MedCrave
Step into the World of Research

Journal of Neurology & Stroke

Review Article

Speech-to-spelling encoding of Brazilian Portuguese: predicting misspellings to the dictation of rare words

Volume 16 Issue 2 - 2026

Fernando César Capovilla
Experimental Psychology Department, University of São Paulo, Brazil

Correspondence: Fernando César Capovilla, Full Professor. Experimental Psychology Department, Institute of Psychology, University of São Paulo, USP, São Paulo, SP, Brazil

Received: March 18, 2026 | **Published:** April 10, 2026

Abstract

Predicting misspellings of rare words in a given orthography involves understanding that spelling errors are not random, but rather follow patterns related to the phoneme-grapheme structure of that orthography. The paper advances a new theoretical and experimental model of the encodability of Brazilian Portuguese that permits the estimation of the risk of spelling errors. It measures the degree of spelling difficulty for any given spoken word during auditory dictation tasks. The model uses Portuguese Phoneme-Grapheme Encoding Indexes derived from the analysis of 4.55M Phoneme-Grapheme Links. According to the encoding model, in the task of spelling spoken words under auditory dictation, the degree of spelling difficulty of any given word corresponds to the arithmetic mean of the Encoding Indexes of the Phoneme-Grapheme links involved. A study was conducted to assess whether the model could account for encoding precision. In the study, 154 students (61 from college and 93 from elementary school) spelled under dictation a *corpus* of 560 rare spoken words. The *corpus* comprised combinations of up to 280 Phoneme-Grapheme Links, and, to maximize phonological encoding, the 560 spoken words to be spelled under dictation had low orthographic familiarity. Each student was required to encode on average 3,676 Phoneme-Grapheme Link instances. Regression analysis revealed that word spelling precision increased with the degree of word encodability. Results suggest the encodability model predicts word-spelling precision in Brazilian Portuguese.

Keywords: Portuguese, misspelling, encoding, phoneme, grapheme

Introduction

There are basically two approaches to reading and spelling: the single-route approach[1] and the dual-route approach.[2] The dual-route cascaded model[2–5] proposes two routes: a lexical route and a phonological or sublexical route. The lexical route is based on an interactive activation procedure[6,7] that activates a phonological code in the mental lexicon from a visual word code. In contrast, the phonological or sublexical route is based on grapheme-phoneme correspondence rules: each grapheme is converted into a corresponding phoneme by the application of a set of rules.

The dual-route approach suggests that accessing the meaning of print may use either the lexical[8,9] or the phonological route.[10–17] The *lexical route* is based on the direct visual recognition of the orthographic form of familiar words. It permits reading words that have exceptional or irregular phoneme-grapheme correspondences, provided those words have familiar orthographic forms. The *phonological route* is traditionally said to operate on grapheme-phoneme conversion rules that permit the construction of word pronunciation. It permits reading both non-words and novel words, that is, words whose orthographic form is unfamiliar (i.e., whose representation has not been stored in the orthographic lexicon), provided they are made of regular phoneme-grapheme correspondences. It permits understanding the meaning of novel words (i.e., having access to the semantic lexicon), provided their decoding yields phonological forms familiar to the reader (i.e., whose representations have been stored in the phonological lexicon), so that their utterance sounds familiar to the reader. If the component phoneme-grapheme correspondences are irregular, their decoding produces unfamiliar phonological forms, which end up not being recognized by the reader. Consequently, in that case, the reader is unable to understand the meaning of such irregularly spelled words.

Given that, in the phonological route, access to meaning is necessarily mediated by speech, reading comprehension only occurs when three conditions are met: 1. Words to be read are made of grapheme strings the decoding of which is so regular (i.e., canonical or nonexceptional) that it produces precisely the phoneme sequence of the corresponding spoken word form; 2. Readers have appropriate decoding skills to decode in a precise, fluent, and effortless manner, thus reading aloud words almost as fluently and naturally as if they were naming the respective objects that the words correspond to; and 3. Readers have a sufficiently developed phonological lexicon to recognize their own utterances as corresponding to familiar words.

In the dual-route approach, the writing system is conceived as a reversible code system that permits encoding speech into print for spelling under dictation and decoding print into speech for reading aloud. Phonological processing is the most significant cognitive process underlying the development of reading and spelling skills. The phonological recoding involved in such a process is said to be based on conversion rules. In *reading aloud*, the phonological process consists of converting print into speech by means of the application of *grapheme-to-phoneme conversion rules* (i.e., *pronunciation rules*). Conversely, the phonological process involved in *spelling under dictation* consists of converting speech into print by applying *phoneme-to-grapheme conversion rules* (i.e., *spelling rules*).

To shed light on the nature of the phonological processes underlying the phonological route, the present paper proposes a Recoding Model based on recoding intuition (implicit knowledge) rather than on recoding rules (explicit knowledge). Such intuitions are formed by temporal and spatial contiguities and contingencies based on the mere statistical distributions of bi-directional associations between Speech Units and Print Units at the sub-lexical level (e.g., Grapheme, Digraph, Syllable), as well as at the lexical level (i.e., Word). Speech Units (SU) are arranged in temporal contiguity in speech, and may occur at the phoneme level or above it (e.g., spoken syllables, spoken morphemes, spoken words). Print Units (PU) are arranged in spatial contiguity in print, and may occur at the grapheme level or above it (e.g., written syllables, written morphemes, written words).

According to the model, in incidental experiences, SU and PU may occur in association with temporal and spatial contiguity. In systematic training, their co-occurrence may be established, according to if-then contingencies (i.e., if SU then PU). Such an arrangement tends to constitute, by induction, a gradient of SU-PU associations, which is experienced as encoding intuitions and expectancies regarding the way SU may be usually encoded when spelled under dictation. Such a systematic arrangement of if SU then PU contingencies during spelling acquisition training is constituted by unconscious induction, implicit know-how competence that permits encoding SU sequences in speech into PU strings in print. Encoding rules would be limited to grammar position rules that condition the encoding of a SU into a PU to the position that the SU occupies in the spoken word (i.e., word beginning, word end, in between vowels, preceding a voice or unvoiced consonant). For instance in Portuguese, phoneme [k] is spelled as “c” in 90,00% of the cases (e.g., [ka’loɾ] -”calor”), as “qu” in 7,83% of the cases (e.g., [ki’taɾ] -”quitar”), as “q” in 1,83% of the cases (e.g., [delĩ’kwẽʃsjɐ] -”delinquência”), as “k” in 0,216% of the cases (e.g., [kiu’i] -” kiwi”), as “ck” in 0,0827% of the cases (e.g., [lamaɾ’kistɐ] -”lamarckista”), as “ch” in 0,314% of the cases (e.g., [‘kɾomɐ] -”chroma”), and as “cqu” in 0,0029% of the cases (e.g., [e’kɛɾɪɐ] -”hecquéria”). Thus, upon listening to the phoneme [k], the expectancy of spelling it as “c” is 11,5 times greater than that of spelling it as “qu”. However, a position rule permits reducing dependency on mere chance by establishing that “the phoneme [k] is usually spelled as ‘c’ when it precedes ‘a’, ‘o’ and ‘u’ (e.g., ‘casta’, ‘costa’, ‘custa’) vowels, and as ‘qu’ when it precedes ‘e’ and ‘i’ vowels (e.g., ‘querido’, ‘quimera’).” Drill and practice training develops in a bottom-up fashion the intuitive and unconscious knowhow performance that is primarily basic to spelling competence, whereas grammar rule instruction forms, in a top-down fashion, the formal know-why explanation and decision-making patterns that add, secondarily, great efficiency to spelling competence. The present paper argues that, underlying the so called encoding rules are intuitions formed by probabilistic distributions of SU-PU associations, and that the only true rules underlying phonological encoding are grammar rules (i.e., position rules) pertaining to the cases where the PU to be produced depends on the position that the SU occupies in the spoken word to be spelled under auditory dictation by SU-PU encoding.

By the same token, according to the model, in incidental experiences, PU and SU may occur associated by temporal and spatial contiguity. In systematic training, their co-occurrence may be established according to if-then contingencies (i.e., if PU then SU). Such an arrangement tends to constitute, by induction, a gradient of PU-SU associations, which is experienced as decoding intuitions and expectancies regarding the way PU may be usually uttered when read aloud. Such a systematic arrangement of if PU then SU contingencies during reading acquisition training constitutes implicit *know-how* competence that permits decoding PU strings in print into SU sequences in speech. Decoding rules would be limited to grammar position rules that condition the decoding of a PU into a SU to the position that the PU occupies in the written word (i.e., word beginning, word end, in between vowels, preceding a voice or unvoiced consonant). For instance, in Portuguese, the grapheme “s” sounds like [s] 69% of the time and [z] 31% of the time. Thus, upon seeing the grapheme “s”, the expectancy of sounding it as [s] is 2.2 times greater than that of sounding it as [z]. However, the position rule permits not depending on mere chance and establishes that “the grapheme ‘s’ sounds as [z] when it occurs in between two vowels.”). Drill and practice training develops in a bottom-up fashion the intuitive and unconscious knowhow performance that is primarily basic to reading competence, whereas grammar rule instruction forms, in a top-down fashion, the formal know-why explanation and decision-making patterns that add, secondarily, great efficiency to reading competence. The present paper argues that, underlying the so-called decoding rules, are intuitions formed by probabilistic distributions of PU-SU associations, and that the only true rules underlying phonological decoding are grammar rules (position rules) pertaining to the cases where the SU to be produced depends on the position that the PU occupies in the written word to be read aloud by PU-SU decoding.

Conversion rules tend to be involved in phonological encoding and decoding only when pronunciation and spelling are position-dependent. Yet, even in those cases, pronunciation and spelling may be purely intuitive, based on systematic drill and practice exercises centered on PU-SU and SU-PU associations. Even though position rules pertaining to pronunciation could describe behavior, those rules may never have been taught or even deduced by the student. It is possible that decoding and encoding intuitions are primary and fundamental to reading and spelling, whereas grammar position rules are secondary and complementary to them. Even though conversion rules can provide *post facto* support for intuition, they cannot substitute for it. The present paper proposes that recoding intuition derives from the statistical distributions of bimodal units (auditory and visual) in bidirectional links.

Both reader intuition and writer intuition derive from experience:

(1) The statistical distribution of the relative incidence of all SUs that decode a given PU forms a reading intuition. When that PU occurs, the probability of uttering one or another of all the SUs is a positive function of the SU statistical prevalence in that distribution.

(2) The statistical distribution of the relative incidence of all PUs that encode a given SU forms a writing intuition. When that SU occurs, the probability of encoding one or another of all PU is a positive function of the PU statistical prevalence in that distribution.

PU-SU links are stored in a bimodal (visual-auditory) PU-SU sub-lexicon. Print triggers speech. The Stroop effect (Assef, Capovilla, & Capovilla, 2007)[18] results from that encapsulated triggering. Literacy training creates the bimodal Print-Voice sub-lexicon that affords reading, and the bimodal Voice-Print lexicon that affords spelling.

The degree of decoding development and performance derives directly from the statistical distribution of PU-SU links. The degree of encoding development and performance derives directly from the statistical distribution of SU-PU links. The computerized mapping of those distributions[12,14,19–22] permits the prediction of both reading and spelling error distributions.

Natural environments produce overwhelming variation in experience. Experimental and statistical control (via longitudinal and cohort studies) upon children’s reading and spelling ontological histories in natural environments is relatively limited. In contrast, mapping the distribution of PU-SU links and the distribution of SU-PU links in one’s mother language is much more feasible and potentially effective. The present Recoding (Decoding-Encoding) Model uses such a mapping.

Recently, a series of papers have proposed a theoretical and experimental model for measuring Portuguese encoding, and estimating the degree of difficulty involved in spelling under auditory dictation any given spoken word in Portuguese.[8,10–14,19,20–26] The authors implemented that model in a software named *Brazilian Voice in the New Orthography*. The software permits transcribing, using International Phonetic Alphabet (IPA) characters, all regional pronunciations of any given Portuguese word. Using the software, the authors transcribed in IPA characters more than 60,000 words in more than 315,000 different pronunciations. Computerized analysis of the statistical distribution of combinations between SU (in IPA characters) and PU provided a basis for a seminal mapping of the bi-directional links between Phonology and Orthography in Portuguese.

The data *corpus* arising from such a mapping allows for explaining how the phonological route can decode unfamiliar words in reading-aloud tasks and encode them in spelling under auditory dictation tasks.

The Orthography-Phonology mapping involved two phases:

(1) Identifying all possible SUs that may utter any given PU in reading aloud tasks;

(2) Computing the relative percentage of occurrence with which each alternative SU utters any given PU in a representative *corpus* of the Portuguese lexicon. Such a relative percentage corresponds to the Portuguese PU-SU Decoding (GPD) Index, which measures the decoding difficulty of all PUs that comprise Portuguese Orthography.

The present model permits calculation of the Word Decoding Degree (WDD) for any given written word to be read aloud. WDD corresponds to the average mean of the GPD indexes involved in pronouncing the written word in a reading-aloud task. This involves two steps:(1) Summing up the Decoding (GPD) Indexes of all PU-SU links involved in pronouncing the written word; (2) Dividing that sum by the number of GPD indexes (each for a different PU-SU link) involved in pronouncing that written word.

The Phonology-Orthography mapping also involved two phases:

(1) Identifying all possible PUs that may spell (i.e., encode) any given SU while writing in a spelling under dictation task;

(2) Computing the relative percentage of occurrence with which each alternative PU encodes any given SU in a representative *corpus* of the Portuguese lexicon. Such a relative percentage corresponds to the Portuguese SU-PU Encoding Index, which measures the degree of encoding difficulty for all SUs that comprise the Portuguese Phonology.

The present model permits calculating the Word Encoding Degree (WED) of any given spoken word for spelling under dictation. By the way, such a spelling under dictation occurs even in spontaneous writing. In this case, the inner ear responds (via the audibilization process, cf. Baddeley, Gathercole, & Papagno, 1998) to one's own inner speech. WED corresponds to the average mean of the SU-PU Encoding indexes involved in spelling the spoken word while writing it down under dictation. This involves two steps:(1) summing up the SU-PU Encoding indexes of all SU-PU links involved in spelling the spoken word; (2) dividing that sum by the number of SU-PU Encoding indexes (each for a different SU-PU link) involved in spelling that spoken word.

The software *Brazilian Voice* provides 367 SU-PU Encoding Indexes and their corresponding 367 PU-SU Decoding Indexes. Tables 1 to 3 summarize the SU-PU Encoding Indexes involved in the word list used in the present study. SU-PU Encoding Indexes permit calculating by hand the degree of difficulty of spelling under dictation any given spoken word in Portuguese. PU-SU Decoding Indexes permit calculating by hand the degree of difficulty of reading aloud any given printed word in Portuguese.

Table 1 Voice-Spelling Links (VS-Link) that make up the 560 words and their respective Voice-Spelling Encoding Indexes (VSEI). Part 1: Links 1-90

N	**VS-Link**	**VSEI**	**N**	**VS-Link**	**VSEI**	**N**	**VS-Link**	**VSEI**
1	[\a\]-"a"	90.613	31	[\b\]-"b"	100.000	61	[\ew\]-"eu"	20.225
2	[\a\]-"á"	9.038	32	[\ɔ\]-"ó"	66.945	62	[\ew\]-"êu"	2.313
3	[\a\]-"ha"	0.333	33	[\ɔ\]-"o"	32.604	63	[\ew\]-"hel"	0.715
4	[\a\]-"à"	0.001	34	[\ɔj\]-"oi"	50.122	64	[\ɛ\]-"e"	53.619
5	[\ɐ\]-"a"	97.651	35	[\ɔj\]-"ói"	49.878	65	[\ɛ\]-"é"	40.930
6	[\ɐ\]-"â"	2.314	36	[\ɔw\]-"ol"	88.800	66	[\ɛ\]-"hé"	0.458
7	[\ɐ\]-"u"	0.015	37	[\ɔw\]-"ól"	9.200	67	[\ɛ\]-"a"	0.056
8	[\ɐ\]-"hâ"	0.008	38	[\d\]-"d"	100.000	68	[\ɛ\]-"ê"	0.012
9	[\ɐ̃\]-"an"	80.331	39	[\dʲ\]-"d"	100.000	69	[\ɛj\]-"ei"	81.986
10	[\ɐ̃\]-"am"	10.814	40	[\dʒ\]-"d"	99.875	70	[\ɛj\]-"éi"	18.014
11	[\ɐ̃\]-"ân"	6.839	41	[\dʒ\]-"j"	0.028	71	[\ɛw\]-"el"	63.930
12	[\ɐ̃\]-"ã"	1.420	42	[\e\]-"e"	97.259	72	[\ɛw\]-"éu"	22.388
13	[\ɐ̃\]-"âm"	0.524	43	[\e\]-"ê"	1.386	73	[\ɛw\]-"éo"	5.970
14	[\ɐ̃\]-"un"	0.015	44	[\e\]-"he"	1.343	74	[\ɛw\]-"él"	3.980
15	[\ɐ̃\]-"ham"	0.015	45	[\e\]-"é"	0.009	75	[\ɛw\]-"eo"	1.493
16	[\ɐ̃\]-"in"	0.005	46	[\e\]-"er"	0.001	76	[\ɛw\]-"eu"	0.746
17	[\aj\]-"ai"	98.285	47	[\ẽ]-"en"	85.127	77	[\ɛw\]-"hél"	0.249
18	[\aj\]-"i"	0.591	48	[\ẽ]-"em"	7.871	78	[\f\]-"f"	99.992
19	[\aj\]-"ái"	0.355	49	[\ẽ]-"ên"	6.833	79	[\f\]-"ph"	0.008
20	[\aj\]-"y"	0.355	50	[\ẽ]-"hen"	0.100	80	[\g\]-"g"	94.396
21	[\ɐj\]-"ai"	100.000	51	[\ẽ]-"êm"	0.070	81	[\g\]-"gu"	5.555
22	[\ɐ̃j̃]-"ãe"	50.000	52	[\ej\]-"ei"	99.579	82	[\gz\]-"x"	100.000

Citation: Capovilla FC. Speech-to-spelling encoding of Brazilian Portuguese: predicting misspellings to the dictation of rare words. *J Neurol Stroke*. 2026;16(2):78–89. DOI: 10.15406/jnsk.2026.16.00656

Table 1 Continued...

23	[\ɐ̃j̃]-"ãi"	5.114	53	[\ej\]-"êi"	0.211	83	[\h\]-"r"	80.550
24	[\aw\]-"al"	67.854	54	[\ej\]-"a"	0.164	84	[\h\]-"rr"	19.033
25	[\aw\]-"au"	26.783	55	[\ej\]-"em"	94.103	85	[\h\]-"h"	0.417
26	[\aw\]-"ál"	2.432	56	[\ej\]-"en"	5.897	86	[\i\]-"i"	89.300
27	[\aw\]-"áu"	1.973	57	[\ej̃]-"ém"	83.133	87	[\i\]-"í"	10.624
28	[\aw\]-"ao"	0.345	58	[\ej̃]-"én"	4.819	88	[\i\]-"y"	0.037
29	[\aw\]-"hau"	0.287	59	[\ew\]-"eo"	38.289	89	[\i\]-"hi"	0.015
30	[\ɐ̃w̃\]-"ão"	99.986	60	[\ew\]-"el"	38.107	90	[\i\]-"hí"	0.013

Table 2 Voice-Spelling Links (VS-Link) that make up the 560 words and their respective Voice-Spelling Encoding Indexes (VSEI). Part 2: Voice-Spelling Links 91-180

N	VS-Link	VSEI	N	VS-Link	VSEI	N	VS-Link	VSEI
91	[\i\]-"ie"	0.002	121	[\jo\]-"iô"	0.134	151	[\õj\]-"õe"	100.000
92	[\i\]-"hy"	0.001	122	[\ju\]-"io"	71.184	152	[\ow\]-"ou"	61.000
93	[\ɪ\]-"e"	68.546	123	[\ju\]-"eo"	27.221	153	[\ow\]-"ol"	38.520
94	[\ɪ\]-"i"	28.556	124	[\ju\]-"iu"	1.345	154	[\ow\]-"ôl"	0.120
95	[\ĩ\]-"in"	48.160	125	[\ju\]-"iú"	0.219	155	[\p\]-"p"	100.000
96	[\ĩ\]-"en"	30.297	126	[\ju\]-"yu"	0.010	156	[\ř\]-"r"	57.743
97	[\ĩ\]-"im"	10.025	127	[\k\]-"c"	89.997	157	[\ř\]-"rr"	42.257
98	[\ĩ\]-"em"	9.166	128	[\k\]-"qu"	7.842	158	[\ɹ\]-"r"	100.000
99	[\ĩ\]-"ín"	2.068	129	[\k\]-"q"	1.829	159	[\ɾ\]-"r"	100.000
100	[\ĩ\] -"ím"	0.163	130	[\k\]-"k"	0.216	160	[\s\]-"s"	54.052
101	[\iw\]-"il"	88.337	131	[\k\]-"ck"	0.083	161	[\s\]-"c"	27.430
102	[\iw\]-"íl"	4.968	132	[\k\]-"ch"	0.030	162	[\s\]-"ç"	9.308
103	[\iw\]-"iu"	1.728	133	[\k\]-"cqu"	0.003	163	[\s\]-"ss"	5.372
104	[\ɪw\]-"el"	93.937	134	[\ks\]-"x"	91.826	164	[\s\]-"sc"	1.822
105	[\ɪw\]-"il"	5.053	135	[\ks\]-"cs"	3.524	165	[\s\]-"x"	1.492
106	[\ɪw\]-"hil"	0.040	136	[\ks\]-"cc"	2.643	166	[\s\]-"z"	0.392
107	[\ja\]-"ia"	69.181	137	[\l\]-"l"	98.689	167	[\s\]-"xc"	0.121
108	[\ja\]-"ea"	30.278	138	[\l\]-"ll"	1.311	168	[\s\]-"xs"	0.008
109	[\jɐ\]-"ia"	95.343	139	[\m\]-"m"	100.000	169	[\s\]-"sç"	0.003
110	[\jɐ̃\]-"ian"	73.485	140	[\n\]-"n"	99.998	170	[\ʃ\]-"ch"	10.024
111	[\jɐ̃\]-"iam"	14.394	141	[\ɲ\]-"nh"	99.954	171	[\ʃ\]-"x"	7.382
112	[\jɐ̃\]-"yan"	2.273	142	[\o\]-"o"	97.400	172	[\ʃ\]-"sh"	0.030
113	[\jɐ̃\]-"iâm"	1.515	143	[\o\]-"ô"	1.872	173	[\t\]-"t"	99.974
114	[\jaw\]-"iau"	100.000	144	[\o\]-"ho"	0.714	174	[\tʲ\]-"t"	99.918
115	[\jɔ\]-"ió"	96.591	145	[\o\]-"eau"	0.006	175	[\ts\]-"zz"	22.222
116	[\je\]-"ie"	98.390	146	[\õ\]-"on"	81.060	176	[\tʃ\]-"t"	99.893
117	[\je\]-"iê"	1.073	147	[\õ\]-"om"	16.270	177	[\tʃ\]-"tch"	0.038
118	[\jɛ\]-"ié"	42.857	148	[\õ\]-"ôn"	2.107	178	[\u\]-"u"	93.108
119	[\jo\]-"io"	76.233	149	[\õ\]-"ôm"	0.340	179	[\u\]-"ú"	6.288
120	[\jo\]-"eo"	23.615	150	[\oj\]-"oi"	100.000	180	[\u\]-"hu"	0.370

Table 3 Voice-Spelling Links (VS-Links) that make up the 560 words and their respective Voice-Spelling Encoding Indexes (VSEI). Part 3: Voice-Spelling Links 181-228

N	VS-Link	VSEI	N	VS-Link	VSEI	N	VS-Link	VSEI
181	[\u\]-"w"	0.163	197	[\uw\]-"hul"	0.057	213	[\wẽ\]-"uên"	39.035
182	[\u\]-"oo"	0.035	198	[\ʊw\]-"ol"	100.000	214	[\wɛ\]-"ue"	34.884
183	[\u\]-"ou"	0.011	199	[\v\]-"v"	99.883	215	[\wi\]-"ui"	72.662
184	[\ũ\]-"un"	53.296	200	[\v\]-"w"	0.117	216	[\wi\]-"uí"	25.360
185	[\ũ\]-"um"	14.072	201	[\wa\]-"ua"	75.745	217	[\wɪ\]-"ue"	81.481
186	[\ũ\]-"ún"	4.894	202	[\wa\]-"oa"	21.980	218	[\wĩ\]-"uin"	90.291
187	[\ũ\]-"úm"	0.482	203	[\wa\]-"uá"	1.922	219	[\wo\]-"uo"	97.361
188	[\ũ\]-"hom"	0.076	204	[\wɐ\]-"ua"	75.325	220	[\wo\]-"uô"	2.639
189	[\ʊ\]-"o"	98.354	205	[\wɐ̃\]-"uan"	60.714	221	[\ʎ\]-"lh"	78.630
190	[\ʊ\]-"u"	1.636	206	[\wɐ̃\]-"uam"	23.214	222	[\ʎ\]-"li"	16.749

Citation: Capovilla FC. Speech-to-spelling encoding of Brazilian Portuguese: predicting misspellings to the dictation of rare words. *J Neurol Stroke.* 2026;16(2):78‒89. DOI: 10.15406/jnsk.2026.16.00656

Table 3 Continued...

191	[\uj\]-"ui"	99.008	207	[\wɐ̃\]-"uã"	5.357	223	[\ʎ\]-"le"	4.622
192	[\uj\]-"úi"	0.496	208	[\waj\]-"uai"	100.000	224	[\z\]-"s"	73.976
193	[\uj\]-"ue"	0.331	209	[\waw\]-"uau"	12.000	225	[\z\]-"z"	23.698
194	[\uj\] -"ui"	100.000	210	[\we\]-"ue"	90.000	226	[\z\]-"x"	2.280
195	[\uw\]-"ul"	97.434	211	[\we\]-"uê"	4.737	227	[\ʒ\]-"g"	50.672
196	[\uw\]-"úl"	2.452	212	[\wẽ\]-"uen"	60.965	228	[\ʒ\]-"j"	18.436

Capovilla and coworkers[8,11–14,23–30] have conducted a number of experiments that provide evidence of the model's validity in predicting the precision of spelling under dictation among college students. They have also developed a second version of the software based on cloud computing, featuring greatly improved index calculation algorithms derived from the analysis of 4,55 million SU-PU link instances. It allows doing without tables, since it automatically calculates the degree of encoding for any spoken word. The present study used such improved estimates, as summarized in Tables 1 to 5.

Table 4 comprises six columns. Column 2 (Transcript) presents the IPA renditions of the spoken words to be written under dictation. Column 3 (Word) presents the corresponding words to be written under dictation. Column 4 (Encoding Degree) presents the degree of ease expected in spelling the word to auditory dictation. Column 5 (Total Encoding) presents the degree of ease actually obtained from subjects when spelling the word to auditory dictation. The higher the correlation between Columns 4 and 5, the greater the model's predictive power. Column 6 (Familiarity) presents the incidence of the word on the Internet for two consecutive years. Zero corresponds to 1 occurrence; 1, to 10 occurrences; 2, to 100 occurrences; 3, to 1,000 occurrences; 4, to 10,000 occurrences; 5, to 100,000 occurrences; and 6 to 1,000,000 occurrences. For instance, "hecquéria" ("voluntarismo"), from the Greek ἑκὼν: *hekôn*: "conscious motivation and the voluntary pursuit of knowledge". It is a very rare word, because it appeared less than 10 times over a period of two years. In contrast, "epizeuxe" (English: "epizeuxis") from Ancient Greek ἐπίζευξις : *epízeuxis*, "a fastening upon" (from ἐπιζευγνύναι: *epizeugnúnai*, from ἐπί: *epí*: "upon" plus ζευγνύναι: *zeugnúnai*: "to yoke"): "repetition". It occurred close to 1,000 times on the Internet over two years.

Table 4 Sample of the list of 560 words with respective IPA transcription, Encoding Degree, Total Encoding, and Familiarity (Base-10 logarithm of Annual Mean frequency). Part 1

Item	Transcript	Word	Encod degree	Tot encod	Fam
1	[\v\i\.\'s\e\.\n\ʊ\]	"vis-**ce**-no"	81.10	79.60	1
2	[\a\.\s\e\.\'t\ɛ\.\ɾ\ju\]	"as-ce-**té-**rio"	71.68	72.92	1
3	[\o\.\m\e\.\'z\i\.\ʊ\]	"ho-me-**zi-**o"	68.22	69.48	1
4	[\a\.\l\i\.\'a\s\.\t\u\ɾ\]	"ha-li-**ás-**tur"	68.06	67.18	1
5	[\n\o\.\'n\i\.\l\ju\]	"no-**ní-**lio"	79.65	78.41	1
6	[\e\.\'k\ɛ\.\ɾ\ja\]	"hec-**qué-**ria"	42.29	46.02	1
7	[\e\.\p\i\.\'z\ew\.\ks\ɪ\]	"e-pi-**zeu-**xe	70.13	73.61	3
8	[\a\.\s\i\.\n\a\.\s\i\.\'f\ɔ\.\l\i\.\ʊ\]	"a-ci-na-ci-**fó-**li-o"	80.67	78.65	1
9	[\'p\a\ɾ\.\s\ju\]	"**pár**-seo"	58.06	57.59	1
10	[\a\.\'l\i\.\o\.\f\ɪ\s\]	"ha-**lí-**o-fis"	55.66	54.32	1
11	[\'e\ɾ\.\s\ɪ\]	"**her-**se"	55.98	55.24	3
12	[\'i\k\.\s\ʊ\]	"**hic-**so"	33.95	34.52	2
13	[\o\.\'m\a\.\z\ju\]	"o-**má-**seo"	61.53	58.70	1
14	[\e\.\l\i\.\'ɔ\.\k\ɾ\a\.\tʲ\ɪ\]	"he-li-**ó-**cra-te"	78.38	76.43	1
15	[\dʲ\i\k\.\s\o\.\n\i\.\'ɛj\.\ɐ\]	"dic-so-ni-**ei-**a"	82.39	80.32	1
16	[\uw\.\'s\i\.\t\ɐ\]	"hul-**si-**ta"	68.21	71.61	1
17	[\a\.\l\i\.\'aw\]	"ha-li-**al**"	64.06	61.63	3
18	[\k\wĩ:\.\'t\aw\]	"cuin-**tau**"	71.61	75.38	1
19	[\b\a\.\k\u\.\'ɔw\]	"ba-cu-**ol**"	92.51	90.43	1
20	[\a\s\.\k\u\.\'a\ɾ\]	"as-cu-**ar**"	86.40	82.05	1
21	[\'s\ej\.\s\ɪ\]	"**sei-**ce"	62.40	58.56	3
22	[\e\.\'b\aw\.\tʲ\ɪ\]	"he-**bau-**te"	59.32	61.40	1
23	[\e\.\s\e\.\'lĩ:\]	"es-ce-**lim**"	61.01	59.74	1
24	[\e\.\l\i\.\'ɛj\.\ɐ\]	"he-li-**ei-**a"	73.79	69.99	1
25	[\e\.\'z\ɛ\.\ɾ\ɪ\.\z\ɪ\]	"e-**xé-**re-se"	64.51	65.18	4
26	[\ew\.\b\a\.\'z\ɛ\.\l\ja\]	"eu-ba-**sé-**lea"	64.96	67.21	1
27	[\'wi\s\.\tʲ\ɪ\]	"**uís-**te"	61.99	65.50	1
28	[\t\a\k\.\'s\o\.\n\ɪ\.\ɐ\]	"tac-**sô-**ni-a"	60.85	61.14	1
29	[\'aw\.\z\i\.\ʊ\]	"**au-**si-o"	72.11	68.28	2
30	[\a\.\'ʎ\ɛ\.\l\ɐ\]	"a-**lhe**-la"	83.83	83.17	2

Table 5 Sample of the list of 560 words with respective IPA transcription, Encodability, Total Encoding, and Familiarity (Base-10 logarithm of Annual Mean frequency). Part 2

Item	Transcription	Word	Encod degree	Tot Encod	Fam
31	[\a\.\k\u\.\'o\.\m\e\.\t\ɾ\ʊ\]	"a-cu-**ô**-me-tro"	85.69	82.92	1
32	[\k\u\.\ɾ\a\.\'n\aw\]	"cu-ra-**nau**"	83.42	81.71	1
33	[\s\e\.\'z\ɛ\.\l\ju\]	"se-**sé**-lio"	72.68	70.51	1
34	[\o\.\z\e\.\'aɾ]	"o-ze-**ar**"	81.80	83.40	1
35	[\s\i\.\'p\a\.\l\ja\]	"si-**pá**-lia"	70.04	69.29	1
36	[\'ʃ\ɔ\.\k\ʊ\.\ɪ\]	"**chó**-cu-e"	47.43	50.71	1
37	[\s\ej\.\'s\aw\]	"sei-**çal**"	57.72	52.11	1
38	[\'ɛw\.\s\i\.\õ:\]	"**hél**-ci-on"	49.57	41.98	1
39	[\a\.\'s\i\.\dʲ\jo\]	"as-**cí**-dio"	55.86	50.23	1
40	[\a\.\l\i\.\'a\ɾ\.\s\ju\]	"ha-li-**ár**-cio"	56.57	55.99	1
41	[\a\.\l\o\.\'ks\u\.\ɾ\ja\]	"a-lo-**xú**-ria"	79.15	78.02	1
42	[\k\wo\.\'ʒ\ɛ\.\l\ʊ\]	"cuo-**je**-lo"	78.51	75.64	1
43	[\l\aw\.\'d\ɛ\.\ks\ju\]	"lau-**dé**-xio"	71.58	74.40	1
44	[\aw\.\'t\a\.\l\ja\]	"au-**tá**-lia"	60.74	62.46	1
45	[\o\.\m\i\.\'zi\.\ʊ\]	"ho-mi-**zi**-o"	66.89	68.65	2
46	[\s\i\.\'aw\]	"si-**au**"	56.71	56.89	4
47	[\m\a\.\ʎ\i\.\'s\o\ɾ\]	"ma-lhis-**sor**"	80.19	81.99	1
48	[\l\i\.\'ɔ\.\s\e\.\l\ʊ\]	"li-**ós**-ce-lo"	78.72	76.73	1
49	[\'s\aw\.\s\ju\]	"**sál**-seo"	34.44	32.43	1
50	[\k\u\.\'e\.\ɾ\ɐ\]	"cu-**e**-ra"	95.60	92.06	3
51	[\b\u\ks\.\'b\aw\.\m\ja\]	"bux-**báu**-mia"	79.45	78.01	1
52	[\'tʃ\i\.\ʎ\ɐ\]	"**tí**-lea"	53.20	51.12	1
53	[\e\.\s\e\.\'k\a\\ɾ\ja\]	"ex-ce-**cá**-ria"	66.12	64.18	1
54	[\'f\u\k\.\s\ja\]	"**fúc**-sia"	44.73	45.17	4
55	[\p\a\.\'ɾ\ɛ\.\lj\o\]	"pa-**ré**-lio"	84.41	83.03	3
56	[\'iḛ̃:\.\dõ:\]	"**iân**-dom"	53.10	56.54	3
57	[\k\õ:\.\ 'p\a\s\.\k\wo\]	"com-**pás**-cuo"	65.24	63.41	1

In Column 2, IPA transcriptions are divided into syllables, and each syllable is divided into phones. Each phone is represented by a discrete phonetic alphabet character. For instance, the word "haliástur": "bird of prey". The IPA transcription of the word "haliástur" is [\ali'asturɾ\]. It is divided into spoken syllables: [\a\.\li\.\'as\.\tuɾ\]. Each spoken syllable is divided into SU (phonemes): [\a\.\l\i\.\'a\s\.\t\u\ɾ\].

In Column 3, written words are divided into written syllables to facilitate the identification of SU-PU links (phoneme-grapheme links). For instance, the written word "haliástur" is divided into four written syllables: "ha-li-**ás**-tur". Each written syllable is made of PU (graphemes): "h", "a", "l", "i", "á", "s", "t", "u", "r". Each PU (grapheme) corresponds to a given SU (phoneme) [\a\] . [\l\] [\i\] . [\'a\] [\s\] . [\t\] [\u\] [\ɾ\]. This helps identify SU-PU links (phoneme-grapheme links): [\a\]-"ha", [\l\]-"l", [\i\]-"i", [\'a\]-"á", [\s\]-"s", [\t\]-"t", [\u\]-"u", [\ɾ\]-"r". Each SU-PU link has its own encoding index, summarized in Tables 1-3.

In Column 4, the Encoding Degree represents the average of the SU-PU Encoding Indexes of all the SU-PU links contained in the spoken word to be spelled under auditory dictation. The Encoding Degree is obtained by summing up the SU-PU Encoding Indexes of each SU-PU link that makes up each spoken word, and dividing that sum by the number of SU-PU links in that word.

In Column 5, the Total Encoding that was obtained from the subjects. It corresponds to the percentage of words correctly encoded. In Column 6, the Familiarity Degree of the word corresponds to the base 10 log of the annual incidence of that word on the Internet. The lower the Familiarity Index, the lower the efficacy of the lexical route and the higher the need to use the sublexical encoding route.

The present study

The present study uses the improved model of Portuguese Encoding based on Word Encoding Degrees (WED) and SU-PU Encoding Indexes, and provides a preliminary evaluation of its validity. According to the model, Word Encoding Degree (WED) measures how easy it is to spell a word when dictated audibly. The greater the Word Encoding Degree, the easier it is to spell that word under auditory dictation. For any given spoken word, WED corresponds to the arithmetic mean of the SU-PU Encoding Indexes involved in spelling the SU contained in that spoken word. Each SU that comprises a given spoken word has a given SU-PU Encoding Index. The greater the SU-PU Encoding Index of a given SU, the easier it is to encode that SU in writing (i.e., to spell it down in writing).

In the study, 154 students (61 from college and 93 from elementary school) spelled, under dictation, 560 very rare spoken words comprising 280 distinct SU-PU links. In this research, the dependent variable was the total encoding of each word. The total encoding was calculated by summing up the observed percentage of total encoding of each SU-PU link that made up each spoken word and dividing that sum by the number of the SU-PU links present in that word. The independent variable was the degree of encoding for each word. The procedure used to establish the encoding degree consists of summing

Citation: Capovilla FC. Speech-to-spelling encoding of Brazilian Portuguese: predicting misspellings to the dictation of rare words. *J Neurol Stroke.* 2026;16(2):78‒89. DOI: 10.15406/jnsk.2026.16.00656

up the SU-PU Encoding Indexes of each SU-PU link that makes up each spoken word, and dividing that sum by the number of SU-PU links in that word.

In sum, the present study analyzed students' spelling errors during live dictation under both auditory and visual conditions. A sample of 154 volunteer students took part in the experiment. There were 61 college students and 93 elementary-school students from 5th to 9th grade levels. The task consisted of spelling under dictation 560 spoken words with very low orthographic familiarity to maximize phonological encoding. The list of 560 words to be spelled under dictation had very low orthographic familiarity, as ascertained by the mean annual search frequency in the Google AdWords database (Gross et al., 2014; available at http://www.altgrupo.com.br/adwords/)[31] across 16 searches. The task of spelling under dictation consisted of encoding SU into PU. Encoding precision was analyzed as a function of the word's mean encoding degree, which was the average of the SU-PU Encoding Indexes of all the SU-PU links contained in the spoken word to be spelled. The present study analyzed spelling errors committed during auditory-visual dictation of 560 uncommon spoken words by 154 students (61 college students and 93 5th-9th-grade elementary school students). Encoding precision was analyzed as a function of the word's Mean SU-PU Encoding Degree, i.e., the average of the SU-PU Encoding Indexes of all the SU-PU links contained in the spoken word to be spelled under auditory-visual dictation.

Method

Subjects

A sample of 154 volunteer students took part in the study. Of the 154 students, 61 were college students, and 93 were elementary school students (18 from 6th grade, 20 from 7th grade, 28 from 8th grade, and 27 from 9th grade).

Instruments

Tables 4 and 5 present a sample of 59 out of the 560 uncommon words for spelling under dictation, along with their precise pronunciation (IPA transcription) and corresponding SU-PU Encoding Degree. The tables list words along with their respective IPA transcriptions, SU-PU Encoding Degree, Total Encoding Percentage, and Familiarity Degree (measured as the base-10 logarithm of the Annual Mean Frequency in the Google AdWords database). The use of relatively novel or rare words aims to reduce lexical recognition and increase phonological decoding. The use of IPA transcriptions helps standardize the experimenter's pronunciation across sections. During the spelling under dictation task, students used a personal response kit containing 20 sheets of paper, each with blank lines numbered from 1 to 560.

Procedure

Students collectively responded to the spelling under the auditory dictation test in their classroom over several 30-minute sessions. In each session, students received their individual response kits. The experimenter instructed students to spell each spoken word in its corresponding line. At each of the 560 items, the experimenter uttered the word preceded by its number. Thus, the experimenter uttered "Word number __ is __" followed by the word. The experimenter repeated the word five times, with a 2-second interval between repetitions. The experimenter dictated words at a rate of 1 per 15-25 seconds. The number of students per session varied. Each student was required to spell 560 words, drawn from different combinations of 228 types of voice-spelling links, across approximately 3,676 instances of voice units to be encoded.

Results

There were 521,640 discrete SU-PU encoding opportunities, 285,943 (54.82%) vowel SU, and 235,697 (45.18%) consonant SU. Out of the 521,640 encoding opportunities, 421,126 (80.73%) were appropriate spellings (231,546 vowel SU [80.98%] and 189,580 [80.43%] consonant SU), and 100,514 [19.27%] were misspellings (54,397 [19.02%] vowel SU and 46,117 [19.57%] consonant SU). Out of the 100,514 vowel and consonant misspellings, there were 96,649 [96%] substitutions (51,528 [53%] vowel SU and 45,121 [47%] consonant SU), and 3,865 [4%] were omissions (2,577 pertaining to vowels [67%] and 1,288 [33%] pertaining to consonants). Of the 96,649 substitutions, 79,518 [82%] were canonical (cf. modeled by Capovilla and coworkers (Capovilla, 2011, 2013, 2015a, 2015b; Capovilla & Casado, 2014; Capovilla et al, 2017, 2019), and 17,131 [18%] were non-canonical ones, which reflect phonetic processes. Out of the 51.528 vowel substitutions, 40,843 [79%] were canonical ones, and 10,685 [21%] were non-canonical ones. Out of the 45,121 consonant substitutions, 38,675 [86%] were canonical ones, and 6.446 [14%] were non-canonical ones.

Figure 1 represents the total encoding percentage of SU-PU links as a function of the SU-PU Encoding Indexes of the SU-PU links contained in the spelling of those spoken words. Regression analysis of total encoding of SU-PU links as a function of SU-PU Encoding Indexes for those SU-PU links was significant, $F(1, 558) = 3097.96$, $p < .0001$, $r = .921$, $r^2 = .847$, standardized coefficient Beta = .921, $t = 55.66$, $p < .0001$.

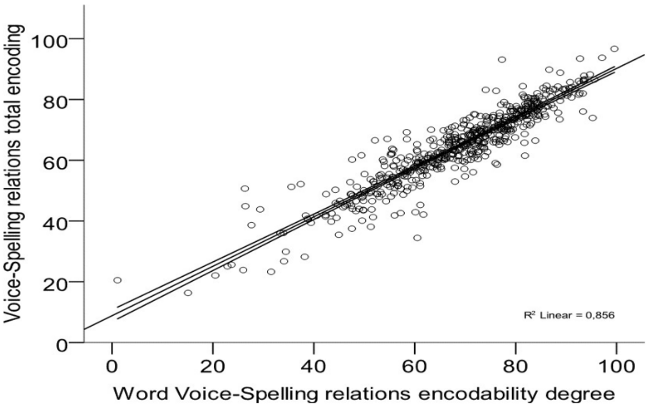

Figure 1 Scatter plot with regression line (confidence interval = .95) of percentage of total encoding of SU-PU links as a function of SU-PU Encoding Indexes of those SU-PU links contained in the spoken words to be written under auditory dictation.

Discussion

The total encoding of SU-PU links in the spoken words to be spelled under dictation increased with the SU-PU Encoding Indexes of the SU-PU links contained in the spoken words to be written under auditory dictation, as computed in accordance with the present SU-PU Encoding Degree Model. The precision of SU-PU Encoding was directly proportional to the magnitude of the SU-PU Encoding Index in Brazilian Portuguese. Over the 521,640 discrete SU-PU Encoding opportunities obtained by the 154 students while spelling the 560 words, it was found that the greater the SU-PU Encoding Index of a given SU-PU link in Brazilian Portuguese, the greater the precision with which a SU was encoded with its corresponding target PU. The greater the dominance of SU-PU links in Brazilian Portuguese, the greater the encoding precision in students' spelling under dictation. The lower the prevalence of a given SU-PU link in Brazilian Portuguese, the greater the incidence of misspellings involving that link. The frequency distribution of SU-PU links obtained in the spelling under dictation task was directly proportional to the probability distribution of SU-PU link patterns in Brazilian Portuguese. Spelling intuition (as measured by the probability of encoding a given SU with a given PU) was directly proportional to the probability distribution of SU-PU link patterns in Brazilian Portuguese.

Conclusion

The present study explained how to calculate the Word Encoding Degree of any given Portuguese word by hand. It provided a sample of 228 SU-PU Encoding Indexes from the 367 existing in Portuguese. The 228 SU-PU Encoding Indexes presented pertained to the 228 SU-PU Encoding links existing in the list of 560 words to be spelled under auditory dictation in the present task. It provided the Word Mean Encoding Degree of a sample of 57 words out of the list of 560 words to be spelled under dictation in the task of the present study.

Using the 560-word list to assess spelling under dictation in students, the present study demonstrated that the degree of encoding that characterizes each of the many SU-PU Encoding links in spoken words to be written under auditory dictation may predict behavior in choosing among different PUs during encoding.

Results revealed that the probability of choosing a given PU (e.g., grapheme) to encode a given SU (e.g., phoneme) was a positive function of the magnitude of the SU-PU Encoding Index for the target SU-PU Encoding link, relative to distracter SU-PU Encoding links. That is, the probability of choosing a target PU to encode a target SU is a positive function of the relative proportion with which that target PU encodes that target SU, as compared to other distracter PUs that may encode that same target SU, in a sample of 4.55 million link instances.

Since the SU-PU Encoding Index Model can predict the behavior of choosing from among different PUs to encode Portuguese speech, a number of technological and theoretical implications follow. Word Encoding Degree is a scalar variable that allows systematic control in experimental and statistical designs, serving as both an independent variable and a covariate. By calculating the Word Encoding Degree of words pertaining to different word lists, one may be able to compare reading and spelling data obtained when using those lists. By doing so, one may help establish productive communication among researchers who use different research instruments. One may also produce word lists free from arbitrary classification of word types, and capable of fine, precise tuning. The perspective of systematically regulating word reading and word spelling difficulties also benefits literacy procedures and instruments in schools, as well as reading and spelling rehabilitation procedures in clinical practice.

In terms of theory, results seem to be compatible with the suggestion that the explanation of the primary mechanisms underlying phonological recoding in reading and spelling may be found in intuition based on statistical distributions of SU-PU links, in addition to conversion rules. If conversion rules are to provide secondary support for reading and spelling intuition, their effectiveness likely depends on establishing a baseline of intuitive performance. The greater the intuition, the larger the ground for supporting rules.

Encoding intuition may play an important role in spelling under dictation. It may provide the primary substrate necessary for spelling rules to start making sense and to become an effective secondary support factor. The present study provides support to the idea that encoding intuition results from the statistical distribution of Phoneme-Grapheme associations in the mother language of literate writers, and that the computerized mapping of that distribution of Phoneme-Grapheme associations in a representative *corpus* of lexical items can provide Encoding Indexes useful to calculate the Mean Encoding Degree of any given word.

Implications of the present model for the analysis of spelling and reading processes

The present model challenges the assumption that encoding is essentially a rule-governed behavior that depends on the application of phoneme-to-grapheme conversion rules painstakingly learned through instruction. Instead, it seems to be an intuitive process based on contiguity derived from that particular frequency distribution of phoneme-to-grapheme links, which pervades the world of print in a given universe of speech-to-print links. Fully developed spelling competence involves an equilibrium among several semantic, phonological, orthographic, grammatical, etymological, and phonetic factors, amongst others.[23,25,32] Yet underlying encoding competence is an encoding performance based on intuitive know-how processes that operate below the awareness threshold, whereas know-why rules pertaining to semantics, grammar, etymology, and so on operate above it.

As with the present study on encoding, a parallel study on decoding challenges the assumption that decoding is an essentially rule-governed behavior that depends on the application of grapheme-to-phoneme conversion rules painstakingly learned through instruction. Instead, it seems to be an intuitive process based on contiguity derived from that particular frequency distribution of grapheme-to-phoneme links, which pervades the world of speech in a given universe of print-to-speech links. Fully developed reading competence involves an equilibrium among several semantic, phonological, orthographic, grammatical, etymological, and phonetic factors, amongst others. Yet underlying decoding competence is a decoding performance based on intuitive know-how processes that operate below the awareness threshold, whereas know-why rules pertaining to semantics, grammar, etymology, and so on operate above it.

Capovilla's[14] study investigated the precise nature of the phonological route decoding processing. It argues that, contrary to what is assumed in the literature, the decoding process is not based on conversion rules but rather on conversion-based unconscious intuition derived from contingencies. It argues that precise decoding may occur in the absence of rules, simply through intuition induced by contingencies, i.e., by analogy or stimulus generalization. The study discovered that phonological decoding may happen in direct opposition to conversion rules. To ensure reading was phonological rather than lexical, the study used a list of nonwords to be read aloud, one by one. In that study, 40 elementary school students (10 from each grade: 3rd, 4th, 5th, and 6th, ages 8 to 11) from a private school were individually exposed to an 18 2-syllable non-word reading-aloud task designed to engage phonological route reading. The purpose was threefold: (1) Assessing whether students exhibit the appropriate reading-aloud performance, which is to pronounce “s” as /s/ when preceding unvoiced consonants and /z/ when preceding voiced consonants. (2) Assessing whether students can recite, evoke, or formulate a rule that can explain or at least describe their own behavior. (3) Assessing whether the teachers had ever taught their students any rule regarding that case and whether they had ever been aware of the existence of any rule regarding that case. In the study, half of the students were interviewed before being presented with the non-word reading task (Instruction Group). The other half of the students were interviewed only after being presented with the non-word reading task (Naive Group). Results were compared to assess

Citation: Capovilla FC. Speech-to-spelling encoding of Brazilian Portuguese: predicting misspellings to the dictation of rare words. *J Neurol Stroke.* 2026;16(2):78‒89. DOI: 10.15406/jnsk.2026.16.00656

the effect of instruction on rule formulation and sensitivity to one's own behavior during the reading-aloud task.

If reading and spelling are fluent and precise, the assumption is that conversion rules have been explicitly taught or at least extracted. Yet, there are cases in which performance is based on conversion intuition derived from conversion induction rather than on conversion rules established by conversion instruction. In those cases, no explicit rule formulation takes part, neither during teaching procedures nor during learning processes. A simple example can illustrate this point. For phonetic reasons,[10–14,20–27] in Portuguese, grapheme "s" is uttered: as /s/ when preceding unvoiced consonants (e.g., /p/, /t/, /k/, /ʃ/, /f/); and as /z/ when preceding voiced consonants (e.g., /b/, /d/, /g/, /ʒ/, /v/). This phonetic phenomenon operated below the threshold of awareness. Speakers and readers are completely unaware of it. When asked to produce the phone corresponding to the grapheme "s", they will always sound /s/. When asked whether "s" could sound differently, they will reply, "Yes! 's' sounds as /z/ when it appears in between two vowels". When asked if "s" sounds as /z/ on any other occasion, they will usually reply: "No." So there is a straightforward grammatical orthoepic rule that establishes the way the grapheme "s" is to be pronounced when reading aloud.

Even when presented with words in which "s" precedes a mixed sequence of voiced consonants (e.g., /b/, /d/, /g/, /ʒ/, /v/) and unvoiced consonants (e.g., /p/, /t/, /k/, /ʃ/, /f/); students still will maintain that it always sounds as /s/. However, when asked to read aloud (i.e., to sound aloud) those words, they start noticing that "s" sounds differently, sometimes as /s/, and sometimes as /z/. Unless they have had training in linguistics or speech-language pathology, they are usually unable to identify the single feature common to all cases in which "s" sounds as either /s/ or /z/. That is, they are usually unable to state a simple rule such as: "The letter 's' sounds as /s/ when it precedes unvoiced consonants, and as /z/ when it precedes voiced consonants."

Native good readers tend to perform perfectly well at decoding novel words and non-words, as if they knew the specific conversion rule. That is as if their behavior was controlled specifically by that rule. Yet, their behavior is not rule-governed but rather contingency-shaped. It is induced intuitively by the statistical distribution of grapheme-phone combinations in Portuguese.

During literacy acquisition,[15–17,33] Brazilian children read aloud perfectly well any non-word one can invent involving the grapheme "s" followed by unvoiced consonants ("sc", "sf", "sk", "sp", "sq", "sr", "st", "sx") or voiced consonants ("sb", "sd", "sg", "sj", "sl", "sm", "sn", "sv", "sz"). In reading-aloud tasks, when asked to pronounce those non-words they had never seen or heard before, Brazilian children usually sound out the grapheme "s" as /s/ when preceding unvoiced consonants ("sc", "sf", "sk", "sp", "sq", "sr", "st", "sx"). And they usually sound out the grapheme "s" as /z/ when preceding voiced consonants ("sb", "sd", "sg", "sj", "sl", "sm", "sn", "sv", "sz"). They usually get it right: they demonstrate that they know how to sound out correctly, but almost never know why they do it the way they do. Such a perfect performance is completely intuitive-based and not rule-governed at all. In such a non-word reading-aloud task, their competence in pronouncing "s" as either /s/ or /z/ is an unconscious tacit knowledge, an implicit phonetic know-how skill.

Native good readers have tacit know-how competence but not declarative, explicit know-why knowledge. They know how to perform, but they do not know why they perform the way they do. If their know-how performance is impeccable despite the absence of any explicit know-why knowledge, then we can state that their conduct is not at all rule-governed behavior, but rather intuitive, contingency-shaped performance. Namely, an implicit phonetic know-how skill. They have never extracted the rule, have never been taught the rule, and have never been aware of its existence. Upon interview, their teachers admit to never having taught the rule to their students. They admit they were not even aware of the rule's existence. In fact, despite performing correctly, both students and teachers showed they did not realize what they were doing. Before the task, upon being asked how "s" is to be pronounced, they uniformly affirm that "s" is to be pronounced as /s/, except in between two vowels, when it is to be sounded as /z/. Upon being asked whether "s" sounds like /z/ in any other circumstance, they uniformly reply "No". Yet, during the non-word reading task, the very same readers pronounce "s" as /z/ when it precedes a voiced consonant. Before being asked to listen to their own reading-aloud performance, they are not even aware of how they have been pronouncing. When asked repeatedly to describe how they pronounced in the task, they admitted that sometimes they pronounced "s" as /z/ and other times as /s/, but they failed to explain why. They were unable to formulate a rule. Upon being debriefed on what they did and the rule that could account for it, they almost always act surprised and report that they thought they pronounced as expected: "s" always sounds as /s/ when preceding a consonant and always sounds as /z/ in between vowels, but only in between vowels. Thus, good readers and their teachers report being unaware not only of the conversion rule (involving voiced and unvoiced consonantal attacks) but also of their own appropriate decoding behavior. All these features suffice to qualify recoding performance as being automatic, intuitive, and subconscious, a know-how performance established by recoding induction via imitation practice rather than a know-why knowledge established by recoding instruction based on meta-language conversion rules.

In Capovilla's[14] study, students who were interviewed before the reading task stated the same grammar inter-vowel-position rule: "We always utter 'S' as /s/, except when it occurs in between vowels. In that case, and only in that case, we utter 'S' as /z/." During reading aloud task, younger students failed to perform phonetic junctions due to syllabification. Long inter-syllabic pauses prevented phonetic junction and, thence, "S"-coda assimilation. In contrast, older students showed perfectly fluent and precise know-how decoding performance. Due to their extensive history of drill-and-practice modeling and shaping, they performed phonetic junction and thus presented perfect "S"-coda voicing in sharp contrast to what they expected to do, based on their grammatical assumptions. They were surprised to realize they had performed perfectly, in direct contrast to the grammar rule they had just stated. Their teachers were subjected to the same task and were equally puzzled and even ashamed of failing to perform in accordance with their own stated grammar rules.

In conclusion, when reading fluency emerges, syllabification is replaced by a phonetic junction. Therefore, the emergence of the phonetic assimilation process of "S"-coda voicing may be regarded as a reading-fluency landmark. Once reading fluency emerges, the phonetic assimilation process of "S"-coda voicing emerges with it as an unconscious process that occurs independently from grammar position rules, and in complete opposition to them. Know-how decoding performance proved to be compatible with intuition derived from drill-and-practice modeling and shaping, rather than with grammar rules acquired through painstaking instruction. Know-how decoding performance may be dissociated from know-why decoding knowledge, which is based on orthoepic grammatical rules that specify how grapheme pronunciation depends on grapheme position. Know-how decoding performance may occur in the absence of rules, or even in total contrast to the rules that readers subscribe to, due to phonetic processes that readers are unaware of, such as assimilation.

Therefore, the assumption that "the phonological route works based on the application of conversion rules" is unwarranted. Instead, the phonological route seems to operate primarily on intuitive decoding derived from the statistical distribution of grapheme-phoneme patterns in language, implemented through drill-and-practice modeling and shaping. Orthoepic grammar rules are added secondarily to inform intuition, making conversion dependent on grapheme position. The same process applies to spelling.

The findings suggest a dissociation between know-why behavior (declarative knowledge based on rules acquired through conscious effort during instruction) and know-how performance (on-task, intuitive skill acquired unconsciously through repeated exposure to instances in drill-and-practice training).

To maximize the engagement of the phonological (sublexical) route, the present encoding study used a spelling-to-auditory-dictation task with rare words, whereas Capovilla's[14] decoding study used a reading-aloud task with nonwords. Because non-words do not exist in any lexicon, if correct reading occurred, then the phonological-route reading could be held responsible for it. But if, as the dual-route reading model affirms, decoding is based on conversion rules, then it should be possible to demonstrate the operation of rules. If students perform perfectly, and yet are unable to recite any rule that might explain why they did the way they did, or even describe what they did, this suggests their performance is automatic, habitual, contingency-shaped by their idiom, and not rule-governed by grammatical position rules. That conclusion would be supported if students prove unable to either evoke or formulate a rule. The support would be strengthened if, upon being questioned, their teachers report that they have never taught that rule to their students, and the teachers themselves confess that they never realized that the letter "s" may sound as /z/ when preceding specific consonants (i.e., the voiced ones). If, upon being questioned, the teachers show that they are unaware of any change in the sound of "s" except when "s" appears in between two vowels, then our confidence that the process is an unconscious phonetic one would increase. If teachers indicate that they too assumed "s" sounds as /z/ only in between vowels, and if they declare that that was all they ever knew and taught their students regarding that specific case of how to pronounce "s", then we could rule out the effect of any rule.

Since it is usually supposed[2,5] that the phonological route works based on the application of conversion rules, it is usually assumed that grapheme-phone decoding is a rule-governed behavior. Capovilla's[14] study offered an alternative explanation based on the classic dichotomy between rule-governed behavior (declarative know-why behavior) and contingency-shaped performance (intuitive know-how response patterns and tendencies).

Capovilla[14] argues that the processes are distinct, with a distinct specific neuroanatomical substrate in each case. Rules are conscious declarative statements (regarding what to do in given circumstances) that are acquired painstakingly as a consequence of conscious effort under explicit instruction. The hippocampus is the center of explicit (declarative) long-term memory and is responsible for storing rules essential for rule-governed behavior. In contrast, intuitive tendencies, response patterns, and on-the-spot performance are usually shaped unconsciously by the statistical distribution of passive SS (stimulus-stimulus: grapheme-phone) associations and active SRS (stimulus-response-stimulus) contingencies. The basal ganglia are involved in procedural skills and habits that constitute implicit long-term memory, including response patterns and intuitions formed through task-based drill-and-practice activities based on instances. The paper argues that non-word reading is not rule-governed but, rather, is contingency-shaped performance based on intuitive tendencies formed by the statistical distribution of print-to-speech associations, which can be predicted from language-structural properties. For reading aloud, those properties pertain to the probability distribution of grapheme-phone associations, with which each of several phone sounds is associated with a given grapheme.

According to this interpretation, the statistical distribution of speech units, as well as of print-to-speech associations the child is exposed to from birth, tends to form, in the long run, intuitions and habits regarding reading and spelling. Those intuitions and habits interact with phonetic processes. Phonetic processes affect pronunciation, reading aloud, and spelling under dictation. That happens because spelling involves reverberation in the phonological loop of the phonological image in the phonological buffer.[34] Dyslexics differ from neurotypical readers (i.e., good readers) in terms of hyperactivation of several areas, including the inferior frontal gyrus (IFG), basal ganglia and thalamus, and precentral gyrus. Dyslexia is typically associated with reduced activity (hypofunction) in the left temporoparietal regions involved in phonological processing, particularly in adults.[35] However, increased activation (hyperactivation) in frontostriatal regions, e.g., the inferior frontal gyrus, pre-central gyrus, and striatum, has also been reported.[36,37] The striatum is involved in phonological processing during reading development and continues to be involved in phonological processing into adulthood.[38–41] Although regional increases in activation are often interpreted as being compensatory, hyper-activation could also indicate dysregulation of frontostriatal pathways, as in developmental stuttering (Giraud et al., 2008).[42] Several studies report hyperactivation of the frontal and striatal regions in patients with dyslexia during reading tasks. Hyper-activation in these regions is typically interpreted as a form of neural compensation associated with articulatory processing. However, according to Hancock, Richlan, and Hoeftl,[43] frontostriatal hyper-activation in dyslexic patients can also arise from fundamental impairment in phonological reading and implicit sequence learning relevant to early language acquisition. Recent models with increased temporal sampling rates have been proposed to account for phonological deficits in dyslexia. The frequency of ongoing neural oscillations in auditory brain regions closely matches the rate of formant transitions in speech and may correspond to the neural coding of phonemic features.[44] In dyslexia, these oscillations occur at a higher frequency, which could reflect a finer-grained temporal representation, i.e., so that speech information is 'packaged' into more units in dyslexia than in typical readers.[45,46] According to this model, hyperactivity in frontostriatal regions in dyslexics could reflect an additional, but non-compensatory, demand for sequence or phonological processing in the basal ganglia due to processing phonological representations at a finer level of detail than typical readers. In this case, one could expect a neuroanatomical correspondence between regions associated with phonological processing and those showing hyperactivity in dyslexia.

To understand the functional significance of hyper-activation in dyslexic patients, Hancock, Richlan, and Hoeftl[36] reviewed current evidence supporting the compensation hypothesis and investigated the anatomical overlap between hyper-activated regions and neural systems involved in articulation, phonological processing, and implicit sequence learning. They found anatomical convergence between hyper-activation regions and regions supporting articulation, consistent with the proposed compensatory role of these regions. However, they also found low anatomical convergence with regions involved in phonological and implicit sequence learning. Stronger support was found for the compensatory hypothesis. In a quantitative meta-analysis of eight imaging studies that included an intervention

Citation: Capovilla FC. Speech-to-spelling encoding of Brazilian Portuguese: predicting misspellings to the dictation of rare words. *J Neurol Stroke.* 2026;16(2):78–89. DOI: 10.15406/jnsk.2026.16.00656

component, Barquero, Davis, and Cutting (2014)[47] reported increases in intervention-related activation in the right inferior frontal gyrus/insula, left inferior frontal gyrus, left thalamus/basal ganglia, and other regions during reading tasks. According to Barquero et al., these regions could reflect a wide range of processes that may contribute to gains in reading ability. Several studies have reported hyperactivation in the frontal and striatal regions in individuals with reading disorder (RD) during reading-related tasks. Both developmental and intervention studies suggest that hyperactivation of the right inferior frontal gyrus has a compensatory role. The evidence reviewed so far supports the role of the basal ganglia in the automatization of phonological and articulatory processes in reading. Incidentally, we notice the involvement of prosody sensitivity disorders in dyslexia[48] as well as therapeutic gains accruing from musicalization and dancing[49] upon sensory integration and temporal processing, which are under the control of the vestibule-cerebellum processing.[40]

The theoretical and experimental model described in the present study may serve as a sensitive, precise, and reliable framework for advancing research on reading and spelling errors in both neurotypical and neuroatypical students.[50]

Acknowledgements

None.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. Seidenberg MS, McClelland JL. A distributed, developmental model of word recognition and naming. *Psychol Rev.* 1989;96(4):523–568.
2. Coltheart M. Modeling reading: the dual-route approach. In: Snowling MJ, Hulme C, editors. *The Science of Reading: A Handbook.* Oxford, GB: Blackwell; 2005:6-23.
3. Coltheart M, Curtis B, Atkins P, et al. Models of reading aloud: dual--route and parallel-distributed-processing approaches. *Psychol Rev.* 1993;100(4):589–608.
4. Coltheart M, Rastle K, Perry C, et al. DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol Rev.* 2001;108(1):204.
5. Swanson HL, Harris KR, Graham S, eds. *Handbook of Learning Disabilities.* 2nd ed. New York, NY: Guilford Press; 2014.
6. McClelland JL, Rumelhart DE. An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychol Rev.* 1981;88(5):375–407.
7. Rumelhart DE, McClelland JL. An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychol Rev.* 1982;89(1):60–94.
8. Capovilla FC. Producing and recognizing speech, writing, and sign language. *Rev Psicopedagogia.* 2023;40(122):122–158.
9. Capovilla FC, Capovilla AGS, Macedo EC. Analisando as rotas lexical e perilexical na leitura em voz alta. In: Pasquali L, org. *Instrumentação Psicológica: Fundamentos e Práticas.* Porto Alegre, RS: ArtMed; 2008:1–24.
10. Capovilla FC. Paradigma neuropsicolinguístico: refundação conceitual e metodológica. In: Capovilla FC, org. *Transtornos de Aprendizagem.* 2nd ed. São Paulo, SP: Memnon; 2011:42–131.
11. Capovilla FC. Paradigma neuropsicolinguístico para refundação conceitual e metodológica. In: Capovilla FC, Raphael WD, Mauricio AC, orgs. *Novo Deit-Libras.* 3rd ed. São Paulo, SP: Edusp; 2013:73–156.
12. Capovilla FC. Nossa língua portuguesa: como avaliar o grau de cifrabilidade. In: Pereira R, org. *Abordagem Multidisciplinar da Aprendizagem.* Viseu, Portugal: Qualconsoante; 2015:481–503.
13. Capovilla FC. Paradigma matricial de linguagem oral, escrita e de sinais. In: Santos FH, Andrade V, Bueno O, orgs. *Neuropsicologia Hoje.* 2nd ed. Porto Alegre, RS: Artmed; 2015.
14. Capovilla FC. Phonological reading route in Portuguese. *Int J Psychol Neurosci.* 2023;9(1):1–28.
15. Seabra AG, Capovilla FC. *Problemas de Leitura e Escrita.* 6th ed. São Paulo, SP: Memnon; 2011.
16. Seabra AG, Capovilla FC. *Alfabetização: Método Fônico.* 6th ed. São Paulo, SP: Memnon; 2021.
17. Seabra AG, Capovilla FC. *Teste de Competência de Leitura de Palavras e Pseudopalavras.* 2nd ed. São Paulo, SP: Memnon; 2023.
18. Assef ECS, Capovilla AGS, Capovilla FC. Computerized Stroop test to assess selective attention in children with ADHD. *Span J Psychol.* 2007;10(1):33–40.
19. Capovilla FC, Cecconi C. Leitura e escrita da língua portuguesa. In: Associação Brasileira de Dislexia, ed. *Teoria e Prática em Dislexia.* Campinas, SP: Memnon; 2024:195–201.
20. Capovilla FC, Casado K. *Quebrando o Código do Português Brasileiro.* São Paulo, SP: Memnon; 2014.
21. Capovilla FC, Coelho R, Luz TT, et al. Unravelling Portuguese phonological encodability. *Int J Psychol Neurosci.* 2017;3(1):1–47.
22. Capovilla FC, Luz TT, Graton-Santos LE, et al. Phonological route processing. *Rev (Con)Textos Linguísticos.* 2019;13(24):24–35.
23. Capovilla FC. Avaliando fluência via processo fonético de vozeamento. In: Salles JF, Navas ALPGP, orgs. *Avaliação da Linguagem.* São Paulo, SP: Vetor; 2021:145-154.
24. Capovilla FC. Compreendendo a fala recebida por audição, visão e tato. In: Soares AM, Capovilla FC, Simão JROR, Neves LM, orgs. *Caminhos da Aprendizagem.* Belo Horizonte, MG: Artesã; 2021:183-198.
25. Capovilla FC. Dificuldades e desafios na alfabetização. In: Brasil, org. *RENABE.* Brasília, DF: MEC; 2021:207–241.
26. Capovilla FC. How to conduct a systematic and comprehensive assessment of the language development of deaf Brazilian schoolchildren. *Rev Psicopedagogia.* 2024;41(126):444–488.
27. Capovilla FC. Contribuições para atualizar a definição de dislexia. In: Associação Brasileira de Dislexia, ed. *Teoria e Prática em Dislexia.* Campinas, SP: Memnon; 2024:150-188.
28. Capovilla FC, Cecconi C. Portuguese and Libras (Brazilian Sign Language). *Rev Psicopedagogia.* 2022;39(119):153–171.
29. Capovilla FC, Cecconi C. Língua portuguesa brasileira falada. In: Associação Brasileira de Dislexia, ed. *Teoria e Prática em Dislexia.* São Paulo, SP: Memnon; 2024:189-194.
30. Capovilla FC, Nico MAN. Teste de detecção de acentuação tônica. In: Associação Brasileira de Dislexia, org. *Transtornos de Aprendizagem.* Viseu, Portugal: QualConsoante; 2019:163–191.
31. Gross MS, Liu NH, Contreras O, et al. Using Google AdWords for recruitment. *J Med Internet Res.* 2014;16(1):295–305.
32. Capovilla FC. Pragmática da comunicação linguística. In: Valle LER, Capovilla FC, orgs. *Perspectivas em Transtornos.* São Paulo, SP: Memnon; 2023:33–6356.
33. Seabra AG, Capovilla FC. *Alfabetização Fônica.* 5th ed. São Paulo, SP: Pearson Clinical Brasil; 2020.
34. Baddeley A, Gathercole S, Papagno C. The phonological loop as a language learning device. *Psychol Rev.* 1998;105(1):158–173.

35. Linkersdörfer J, Lonnemann J, Lindberg S, et al. Grey matter alterations in dyslexia. *PLoS One.* 2012;7(8):e43122.
36. Hoeft F, McCandliss BD, Black JM, et al. Neural systems predicting dyslexia outcomes. *Proc Natl Acad Sci U S A.* 2011;108(1):361–366.
37. Kronbichler M, Hutzler F, Staffen W, et al. Dysfunction of reading areas in dyslexia. *Neuropsychologia.* 2006;44(10):1822–1832.
38. Abdullaev YG, Melnichuk KV. Cognitive operations in the caudate nucleus. *Neurosci Lett.* 1997;234(2):151–155.
39. Abdullaev YG, Bechtereva NP, Melnichuk KV. Neuronal activity in cognitive tasks. *Behav Brain Res.* 1998;97(1):159–177.
40. Booth JR, Wood L, Lu D, et al. Basal ganglia and cerebellum in language processing. *Brain Res.* 2007;1133(1):136–144.
41. Crosson B, Benefield H, Cato MA, et al. Basal ganglia and frontal activity in language. *J Int Neuropsychol Soc.* 2003;9(7):1061–1077.
42. Giraud AL, Neumann K, Bachoud-Levi AC, et al. Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. *Brain Lang.* 2008;104(2):190–199.
43. Hancock R, Richlan F, Hoeft F. Frontostriatal circuits in reading disorder. *Neurosci Biobehav Rev.* 2017;72:243–260.
44. Boemio A, Fromm S, Braun A, et al. Hierarchical and asymmetric temporal sensitivity in auditory cortices. *Nat Neurosci.* 2005;8(3):389–395.
45. Lehongre K, Morillon B, Giraud AL, Ramus F. Impaired auditory sampling in dyslexia. *Front Hum Neurosci.* 2013;7:454.
46. Lehongre K, Ramus F, Villiermet N, et al. Altered gamma sampling in dyslexia. *Neuron.* 2011;72(6):1080–1090.
47. Barquero LA, Davis N, Cutting LE. Neuroimaging of reading intervention. *PLoS One.* 2014;9(1):e83668.
48. Nico MAN, Belli A, Capovilla FC, et al. Overview of dyslexia in Brazil. In: Elbeheri G, Siang L, editors. *Routledge International Handbook of Dyslexia in Education.* Routledge; 2022:35–42.
49. Nadalim CF, Capovilla FC, Marques F, et al. *Maravilhamento.* São Paulo, SP: Desvendério; 2020.
50. Oliveira JBA, Navas AL, Buchweitz A, et al. O Limiar da Decodificação. Rio de Janeiro, RJ: Instituto IDados; 2025.

Citation: Capovilla FC. Speech-to-spelling encoding of Brazilian Portuguese: predicting misspellings to the dictation of rare words. *J Neurol Stroke.* 2026;16(2):78–89. DOI: 10.15406/jnsk.2026.16.00656