Mini Mental State Examination (MMSE)]. *Revista Portuguesa de Neurologia*, *1*, 9-10.

Haggard, M., Summerfield, Q., & Roberts, M. (1981). Psychoacoustical and cultural determinants of phoneme boundaries: Evidence from trading F0 cues in the voiced–voiceless distinction. *Journal of Phonetics*, *9*, 49–62

Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2011). Categorical perception of voicing, colors and facial expressions:

A developmental study. *Speech Communication*, *53*, 417–430. https://doi.org/10.1016/j.specom.2010.11.005.

Horlyck, S., Reid, A., & Burnham, D. (2012). The relationship between learning to read and language-specific speech perception: Maturation versus experience. *Scientific Studies of Reading*, *16*(3), 218-239. https://doi.org/10.1080/10888438.2010.546460.

Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to speech analysis. The distinctive features and their correlates*. Cambridge, Mass.: M.I.T. Press.

Jardim-Azambuja, R. (2004). *Estudo longitudinal sobre a emergência dos contrastes de sonoridade e de ponto de articulação na aquisição fonológica do português brasileiro _ crianças de 1:0 a 1:6*. [Longitudinal study on the emergence of the voicing and place of articulation contrasts in the phonological acquisition of Brazilian Portuguese – children from 1:0 to 1:6]. Master Dissertation. Porto Alegre: PUCRS.

JASP Team (2020). *JASP* (Version 0.13.1) [Computer software].

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Kawahara, H., & Matsui, H. (2003). Auditory morphing based on an elastic perceptual distance metric in an interference-free time-frequency representation. ICASSP, *IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings*, *1*, 256–259. https://doi.org/10.1109/icassp.2003.1198766.

Kawahara, H., & Morise, M. (2011). Technical foundations of TANDEM-STRAIGHT, a speech analysis, modification and synthesis framework. *Sadhana*, *36*(5), 713–727. https://doi.org/10.1094/PDIS-08-15-0933-PDN.

Kawahara, H., Nisimura, R., Irino, T., Morise, M., Takahashi, T., & Banno, H. (2009). Temporally variable multi-aspect auditory morphing enabling extrapolation without objective and perceptual breakdown. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings*, 3905–3908. https://doi.org/10.1109/ICASSP.2009.4960481.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358–368.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18*, 201–212.

Lisker, L. & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*(3), 384-422.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). London: Lawrence Erlbaum.

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science, 343*, 1006–1010. https://doi.org/10.1126/science.1245994.

Monzalvo, K., & Dehaene-Lambertz, G. (2013). How reading acquisition changes children's spoken language network. *Brain and Language, 127*, 356–365. https://doi.org/10.1016/j.bandl.2013.10.009.

Morais, J., Alegria, J., & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy: An interactive view. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 7*, 415-438.

Morais, J., Araújo, L., Leite, I., Carvalho, C., Fernandes, S., & Querido, L. (2010). Relatório final, Jan 2008 – Out 2010. *Estudo Psicolinguístico. Estabelecimento de Níveis de Referência do Desenvolvimento da Leitura e da Escrita do 1o ao 6o ano de Escolaridade*, Plano Nacional de Leitura do Ministério da Educação, Portugal [Final report Jan 2008 - Oct 2010 study Psycholinguistic. Establishment of Reading and writing Development Reference Levels from the first to the 6th year of Education, National Reading Plan of the Ministry of Education, Portugal]. http://www. planonacionaldeleitura.gov.pt/PNLEstudos/uploads/ficheiros/psico_15fev.pdf.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech analysis. *Cognition*, *24*, 45-64.

Morais, J., Cary, L., Alegria, J. & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*, 323-331.

Morais, J., Castro, S. - L., Scliar-Cabral, L., Kolinsky, R., & Content, A. (1987). The effects of literacy on the recognition of dichotic words. *Quarterly Journal of Experimental Psychology, 39 A*, 451-465.

Morrison, F. J., Smith, L., & Dow-Ehrensberger, M. (1995). Education and cognitive development: A natural experiment. *Developmental Psychology, 31*(5), 789. https://doi.org/10.1037/0012-1649.31.5.789.

Paula, F. V. (2007). *Conhecimento morfológico implícito e explícito na linguagem escrita* [Implicit and explicit morphological knowledge in written language]. Unpublished PhD dissertation, Universidade de São Paulo, São Paulo & Université de Rennes 2, Rennes, França.

Pollack, I., & Pisoni, D. (1971). On the comparison between identification and discrimination tests in speech perception. *Psychonomic Science*, *24*, 299–300.

Powell, D., Stainthorp, R., Stuart, M., Garwood, H., & Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology, 98*(1), 46–68. https://doi.org/10.1016/j.jecp.2007.04.003.

Raftery, A. E. (1995). Bayesian model selection in social research. In P. V. Marsden (Ed.), *Sociological methodology* (pp. 111–196). Cambridge, MA: Blackwell.

Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition*, *24*, 31–44.

Repp, B. H. (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Language and Speech, 22*, 173–189.

Schouten, B., Gerrits, E., & van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication, 41*, 71–80. https://doi.org/10.1016/S0167-6393(02)00094-8.

Serniclaes, W., Ventura, P., Morais, J., & Kolinsky, R. (2005). Categorical perception of speech sounds in illiterate adults. *Cognition, 98*, B35–B44. https://doi.org/10.1016/j.cognition.2005.03.002.

Scliar-Cabral, L., Morais, J., Nepomuceno, L., & Kolinsky, R. (1997). The awareness of phonemes: So close-so far away. *International Journal of Psycholinguistics*, *13*, 211–240.

Simon, C., & Fourcin, A. J. (1978). Cross-language study of speech-pattern learning. *Journal of the Acoustical Society of America*, *63*, 925–935.

Stevens, K. N. & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, *64*, 1358-1368.

Stevens, K. N., Klatt, D. H. (1974). Role of formant transitions in the voiced–voiceless distinction for stops. *Journal of the Acoustical Society of America, 55*, 653–659.

Summerfield, Q. (1982). Differences between spectral dependencies in auditory and phonetic temporal processing: Relevance to the perception of voicing in initial stop. *Journal of the Acoustical Society of America, 72*, 51–61.

Thompkins, A. C. & Binder, K. S. (2003). A comparison of the factors affecting reading performance of functionally illiterate adults and children matched by reading level. *Reading Research Quarterly, 38*(2), 236–258. https://doi.org/10.1598/RRQ.38.2.4.

van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron, 43*(2), 271–282. https://doi.org/10.1016/j.neuron.2004.06.025.

Ventura, P., Kolinsky, R., Fernandes, S., Querido, L., & Morais, J. (2007). Lexical restructuring in the absence of literacy. *Cognition, 105*, 334–361. https://doi.org/10.1016/j.cognition.2006.10.002.

Werker, J. F. & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology, 37,* 278-286.

Wood, C.C. (1976). Discriminability, response bias, and phoneme categories in discrimination of voice onset time. *Journal of the Acoustical Society of America, 60*,1381–1389.

Zatorre, R., Meyer, E., Gjedde, A., & Evans, A. (1996). PET studies of phonetic processing of speech: Review, replication and reanalysis. *Cerebral Cortex, 6*, 21–30. https://doi.org/10.1093/cercor/6.1.21.

7. Figure captions

Figure 1. Experiment 1: labeling functions recalculated from the mean boundaries of each group. Very poorly literates: plain black line; poorly literates: dotted black line; schooled literates: plain gray line. Horizontally, one sees the overall mean of the identification boundary, with an error bar of one standard deviation.

Figure 2. Experiment 1: average discrimination scores in each group, for observed performance (plain line) and performance predicted on the basis of identification (dotted line).

Figure 3. Experiment 2: labeling functions recalculated from the mean boundaries of each group. Very poorly literates: plain black line; poorly literates: dotted black line; schooled literates: plain gray line; Grade 2 children: dotted gray line. Horizontally, one sees the overall mean of the identification boundary for the adult groups, with an error bar of one standard deviation.

Figure 4. Experiment 2: average discrimination scores in each group, for observed performance (plain line) and performance predicted on the basis of identification (dotted line).

Figure 5. Experiment 2: correlations between the discrimination peak and either word reading fluency (A) or mean phonological awareness (B).