Contextual diversity is a main determinant of word-identification times in young readers

Manuel Perea [1], Ana Paula Soares [2], and Montserrat Comesaña [2]

[1] ERI-Lectura and Departamento de Metodología, Universitat de València, Valencia, Spain

[2] Human Cognition Lab, CIPsi, School of Psychology, Universidade do Minho, Braga, Portugal

Running Head: Contextual Diversity in young readers

Word Count: 3998 words

Correspondence concerning this article should be addressed to: Manuel Perea, Departamento de Metodología, Facultad de Psicología, Av. Blasco Ibáñez, 21, 46010-Valencia, Spain (e-mail: mperea@valencia.edu).

Abstract

Recent research with college-aged skilled readers has revealed that contextual diversity (i.e., the number of contexts that a word appears) is a more critical determinant of visual-word recognition than mere repeated exposure (i.e., word frequency) (Adelman, Brown, & Quesada, 2006, *Psych.Sci.*). Given that contextual diversity has been claimed to be a relevant factor to word acquisition in developing readers, the effects of contextual diversity should also be a main determinant of word-identification times in developing readers. A lexical decision experiment was conducted to examine the effects of contextual diversity and word frequency in young readers (4th Grade children). Results revealed a sizeable effect of contextual diversity, but not of word-frequency, thus generalizing the Adelman et al. data to a children population. These findings call for the implementations of dynamic, developmental models of visual-word recognition that go beyond a learning rule by mere exposure.

Key words: word identification, word-frequency, lexical decision, developing readers

Word-identification times in laboratory tasks and eye fixation times during normal reading are shorter when reading a word that occurs frequently in print like heart than when reading an infrequent word like elbow. This is the case both in adult readers (e.g., see Balota et al., 2007; Inhoff & Rayner, 1996) and in young readers (e.g., Hyönä & Olson, 1995; Moret-Tatay & Perea, 2011).

Unsurprisingly, word-frequency plays a central role in all current computational models of visual-word recognition and reading. For instance, in the family of localist activation-based models, the resting level of activation of a given word unit depends on its printed word-frequency (e.g., spatial coding model, Davis, 2010; dual-route cascaded model, Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; interactive activation model, McClelland & Rumelhart, 1981; multiple read-out model, Grainger & Jacobs, 1996). Likewise, computational models of reading like the EZ-Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998) or the SWIFT model (Engbert, Nuthmann, Richter, & Kliegl, 2005) employ a similar mechanism of word-frequency in the initial "lexical access" stage.

One obvious shortcoming of these "static" models is that they assume a well-formed and unalterable set of parameters (as in an adult skilled reader), in which word-frequency is included as a critical element (see Baayen, 2010, for criticism of these models). Although there is the implicit assumption that the frequency effect has its origin in the repetition effect, these models provide no information about the underlying nature (or locus) of the word-frequency effect. Thus, one obvious question is how repeated presentations affect a word's accessibility during the course of word learning in an immature word recognition system (see Acha & Perea, 2008; Wang, Castles, Nickels, & Nation, 2011, for developmental analyses of orthographic effects in visual-word recognition).

Is word accessibility merely driven by a count of past presentations (i.e., word-frequency)? Recently, a number of theorists have claimed, using large databases collected from college-aged students, that the main determinant of word-identification times is not "pure repetition" per se, but rather the number of contexts in which a given word occurs (Adelman et al., 2006; see also Baayen, 2010; Brysbaert & New, 2009; Keuleers, Brysbaert & New, 2010, for similar conclusions). Clearly, words that appear in many contexts tend to be quite frequent in print, and this explains the previous reports of a word-frequency effect – they would obtained because of a confound with the number of contexts in which a word occurs. This latter factor has been named Contextual Diversity and it has been defined as "the proportion of texts in which a given word occurs" (Adelman et al., 2006, p. 814). Adelman et al. (2006) indicated that if a given word appears in very different contexts, the likelihood of that word appears in other contexts increases, and this may facilitate the relative accessibility of these words in the "mental lexicon".

Importantly, contextual diversity has also been claimed to be a relevant factor to word acquisition in developing readers (e.g., see Hills, Maouene, Riordan, & Smith 2010). Thus, if words which are initially learned in many contexts have more accessible and enriched internal representations than the words that are learned in a limited number of contexts, then the effects of contextual diversity should be particularly clear for developing readers–note that the children's word learning mechanisms may be even more important for word recognition than in the case of the more "stable" word representations in skilled adult readers (see Ratcliff, Love, Thompson, & Opfer, 2012, for evidence of less efficient word processing in children than in college-aged students).

In the present experiment, we examined the role of contextual diversity and word frequency in word identification times in young readers (4th Grade children). Unsurprisingly, it was not possible to obtain enough low-frequency words with high contextual diversity, and this latter condition was not included in the experiment. The hypotheses are clear. If mere repeated exposure (and not context diversity) is the driving force in lexical access (as stated in currently implemented computational models), then words which occur frequently in print –while keeping contextual diversity controlled– should produce faster word-identification times than words which occur least frequently (i.e., a "word-frequency" effect). Alternatively, if context diversity (and not word-frequency) is the driving force in lexical access (as claimed by Adelman and others), then words which occur in many contexts in print –while keeping word-frequency controlled – should produce faster word-identification times than words which occur in few contexts. If this latter hypothesis is confirmed, a new conceptualization of models of visual-word recognition would be required. This applies not only to the static activation-based models cited above. Connectionist models (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996) are dynamic models that include a learning rule (i.e., word learning is possible). However, as noted by Baayen (2010), this learning rule relies on repeated presentations of the words rather than on contextual diversity (i.e., the PDP model cannot explain the contextual diversity effects reported by Adelman et al., 2006).

Contextual diversity and word-frequency were obtained from the ESCOLEX database (Soares et al., 2012). ESCOLEX is a children's lexical database with grade-level word frequency statistics for 6 to 11 year-old children (1st to 6th Grade) for a total of approximately 48,500 wordforms computed from a 3,2 million word corpus taken from 171 European Portuguese elementary textbooks. Similar to Adelman et al. (2006),

Contextual Diversity was computed as the number of textbooks in which a given word appeared and the word-frequency as the number of occurrences of a given word in the corpus (per million words). In the present study, these indexes were collected from 1st through 4th Grade computations –the participants were 4th Grade children. All 4th Grade children were familiar with all the words in print employed in the experiment, since all these words appeared in textbooks of the initial year of primary school.

In the experiment, we employed the most popular laboratory word identification task: lexical decision ("is the stimulus a word?"; e.g., see Dufau et al., 2011). We opted for the go/no-go procedure (i.e., press a key for "word" and refrain from responding for "nonword") because it is the preferable procedure in experiments with children (i.e., it produces faster lexical decision times and fewer errors than the yes/no procedure; Moret-Tatay & Perea, 2011; see Perea, Moret-Tatay, & Panadero, 2011; Perea, Panadero, Moret-Tatay, & Gómez, 2012, for evidence with young readers; see also Gómez, Ratcliff, & Perea, 2007, for a mathematical model of the go/no-go task).

Experiment

Method

<u>Participants</u>. Twenty-two 4th grade children (mean age: 9.94 years; range: 9.53-10.52; 11 female) participated voluntarily in the experiment –written consent form had been obtained from their parents. The children came from above-average socioeconomic backgrounds in a private school in Porto, Portugal. All participants had normal (or corrected-to-normal) vision and were native speakers of European Portuguese. None of them had any sensory, neurological, or learning disabilities. The experiment took place at the end of the academic year.

Materials. A set of 60 Portuguese words was selected from the ESCOLEX database (Soares et al., 2012). All these words were either nouns or adjectives. The two factors intended in the experiment were contextual diversity (low vs. high) and lexical frequency (low vs. high). (For simplicity, we employ the term low-frequency words, but these words would be better qualified as "medium" frequency words.) Given that it was not possible to obtain enough stimuli in the condition of high contextual diversity and low-frequency, this condition was discarded (see Table 1 for the values [averages and ranges] in each condition). For each of the three remaining conditions (i.e., high contextual diversity - high frequency; low contextual diversity - high-frequency; low contextual diversity - low frequency), there were 20 words each which were well matched in imageability, concreteness, familiarity, number of letters, number of syllables, and number of orthographic neighbors in the ESCOLEX database (see Table 1) (all ps > .17). A set of 60 orthographically legal pseudowords –matched in word length with the word stimuli– was created for the purposes of the lexical decision task. The list of words and pseudowords is presented in the Appendix.

Procedure. The experiment took place in groups of four children in a quiet room. DMDX software (Forster & Forster, 2003) was used to present the stimuli and collect the responses. The scheme of a given trial was the following: A fixation point (+) was presented for 500 ms in the center of the screen. Then, the stimulus (either a word or a nonword) was presented in lowercase 14-pt Times New Roman font. The word/nonword remained on the screen until the participant responded –or 2500 ms had elapsed. The inter-trial interval was 1 sec. Participants were told that real words and "nonsense" words would be displayed on the computer screen and that they should press one button (labeled ''sim'' [yes]) if the stimulus was a real word and should refrain from responding if the stimulus was not a real word. Participants were instructed

to respond as fast as possible while trying not to make too many errors. Each participant received a different random order of stimuli. The session lasted approximately 10 min.

Results

Incorrect responses and response times less than 250 ms or greater than 1500 ms (3.1% of data) were excluded from the latency analysis. The mean response times for correct responses and the error rates per each condition are presented in Table 2. Error rates for word stimuli were very low (less than 0.5%) and were not further analyzed.

Because of the characteristics of the experimental design (i.e., absence of the high contextual diversity - low-frequency condition), we conducted planned comparisons on the lexical decision data rather than an unfocused omnibus analysis in a (see Wilcox, 1987). First, to examine the effect of contextual diversity, we compared the condition with low contextual diversity - high-frequency vs. the condition with high contextual diversity - high-frequency. Second, to examine the effect of word-frequency, we compared the condition with low contextual diversity - high-frequency vs. the condition with low contextual diversity - low-frequency. These planned comparisons were based on the participant ($\underline{t}1$) and item ($\underline{t}2$) mean correct response latencies.

Effect of contextual diversity: On average, words with a high contextual diversity were responded to 53 ms more rapidly than words with a low contextual diversity, $\underline{t}1(21) = 3.37$, $\underline{SE}_{dif} = 15.6$, $\eta^2 = .35$, $\underline{p} = .003$; $\underline{t}2(38) = 2.89$, $\underline{SE}_{dif} = 18.0$, $\eta^2 = .18$, $\underline{p} = .006$.

Effect of word-frequency: The difference between the identification times of low- and high-frequency words was less than 2 ms, both $|\underline{t}|s < 1$.

Similarly to Adelman et al. (2006), and to further examine the impact of contextual diversity and word-frequency on lexical decision times in developing readers, we conducted a regression analysis in which response time was the dependent variable, and $\log_{10}$ of contextual diversity, $\log_{10}$ of word frequency (both from ESCOLEX), number of letters, and imageability were used as predictions. The regression analysis indicated a facilitative effect of $\log_{10}$ of contextual diversity ($\beta = -.41$, $\underline{t} = -3.15$, $\underline{p} = .003$) and imageability ($\beta = -.31$, $\underline{t} = -2.67$, $\underline{p} = .01$), but not of $\log_{10}$ of word-frequency ($\beta = -.04$, $|\underline{t}| < 1$, $\underline{p} > .50$), while the number of letters had a (nonsignificant) inhibitory trend ($\beta = .17$, $\underline{t} = 1.47$, $\underline{p} = .15$). Thus, the regression analyses confirmed the importance of contextual diversity in visual-word recognition with young readers.

Follow-up study

Given the implications of the present findings for developmental, dynamic models of visual-word recognition, it is important to re-examine the effects of contextual diversity and word frequency using another set of words and another sample of young readers. Since both word-frequency and contextual diversity are continuous variables, we selected 60 words widely differing in these factors and conducted a regression analysis parallel to that conducted above. We employed a go/no-go lexical decision task.

Method

Participants. The participants were twenty-seven 4th graders from the same school as in the reported experiment –written consent form had been obtained from their parents. They had not participated in the experiment.

Materials. We selected a set of 60 European Portuguese words (nouns and adjectives) varying in word frequency (M=215 per million; range: 43-934) and contextual diversity

(M=0.48; range: .10-.85 in the ESCOLEX database). We also created a set of 60 orthographically legal pseudowords for the purposes of the lexical decision task. The list of items is provided in the Appendix.

Procedure. It was the same as in the experiment.

Results

Error responses (1.4 %) and response times less than 250 ms or greater than 1500 ms (4.3 %) were excluded from the RT analysis. The predictors of the latency data were the same as in the regression analysis reported above. Results again indicated a facilitative effect of $\log_{10}$ of contextual diversity (β = -.37, t = -2.96, p = .005) and imageability (β = -.30, t = -2.61, p = .01), but not of $\log_{10}$ of word-frequency (β = -.15, t = -1.24, p = .22). As in the experiment, the effect of number of letters was slightly inhibitory (β = .19, t = 1.87, p = .067).

General Discussion

The present study, using two sets of words, has revealed that contextual diversity is a main determinant of word identification times in young readers, thus generalizing the data from Adelman et al. (2006) with college-aged participants to a children population –and a different language (Portuguese). Furthermore, and similarly to the Adelman et al. analyses, word-frequency did not determine word-identification times, at least in the range of frequencies (i.e., medium-high frequency) employed in the present study.

The present data pose obvious problems for static models of visual-word recognition that assume that "word-frequency" per se plays a key role –in particular when applying these models to a children population. Future implementations of

computational models of visual-word recognition should go beyond static word representations and include a learning rule. This learning rule should not just work in the basis of mere count of instances in which a word appears because the more contexts a word is presented, the faster its identification –all other things being equal. Of particular interest here is the Semantic Distinctiveness (SD) model proposed recently by Jones, Johns, and Recchia (2012). In the context of word acquisition, the repetition of a word in the SD model only produces a detectable modification in a word's memory strength when it is accompanied by a change in context. This context can be regarded at a global level (as proposed by Adelman et al., 2006) or at a more local level (on the basis of similarity-based models of word co-occurrence; McDonald & Shillcock, 2001; Hoffman, Rogers, & Lambon Ralph, 2011; see also Landauer & Dumais, 1987). The present study was not designed to disentangle these options. Further research with developing readers should help to separate the importance of these two approaches.

The present developmental data may have implications for models of memory, in particular, on the area of literacy and word acquisition –both in the context of learning the first language (L1) and the second language (L2). The present findings –in conjunction with previous research (e.g., Hills et al., 2010; Jones et al., 2012; Steyvers & Malmberg, 2003) reveal that, in order to facilitate literacy skills, words should be learned in different contexts. Thus, from an educational point of view, and in the context of word learning (e.g., in a second language), the present data suggest that teachers should provide learning opportunities so that the same word can be experienced in different contexts of occurrence rather than confronting them with repeated presentations of words (e.g., nouns, adjectives) in very specific (or similar) contexts.

In sum, we have demonstrated, across two samples of words, that the effect of contextual diversity in visual-word recognition is robust in developing readers. This

strongly suggests that a dynamic, contextual-dependent approach to word acquisition should be included in future implementations of models of visual-word recognition. At an applied level, the present data can also have practical implications on the way new words (in L1 or L2) are initially learned and accessed.

References

Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology, 99*, 245-264.

Adelman, J. S., Brown, G.D.A., & Quesada, J.F. (2006). Contextual diversity, not word frequency, determines word naming and lexical decision times. *Psychological Science, 17*, 814–823.

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B.,…Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445-459.

Baayen, R.H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon, 5*, 436-461.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). The DRC model: A model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-258.

Davis, C.J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117*, 713-758.

Dufau, S., Duñabeitia, J.A., Moret-Tatay, C., McGonigal, A., Peeters, D., Alario, F.–X. … Grainger, J. (2011). Smart phone, smart science: how the use of smartphones can revolutionize research in cognitive science. *PLoS ONE, 6*(9), e24974. doi:10.1371/journal.pone.0024974.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl. R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*. 777-813.

Forster, K.I., & Forster, J.C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124.

Gómez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136*, 389-413.

Grainger, J. & Jacobs, A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*, 518-565.

Hills, T., Maouene, J., Riordan, B., & Smith, L. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language, 63*, 259-273.

Hoffman, P., Rogers, T.T., & Lambon Ralph, M.A. (2011). Semantic diversity accounts for the "missing" word frequency effect in stroke aphasia: Insights using a novel method to quantify contextual variability in meaning. *Journal of Cognitive Neuroscience, 23*, 2432-2446.

Hyönä, Y., & Olson, R.K. (1995). Eye movement patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1430–1440.

Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40*, 431–439.

Jones, M.N., Johns, B.T., & Recchia, G.L. (2012). The role of semantic diversity in
lexical organization. *Canadian Journal of Experimental Psychology, 66*, 121-132.

Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: the Latent
Semantic Analysis theory of acquisition, induction and representation of
knowledge. *Psychological Review, 104*, 211-240.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context
effects in letter perception: Part 1. An account of Basic Findings. *Psychological
Review, 88,* 375-407.

McDonald, S.A., & Shillcock, R.C. (2001). Rethinking the word frequency effect: The
neglected role of distributional information in lexical processing. *Language and
Speech, 44*(3), 295-323.

Moret-Tatay, C., & Perea, M. (2011). Is the go/no-go lexical decision task preferable to
the yes/no task with developing readers? *Journal of Experimental Child
Psychology, 110*, 125-132. DOI: 10.1016/j.jecp.2011.04.005

Perea, M., Moret-Tatay, C., & Panadero, V. (2011). Suppression of mirror
generalization for reversible letters: Evidence from masked priming. *Journal of
Memory and Language, 65*, 237–246. DOI: 10.1016/j.jml.2011.04.005

Perea, M., Panadero, V., Moret-Tatay, C., & Gómez, P. (2012). The effects of inter-
letter spacing in visual-word recognition: Evidence with young normal readers
and developmental dyslexics. *Learning and Instruction, 22*, 420-430. DOI:
10.1016/j.learninstruc.2012.04.001

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996).
Understanding normal and impaired word reading: Computational principles in
quasi-regular domains. *Psychological Review, 103*, 56-115.

Ratcliff, R., Love, J., Thompson, C.A., & Opfer, J. (2012). Children are not like older
adults: A diffusion model analysis of developmental changes in speeded responses.
Child Development, 83, 367-381.

Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye
movement control in reading. *Psychological Review, 105*, 125–157.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of
word recognition and naming. *Psychological Review, 96*, 523-568.

Soares, A.P., Medeiros, J.C., Simões, A., Machado, J., Costa, A.,…Comesaña, M.
(2012, March). ESCOLEX: Uma base lexical com medidas de frequência para
crianças do 1º ao 6º ano de escolaridade [ESCOLEX: A lexical database with
frequency measures for Portuguese children from 1st to 6th grade]. Poster
presented at the *VII Encontro da Associação Portuguesa de Psicologia
Experimental*, Lisbon, Portugal.

Steyvers, M., & Malmberg, K. (2003). The effect of normative context variability on
recognition memory. *Journal of Experimental Psychology: Learning, Memory,
and Cognition, 29*, 760-766.

Wang, H.-C., Castles, A., Nickels, L., & Nation, K. (2011). Context effects on
orthographic learning of regular and irregular words. *Journal of Experimental
Child Psychology, 109*, 39-57.

Wilcox, R.R. (1987). New designs in analysis of variance. *Annual Review of Psychology, 38*, 29-60.

Table 1. <u>Characteristics of the words employed in the experiment. The ranges appear between brackets.</u>

	IMG	FAM	CONC	# Letters	# Syllables	DC	FRQ_MIL	N
High DC, High-Freq	4.55	6.59	4.74	5.65	2.55	0.65	202.85	2.90
	(2.3-5.8)	(6.3-6.8)	(2.6-6.6)	(4-8)	(2-3)	(0.6-0.8)	(154-289)	(0-8)
Low DC, High-Freq	4.51	6.50	4.90	5.85	2.55	0.33	202.04	1.45
	(3.0-5.8)	(6.2-6-8)	(3.6-6.6)	(5-7)	(2-3)	(0.2-0.4)	(112-402)	(0-11)
Low DC, Low-Freq	4.63	6.56	4.95	5.85	2.75	0.32	53.22	2.00
	(3.2-5.7)	(6.2-6.9)	(3.4-6.2)	(5-7)	(2-3)	(0.2-0.5)	(25.5-69.4)	(0-9)

Note: IMG=Imageability (from 1 to 7), FAM=Subjective Familiarity (from 1 to 7), CONC=Concreteness (from 1 to 7), # Letters= Average number of letters, # Syllables = Average number of syllables, DC=proportion of textbooks that include the target word in the ESCOLEX database, FRQ_MIL=frequency per million in the ESCOLEX database, and N=number of orthographic neighbors in the ESCOLEX database (i.e., Coltheart's N).

Table 2. Mean lexical decision times (in ms) and percentage of errors (in parentheses) for words in the experiment

Word frequency	Contextual Diversity	
	Low	High
Low	808 (0.2)	
High	807 (0.9)	754 (0.2)

Note: Error rates for pseudowords were 6.7%.

Appendix

Experiment:

List of words

High contextual diversity - high-frequency: balão; longo; passeio; certo; altura; música; igual; barco; direita; frente; campo; viagem; dinheiro; bonito; modo; animal; leite; cabeça; falta; final

Low contextual diversity - high-frequency: ficha; plural; bloco; poema; metro; dezena; sujeito; divisão; litro; raposa; vírgula; nuvem; autor; manual; macaco; sonho; cálculo; carta; segredo; produto

Low contextual diversity - low-frequency: código; corrida; almoço; escuro; parede; ouvido; pedaço; modelo; tecido; braço; mágico; moinho; doente; pinhal; desejo; minuto; regra; dente; curto; buraco

List of pseudowords

anve; riafe; lasbo; gisbo; nasma; tenfo; vilno; tapra; grica; malpa; borvo; danba; rilba; tarlo; vesca; racla; panto; grica; cuefo; ludia; etura; dusle; gasva; upesta; mislos; nafeta; dunibo; balver; dalpar; cormas; cetupa; anerbo; coltor; culfos; levrol; nosfeu; untipa; triego; muelfo; dipada; adisgo; guinle; denvas; canbra; iselas; verfes; lirato; libaza; forvas; dunfol; dorelta; moltroa; ristura; tengala; corlufa; pábiras; esdutos; rocerço; norjada; canfresa

Follow-up study

List of words: fruto; irmão; festa; melhor; campo; leite; autor; tarde; balão; lápis; planta; porta; manhã; rapaz; praia; mensal; cento; litro; quadra; milhar; painel; régua; metro; plural; banda; verbo; costa; verso; triste; ficha; carro; longo; falta; jornal; certo; regra; conta; cheio; barco; banho; velho; pinta; turma; frente; quinta; pastor; prosa; museu; corvo; local; grilo; grave; pinhal; bloco; pobre; barro; braço; monte; gente; guarda

List of pseudowords: demor; garlu; traquo; conci; tenoi; lanhi; lisna; tonpo; fotru; quarro; conbra; varla; renhos; tinve; riafe; balver; trome; danba; borvo; pazca; merno; falni; mumco; tenfo; tantri; lasbo; zarão; rilba; donli; uvnem; ; panti; tocur; fotei; rizna; dajau; janla; nasma; molti; tuala; gavem; dinceu; galmi; tosas; deida; neira; ; cranba; volme; nector; guaia; tapra; malpa; dalpar; peton; marpo; panto; topre; sotal; reifa; mopri; tonsa