EFL Grapho-Phonemics: The *Teachability* of Stressed Vowel Pronunciation Rules

Despite the existence of a vast and solid heritage supporting their validity and reliability, pronunciation rules that assist the phonemic interpretation of graphemic structures are not usually taught in the EFL classroom at any levels. Since a likely reason for this absence might lie in the intrinsic complexity of the English writing system, a convenient reduction is presented in the form of ten basic rules for the interpretation of vocalic graphemes in stressed syllable. These rules are understood within an EFL oriented conceptual frame that introduces distinctions between oxytone/paroxytone/proparoxytone structures, as well as systemic/specific, post-nuclear/pre-nuclear and adjacent/distant grapho-phonemic contexts. With this, I attempt to generate the kind of grapho-phonemic knowledge that might be useful within the EFL context. The reliability and representativity of these rules have been tested against a wordlist of 5000 frequent English words. Notions of what to teach and in what order can be derived from our findings. A rich array of results is presented that might be further explored and discussed by EFL instructors.

Keywords: English orthography; phonics; grapho-phonemics; reading; EFL teaching; English pronunciation

La Grafo-Fonémica en el Entorno EFL: Sobre la fiabilidad de las Reglas de Pronunciación de Vocales Acentuadas

A pesar de la vastedad y solidez de los conocimientos que a día de hoy validan las reglas de pronunciación que rigen la interpretación fonémica de las palabras inglesas, dichas reglas no son generalmente enseñadas en el ámbito del inglés para extranjeros, en ningún nivel. Dado que una posible razón de esta ausencia radica en la complejidad del sistema ortográfico inglés, en el presente estudio presentamos una conveniente reducción en forma de diez reglas básicas para la interpretación de grafemas vocálicos en sílaba acentuada. Estas reglas se enmarcan dentro de un tejido conceptual sensible a distinciones entre palabras agudas, llanas y esdrújulas, así como a distinciones con respecto al carácter genérico/específico, post-nuclear/pre-nuclear y adyacente/distante de los indicadores grafo-fonémicos. Con ello pretendemos generar el tipo de conocimientos que pueda ser de utilidad en el contexto EFL. La fiabilidad de estas reglas se ha puesto a prueba en una lista de 5000 palabras frecuentes del inglés. De este modo se derivan nociones con respecto a qué resulta oportuno enseñar, y en qué orden. El estudio incluye abundantes resultados que podrán ser libremente explorados por los profesores de inglés.

Palabras clave: Ortografía inglesa; phonics; grafo-fonémica; enseñanza del inglés como lengua extranjera; lectura; pronunciación

1. INTRODUCTION

Any EFL teaching method that at some point relies on the written word very soon has to face the challenges posed by grapheme/phoneme correspondences in English. Grapho-phonemic training is, nevertheless, blatantly absent from most EFL handbooks to this day. When Charles W. Kreidler discussed such absence as early as in 1972, he argued that "[w]e don't teach the elementary student about English orthography because we really don't understand the nature of our spelling system" (Kreidler 1972, 4). However, English grapheme/phoneme correspondences had already been studied quite thoroughly by a number of scholars (Wijk 1966; Venezky 1970). Further effort was invested during the 1970s and 1980s into discerning the systematics of English orthography at the grapho-phonemic level. Extensive corpora were examined and the predictability of phoneme-to-grapheme and grapheme-to-phoneme correspondences were empirically calculated. Uninterrupted though increasingly sparse research looked into the processes of accessing mental lexicon through printed words (Frederiksen and Kroll, 1976), the cognitive aspects of native and non-native dealings with orthography (Schwartz et al. 2007), or the possibility of automatic grapheme-to-phoneme and phoneme-to-grapheme conversions (Daelemans and Bosch 1996).

Concern with orthography within EFL teaching has been rather marginal as a whole, although a few scholars, like H. D. Brown (1970), S. Schane (1970) or C. W. Kreidler (1972) have tackled the issue. More recently, specialists have argued that EFL teachers must "understand the correspondences between English phonology and English orthography" in order to teach learners "to predict the pronunciation of a word given its spelling" (Celce-Murcia et al. 1996). Celce-Murcia and her colleagues build upon the work of W. B. Dickerson (1984, 1987, 1990), which is remarkably revealing in relation to the prediction of stressed-syllable location; and, collaterally, also in the prediction of the value of vocalic graphemes in stressed position; the latter, mostly when it is a matter of choosing between the so-called lax and tense pronunciations[1]. Stressed <a> in a monosyllabic word, for example, will be /æ/ when closed by a single consonant, *rat*, and /eɪ/ when closed by consonant plus silent <e>, *rate*. Wijk's and Venezky's rendering reflected a much more complex panorama where stressed <a> regularly predicts other values in stressed position: /ɑː/ in *spa*, *schwa*, *ah*, *wad*, *wan*, *art*, *car*, *smart*, etc.; /ɔː/ in *war*, *quartz*, *ward*, *dwarf*, etc.; /ɒː/ in

[1] Many approaches to grapho-phonemics focus on this distinction, as if vocalic letters were to have mainly two regular values: the long/tense/free—depending on the author—and the short/lax/checked.

all, *salt*, *bald*, *talk*, etc. Celce-Murcia and her colleague's simplification is, however, common procedure. When incorporated to EFL programs, pronunciation rules for vowels are often reduced to the CV, CVCe, and CVVC rules that predict long pronunciation as in *be*, *Pete* and *seed*; and CVC predicting short pronunciation as in *wet* (Olshtain 2001, 209). Ediger (2001, 157) includes the Vr/CVr environment that guides predictions in the case of *art*, *car*, or *her*, but she overlooks another known, relevant and regular context that involves <r>: *care*, *here*, *fire*, *pure*, which contrast with *car*, *her*, *fir* and *purr*.

At any rate, despite the rich theoretical and empirical background available to assist EFL instructors in their dealings with the grapho-phonemic issue, there seems to be a gap between orthography researchers and teaching specialists. A very small and general part of what is known about English orthography occasionally finds its way into EFL teaching practice. This might be partly due to the predominance, since the early 80s, of communicative and meaning-focused approaches to EFL. A situation that is beginning to change as reliable evidence piles up suggesting that form-focused interventions tend to facilitate learning (Long and Robinson 1998; Lan and Wu 2013). Recent research shows that learners naturally extract statistic regularities from written language, and that such statistic learning can be enhanced through specific training (Doignon-Camus and Zagar 2014). Mere exposure to orthography, on the other hand, has been shown to have a positive effect on the development of phonological awareness (Cheung et al. 2001; Escudero et al. 2008). Explicit teaching of phonics and letter-to-sound correspondences seems to improve learning results not only in relation to EFL pronunciation (Rangriz and Marban 2015), but also in reading comprehension and literacy (Martínez 2011). The time seems ripe for a reconsideration of the role of grapho-phonemic training in the EFL classroom.

A second reason for the general neglect of the grapho-phonemic competence in the EFL classroom has to do with its intrinsic complexity. Book-length developments of the subject such as Cummings' description of American spelling (1988) or Bozman's handbook for students of phonetics (1988) cannot be introduced into the EFL arena without adequate reduction and adaptation. The cumulative nature of these and other fundamental works is far from EFL-friendly. The system encompasses rules overriding rules and exceptions that seem to be regular in their own way, only to have further exceptions that conform to the initial rules, etc. However, these rules are rarely presented within a pedagogical hierarchy distinguishing between essential and complementary rules, or determining logical sequences of

overriding. We know very little about each rule in particular—the relative amount of words they regulate, for example, with more or less regularity.

A third reason for the paltriness of EFL grapho-phonemic training might be the lack of direct applicability of much of the existing empirical research, which is almost entirely concerned with measuring the consistency of English orthography. Hanna's exhaustive exploration of phoneme-to-grapheme correspondences, uncovering an 80% consistency within a corpus of 17, 310 English words, has had no pedagogical applications (Hanna et al. 1966). Berdnt's reversal of Hanna's data, in order to determine grapheme-to-phoneme consistency has remained, over the years, mostly anecdotal (Berdnt et al. 1987). As Berdnt and her colleagues pointed out, the probabilities they calculated were independent from context and therefore, as the authors admitted: "[they] provide a conservative estimate of the extent to which particular letters and letter clusters are pronounced as particular phonemes in English, but they provide no information about the rules responsible for the derivation of these correspondences" (Berdnt et al. 1987, 1). Teachable rules are, after all, what EFL teachers would demand.

More recent research (Stanback 1992; Treiman et al. 1995; Kessler and Treiman 2001, 2003) incorporates a remarkable sophistication of statistical procedures, and continues to add to a perceived need to convince teachers—particularly teachers of L1 English— that English spelling is much more consistent than it seems. However, such calculations of consistency are often derived from a consideration of monosyllabic words. Polysyllabic words are conveniently avoided because they "raise difficult questions of where syllable boundaries lie," and because they "have much higher proportions of foreign, Latinate, and technical words" (Kessler and Treiman 2001, 593).

After decades of increasingly sophisticated research, there can be no doubt today that the English grapho-phonemic system must be consistent. Paradoxically, most specialist also agree that it *seems* chaotic. The famous 80% consistency of the system seems to have worked as an excuse for not teaching pronunciation rules: Since the system is consistent, simple exposure to it will in the long run build up the necessary competence in our EFL students. Besides, a total 20% of recalcitrant words, apparently hiding just about anywhere within the system, may be quite dissuasive. In relation to exceptions and teaching, Axel Wijk wrote: "Since a large number of the irregular spellings are found among the commonest words in the language, it is obvious that there would not be much sense in foreigners making any systematic study of the rules of pronunciation when they begin their study of the language" (Wijk 1966, 11). His suggestion for incrementing the EFL students' competence is the one that has been strictly followed by most

textbooks: “to learn the pronunciation of each new word that they come across, by itself, […] without much reference to any rules for the pronunciation of the various letters or combinations of letters of which the words are made up” (Wijk 1966, 11). From such perspective, there is little hope that grapho-phonemic correspondences may ever become part of the EFL training program. As far as I know, Wijk did not fully substantiate his assertion about irregularity at the basic levels; a notion, on the other hand, that contradicts Kessler and Treiman’s findings (2001). One way or another, we are told by relevant orthographists about the overall consistency of large corpora, but we do not know much about the regularity or frequency[2] of most rules, or about rule-regularity and rule-frequency within particular groups of words, or within smaller samples like the glossary of a handbook for beginners. If we managed to shed some light over such issues, we might be able to offer a more enticing panorama to our EFL teachers. We might, for example, base our choice of a particular reading material on its grapho-phonemic characteristics and the kind or rules that are present in it. A specifically EFL-way of looking at pronunciation rules is required. This would imply the use of a conceptual frame that could help us discern between rules and organize them into a pedagogical hierarchy.

2. Basic Rules for the Interpretation of Stressed Unigraphs

A pronunciation rule is a hypothesis about the predictive properties of a particular graphemic environment in relation to the phonemic value of the letters affected by it.[3] A very well-known pronunciation rule establishes, for example, that a stressed unigraph, in an oxytone structure, followed by a single consonant, and closed by a silent <e>, is to be pronounced with the long version of that particular unigraph (Venezky 1970, 104; Bozman 1988,

[2] When characterizing a rule in terms of ‘frequency’, I will be considering the number of words within this corpus to which a particular rule is applicable; it implies an estimate of how frequently our students will have to apply it. The expressions PR1/2.1/2.2/etc.—to be discussed later—will also be used to classify word-types grapho-phonemically. A particular word-type has a high ‘presence’—for example PR2.2(*matter*)— when the corpus contains many instances of it.

[3] Notice that in talking about *prediction* we are already siding with the EFL student, who lacks a priori grapho-phonemic competence in English, and who is constantly facing new words in the written format whose pronunciation is not to be identified, as in the case of natives, but rather guessed. Instead of prediction, most experts outside EFL talk about phonological decoding or even phonological translation (Aro and Wimmer 2003).

13). [4] The long version coincides with the name of the unigraph, and this rule—often expressed as VC<e># or similar—allows us to predict values /eɪ, iː, aɪ, oʊ, (j)uː/ for the stressed vowels in *pale*, *scene*, *Mike*, *hose*, and *rude*, respectively, and in any other words with the same structure. A rule that, incidentally, fails to predict the pronunciation of *have*, *allege*, *police* or *move*.

If we are to generate useful knowledge about pronunciation rules, the first thing to do is to isolate and describe as strictly as possible a comprehensive set of such rules. The panorama is more complex than that registered by Olshtein (2001) and Ediger (2001), mentioned above, and yet, it is important to reduce the detailed complexity described in Venezky (1970) or Cummings (1988) to manageable terms. It is not easy to navigate through all the different pronunciation rules on offer. We have rules for the interpretation of consonant-letters—like <c> before <o, u, a>—and for the interpretation of vowel-letters; rules for the location of primary, secondary or tertiary stresses; rules for the interpretation of stressed and unstressed syllables; etc. Our first task is to isolate a particular set of rules, governing a particular aspect of orthography, so that the minimum amount of them covers a maximum of occurrences. I will therefore focus on rules that assist the interpretation of single vowel letters—unigraphs—located at the stressed syllable of any English word. For the sake of pedagogical efficiency, we need to isolate the kind of rules that could be applied to as many English words as possible—instead of only to those that have <c> or <y>, for example—provided that they are sufficiently effective in predicting pronunciation. Table 1 represents the set of what I call the ten basic general-systemic rules for stressed unigraphs in American English; a comprehensive set of rules that can be applied to all the words that have a unigraph in their stressed syllable, without exception.

A *general-systemic* rule is one that is capable of predicting the value of at least four of the five vocalic unigraphs. The above mentioned VC<e># rule belongs to this category, but there are pronunciation rules—I will call them *domain-specific*—that are only applicable to one or two unigraphs, or domains; Olshtain mentions one such rule (Olshtain 2001, 210). In my own

[4] Words stressed on the last syllable constitute oxytone structures—like *bet*, *attack*, *introspect*, etc. A paroxytone structure has the primary stress on the penultimate syllable—like *manner*, *indulgent*, *condemnation*, etc. Words stressed on the antepenultimate syllable or before are proparoxytones—*enemy*, *indiscriminate*, *ceremony*, etc. Although experts do not usually make this distinction, I believe there is much to gain in terms of discrimination power by incorporating them, as I will show.

terms, adjacent post-nuclear[5] <ll> has predicting power in relation to stressed unigraph <a> but not in relation to <e, i, u>. Notice that *bet* and *bell*, *bit* and *bill*, *dust* and *dull* are subject to the same general-systemic rule—PR2.1—while the contrasts *bat*/*ball*, *tan*/*tall*, *mad*/*mall*, etc. point to adjacent post-nuclear <ll> as a domain-specific context affecting <a> quite regularly, and <o> in a less regular way—*rot*/*roll*.

Table 1. English General Systemic Rules

	PR-1	<a>, /ɑː/	<e>, /iː/	<i>/<y>, /aɪ/	<o>, /oʊ/	<u>, /(j)uː/
	…V#	*spa*	*me*	*I, my*	*go*	*gnu*
	PR-2.1	<a>, /æ/	<e>, /e/	<i>/<y>, /ɪ/	<o>, /ɑː/	<u>, /ʌ/
a.	…VC#	*cat*	*bet*	*pin*	*hot*	*mud*
b.	…VCC#	*back*	*belt*	*myth*	*block*	*lung*
c.	…VCCC#	*match*	*fetch*	*bitch*	*blotch*	*gulch*
d.	…VCC<e>#	*lapse*	*ledge*	*bridge*	*lodge*	*budge*
	PR-2.2	<a>, /æ/	<e>, /e/	<i>/<y>, /ɪ/	<o>, /ɑː/	<u>, /ʌ/
a.	…VCCv…	*happen*	*question*	*issue*	*office*	*number*
b.	…VCCC…	*android*	*entry*	*hindrance*	*cockle*	*buckle*
c.	…V<ɪr>v…	*narrow*	*merry*	*mirror*	NA	NA
	PR-3.1	<a>, /eɪ/	<e>, /iː/	<i>/<y>, /aɪ/	<o>, /oʊ/	<u>, /(j)uː/
	…VC<e>#	*make*	*Pete*	*time*	*home*	*include*
	PR-3.2	<a>, /eɪ/	<e>, /iː/	<i>/<y>, /aɪ/	<o>, /oʊ/	<u>, /(j)uː/
a.	…VCv…	*paper*	*Peter*	*final*	*open*	*human*
b.	…VC<r>v…	*patriot*	*secret*	*micron*	*copra*	*nutrient*
c.	…VC<l>v…	*able*	NA	*title*	*noble*	*bugle*
	PR-4.1	<a>, /ɑː/	<e>, /ɜː/	<i>/<y>, /ɜː/	<o>, /ɔː/	<u>, /ɜː/
a.	…V<r>#	*car*	*her*	*fir*	*for*	*fur*
b.	…V<r>C#	*part*	*term*	*girl*	*report*	*return*
c.	…V<r>CC#	*arch*	*perch*	*birth*	*scorch*	*burnt*
d.	…V<r>C<e>#	*large*	*serve*	*dirge*	*horse*	*nurse*
	PR-4.2	<a>, /ɑː/	<e>, /ɜː/	<i>/<y>, /ɜː/	<o>, /ɔː/	<u>, /ɜː/
a.	…V<r>Cv…	*party*	*service*	*virtual*	*important*	*surface*
b.	…V<r>CC…	*partner*	*interpret*	*myrtle*	*northern*	*purchase*
	PR-5.1	<a>, /er/	<e>, /ɪr/	<i>/<y>, /aɪ/	<o>, /ɔː/	<u>, /ʊr/
	VC<re>#	*care*	*here*	*fire*	*more*	*ensure*
	PR-5.2	<a>, /er/	<e>, /ɪr/	<i>/<y>, /aɪ/	<o>, /ɔː/	<u>, /ʊr/
	V<r>v…	*parent*	*period*	*virus*	*story*	*jury*
	PR-6 Proparoxytones	<a>, /æ/	<e>, /e/	<i>/<y>, /ɪ/	<o>, /ɑː/	<u>, /ʌ/
a.	…VCv…	*family*	*president*	*significant*	*policy*	NA
b.	…V<r>v…	*charity*	*American*	*miracle*	NA	NA
c.	….VC<r>v…	*African*	*integrity*	*fibrillate*	*Socrates*	NA

PR: Pronunciation Rule; NA: Not Aplicable

[5] In the English system there are also adjacent pre-nuclear grapho-phonemic contexts—like onset <w-, qu-> before <a, o>, as in *and*/*wand*, *horse*/*worse*— and distant post-nuclear contexts—all the suffixes explored by Dickerson (1984)—that tend to determine both where the primary stress is located and how the nuclear vowel is to be pronounced.

The pronunciation rules in table 1 constitute a development of Bozman's general rules for the spelling of long and short pronunciations (Bozman 1988, 14-15). All of them, without exception, will also be found in the works of Wijk, Venezky and Cummings. There are, however, a few aspects where I depart from the traditional positions. Apart from the technical terms and descriptions used above, there is the subdivision into oxytone, paroxytone and proparoxytone structures, the value that PR1 assigns to <a>, and the boxes marked with NA for domains where a particular rule is *Not Applicable*. A word like *al**u**minum*, for example, which would fit into the PR6 type will be interpreted according to PR3.2 because PR6 is not applicable to the <u> domain (Bozman 1988, 48).

Table 1 represents basic grapho-phonemic competence in General American. An RP version would be almost identical; in standard British English the value for <o> in PR2 and 6 would be /ɒ/, and words like *sorry*, *current*, and *oracle* would fill the NA boxes in PR2.2 and 6. On this occasion, our calculations of rule reliability and presence have been made for American English, but our guess is that with slightly different measures in some marginal cases, the overall results would be quite similar for RP.

3. Research Questions

The general idea is that EFL teachers could make use of table 1, and teach it either completely or partially to their students. Before that, however, it is necessary to put these contents to the test and provide an answer to the following research questions:

- How exhaustive is this particular set of rules in relation with the interpretation of English stressed unigraphs? Would we need any/many more rules?
- How do these rules differ in terms of frequency and regularity? Are there any word-types that feature as particularly (in)frequent and/or (ir)regular?
- Are there areas of special difficulty that EFL teachers might consider avoiding?
- Are there significant differences at the different levels (beginners, intermediate, etc.)? How could the teaching of rules be distributed along the different levels?

4. METHOD

4.1 Data

Using Excel 2010 and SPSS 20, I have checked the reliability and frequency of each of the ten general-systemic rules against M. Davies' 5000 wordlist containing the most frequent words in American English (Davies 2015). The corpus has been conveniently reduced by filtering away words with stressed digraphs and trigraphs, which are not covered by the rules whose reliability we are trying to assess. The original wordlist also contains a number of repetitions. The word *work*, for instance, appears in rank 117 as verb, and then again in rank 199 as a noun. I have computed this and all repeated items only once whenever I found exactly the same grapho-phonemic correspondences. Where pronunciation changes along with function, the items have been treated and computed separately—*protest* for example, has been computed as a 2.1 type with stressed <e> when functioning as a verb and as a 3.2 type with stressed <o> when functioning as a noun.

Table 1. Reduction of the wordlist

Total words	5000	
Special words, acronyms, repetitions, non-computed derivatives, compounds, etc.	1066	
Computed words (ref.)	**3934**	**100%**
Words with stressed unigraph	3005	76%
Words with digraphs or trigraphs	929	24%
Regular stressed unigraph words	2514	64%
Irregular stressed unigraph words	491	12%

Special words like *mm-hmm* (rank 2966) and acronyms have also been discarded. Compound words have been separated into their components and reintroduced in the database. The grapho-phonemics of, for example, *understand*, comprises that of *under*—a regular PR2.2-type by itself—and that of *stand*—a regular PR2.1-type. There are few compounds that break this rule. In fact, the corpus only contains three cases: *gentleman*, *freshman* and *businessman*; the words *gentle*, *fresh* and *business* are already present in the corpus. In most cases, in fact, reintroduced items were found to constitute repetitions not to be computed.

As a result of this process of filtering away the irrelevant items for the sake of maximally reliable results, we have obtained 3005 frequent English words against which to test the ten basic rules presented above.

4.2. Procedure

For each of the words contained in the final list, I have registered information concerning the grapheme <a, e, i, y, o, u> to be found in its

stressed syllable, the applicable pronunciation rule in each case—PR1, PR2.1, etc.—and a yes/no tag depending on whether the applicable rule works or not. For the process of tagging each item with their respective PR types I have taken as a reference the American English transcriptions contained in *Longman's Pronunciation Dictionary*, edited by J. C. Wells (1990). When more than one possible pronunciation was offered, I have only considered the first and most standardized one. In the case of function words with strong and weak forms the strong form has been considered.

Table 2. Tagging sample

Frecuency (rank)	Item	Unigraph	Type	Reliability
5	**a**	a	PR1	n
6	**in**	i	PR2.1	y
7	**to**	o	PR1	n
8	**have**	a	PR3.1	n
10	**it**	i	PR2.1	y
11	**I**	i	PR1	y
12	**that**	a	PR2.1	y
13	**for**	o	PR4.1	y

The tagging of derivatives has been quite challenging. It seems obvious that once you have tagged and computed a word like *color* as a PR3.2 type, the word *colorful*, also registered in the original list, looks very much like a repetition. If treated as such, the word *colorful* should not be computed, and all suffixed words should then be subjected to the same consideration. However, not all suffixed words lend themselves so nicely to that treatment. Many common words like *protective*, *defendant* or *evidence* would not be computed with such procedure, and the corpus would have been drastically impoverished.

My solution here has been rather practical. I have discarded as repetitions those derivatives consisting of a root already existing in the database plus affixes –ly, –ing, –ed, –ful, –ness, –ment, –less, –ship, –wise, –hood, –th (in ordinal numbers), un–, and plural –s. These common and frequent derivatives could be easily treated in grapho-phonemic training through a simple elimination strategy: When faced with a word like *gathering*, students could be instructed to eliminate the –ing ending and consider the predictability of *gather*.

When a derivative with any of these affixes was found whose root was missing from the database, it is the root itself that has been tagged and computed. So, the word *frankly* has been tagged and computed as if it was *frank*—i.e. as PR2.1 rather than PR2.2—because *frank* itself was not in the database. The rest of the derivatives, with affixes other than those listed

above, have been computed according to their corresponding type as whole words: *director* as a PR2.2 type, *performance* as PR4.2, *racism* as PR3.2, etc.

Still, a further complication comes up when dealing with derivatives. A word like *additional*, for example, would be categorized as a regular PR6 type. However, regularity here might not be due to the PR6 context, but to the fact that it derives from *addition*, which actually breaks PR3.2. Inversely, a word like *favorable* breaks PR6, but it does so in order to preserve the recognition of *favor*, a regular 3.2 type. It is not convenient then to treat *additional* as a confirmation of PR6, nor to treat *favorable* as an exception to that same rule, since both words are actually subjected to yet another rule: a principle of preservation of etymological traceability (Venezky 1970, 120; Author year, p.).

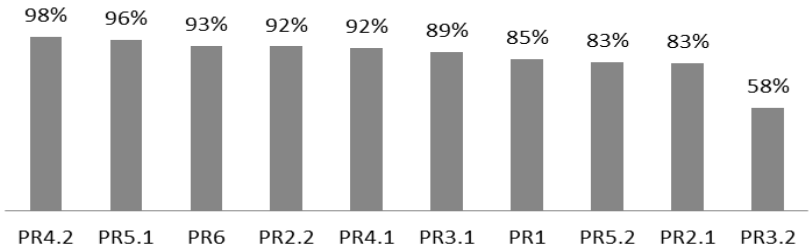
Another practical solution is in order: Derivatives which confirm their rules—as *additional* confirms PR6—have been computed only when their roots also confirm their own rules; derivatives which break their rules—as *favorable* breaks PR6—have been computed only if also their roots break their own rules; and they have not been computed otherwise. In this way, a word like *global* has been computed as confirming PR3.2 because *globe* also confirms PR3.1; *fully* has been computed as breaking PR2.2 because *full* also breaks PR2.1. But neither *additional* nor *favorable* have been computed as either confirming or breaking PR6. In this way, I did not boost the reliability of PR6 with words like *additional*, *clinical*, *columnist* or *developer*; nor did I undermine it with words like *agency*, *behavioral*, *educational* or *frequency*. All of them owe the phonemic value of the stressed vowel to their roots. In most cases, the respective roots—*addition*, *clinic*, *behavior*, *education*, etc.—were in the list and were tagged and computed accordingly, the only exceptions being *theoretical*, *tropical* and *practitioner*. These last three have not been computed in any way; they would have, however, added up to the reliability of PR6.

For operational purposes, I have further divided the words into five levels. Level 1 contains the most frequent 601 words, which would be ideally taught to beginners, with successive levels incrementing 601 words each until reaching the final amount of 3005 words at the end of Level 5. This procedure has allowed me to consider rule regularity and presence from a progressive point of view. Teachers of EFL for beginners might be more interested in rule regularity and frequency at Levels 1 and 2 than at later levels.

5. RESULTS

A total of 2514 items have been found to be predictable from the application of the rules. This rendered a general predictability of 83.6%, out of the 3005 stressed-unigraph words. The mean regularity of pronunciation rules is 87%—derived from the values presented in figure 1. As we can see in this figure, the only value that stands out is that of PR3.2, which registered significantly low regularity.

Figure 1. Rule regularity, percentage. Ranking display



When considering regularity across levels, results show an average of 81%. Rules are less regular at Level 1 (76%), with a clear tendency to increase their regularity as the amount of computed words increases. At Level 2, with 1202 words, regularity is 80%. This percentage increases to 82% at Level 3 (1803 words), and rises above 83% at Levels 4 and 5 (2404 and 3005 words respectively). The more words we include the more regular the set becomes. As we can see in figure 2, this seems to be the case for all rules, except PR5.2. The five bars for each rule, in figure 2, represent levels 1–5, from left to right.

Figure 2. Rule regularity across levels

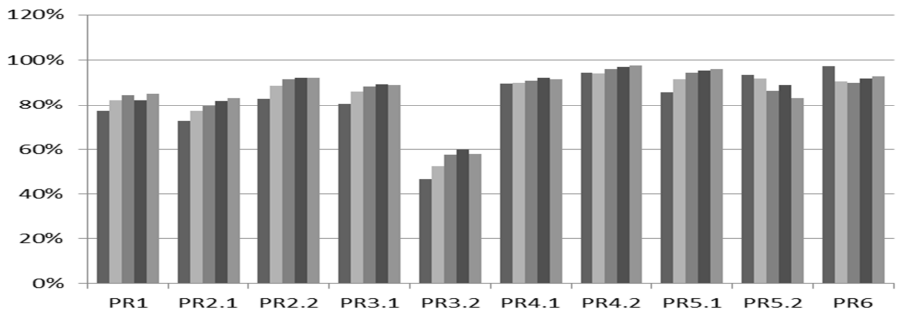


Figure 3 represents the presence of each of the rules within the sample. Frequency data shows more dispersion than regularity data. The average presence was 10%, with a standard deviation of 9%. Values beyond 19% stand out as particularly relevant—PR2.2 and PR2.1. Values around 1%—

PR5.1, PR1—point to a rather anecdotal presence. Put together, 86% of the processed words were found to belong to PR2, 3 and 6.

Figure 3. Frequency of rules, percentage. Ranking display

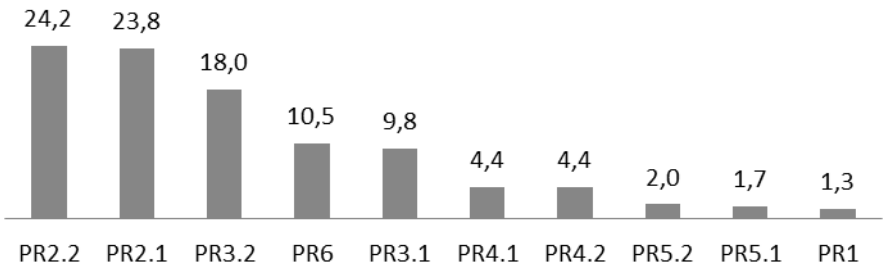
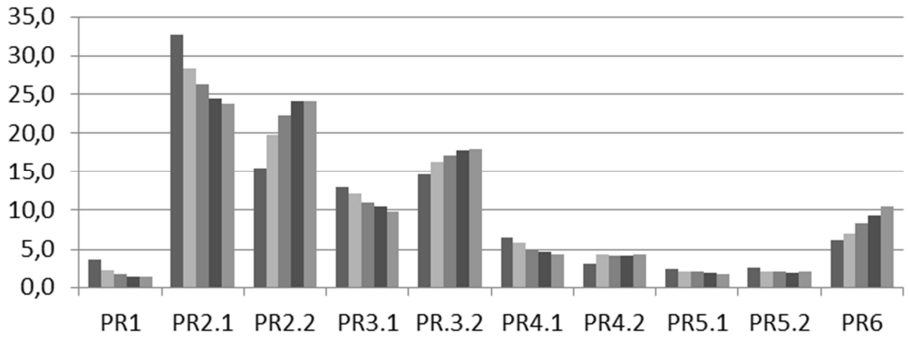


Figure 4 represents the presence of word-types at the different levels. For each rule, the five bars indicate from left to right the percentage values of levels 1–5. We have isolated word-types whose presence decreases, like PR1, PR2.1, PR3.1, PR4.1; word-types whose presence increases, like PR2.2, PR3.2 and PR6; and word-types, like PR4.2, PR5.1 and PR5.2 whose presence does not either increase or decrease in any significant way.

Figure 4. Frequency of rules across 5 levels



If we compare figures 1 and 3, we get confronted with a number of relevant facts: Some rules are very regular, but have very little presence—e.g. PR5.1; one rule, PR3.2, has a rather high frequency that matches a significantly low regularity; and there are some, like PR2.1, PR2.2 and PR.6 that feature both as frequent and regular.

A final set of relevant results has to do with the regularity of the function words contained in the list. The results for both Levels show similarities. At Level 1, only articles (92%), prepositions (90%) and conjunctions (80%) show significant regularity. Similarly, at Level 5, the most regular are, again, articles (92%), prepositions (84%) and conjunctions (80%), as well as interjections (88%). At the initial Level, as far as function words are

concerned, PR2.1 (80%), PR4.1 (100%), PR4.2 (100%), PR5.2 (100%), PR6 (100%) are regular. At the final Level, these rules remain similarly regular, except for PR5.2 that drops to a 50%. At this final Level also PR2.2 (80%) reaches regularity. At both Levels, PR3.2 registers significantly low regularity: 30% at Level 1 and 17% at Level 5. Despite these similarities, the average regularity of function words is significantly lower at Level 1 (69%) than at Level 5 (77%). These low percentages have to do with the irregularity, at both Levels, of pronouns, numbers and demonstratives.

5. Discussion

Concerning the possible *teachability* of our ten basic pronunciation rules, there are at least two fundamental variables to consider: regularity and frequency—or presence. Of these, regularity is clearly a sine qua non: Any rule that had (almost) as many exceptions as regulated cases should no longer be considered a rule. Frequency, on the other hand, raises issues of pragmatism: We might be interested in teaching a rule insofar as our students are likely to meet many chances to apply it. An infrequent rule, however, would still be a rule. The teachable character of four of our ten basic rules seems unquestionable inasmuch as they are both regular and frequent: PR2.1, 2.2, 3.1 and 6, regulating words like *not*, *letter*, *fine* and *animal*.

The consideration of regularity is by no means a simple matter. It is not easy to determine how many exceptions a given rule could have before it turns into a case of anecdotal regularity. It is probably EFL teachers confronting the results of this study who must decide whether a given regularity threshold is acceptable for them or not. Should this threshold be set at 90%, only half the rules—PR2.2, 3.1, 4.1, 4.2, 5.1 and 6—would turn out to be teachable. If the threshold is lowered to 80%, only PR3.2, regulating words like *nation*, *even*, *final*, *over* or *student* would be left out.

Quite visibly, the greatest challenge for teaching our ten adjacent post-nuclear general-systemic indicators is posed by PR3.2. A reliability of 58% for a set of 540 items actually allows us to question PR3.2 as a general-systemic rule, despite its somewhat larger reliability in the <a, u> domains—see Appendix. So, while it is rather clear that a stressed unigraph followed by CC is predictable, the same unigraph followed by Cv is not predictable to a comparable extent. All we can say, at best, is that the long version—PR3 phonemic value—in these cases is slightly more frequent than the short version—or PR-2 phonemic values. That is, words like *lemur* are a somewhat more frequent than words like *lemon*.

This has technical implications for teaching. While revealing pairs like *mat-mate*, *pet-Pete*, *pin-pine*, *cod-code* or *cut-cute* could be used to take

pedagogical advantage of the contrast between PR2.1 and PR3.1, the same procedure would be misguiding in the case of PR2.2 and PR3.2; not because pairs like *matter-mater* or *saddest-sadist* cannot be found, but because PR3.2 is not fully reliable.

Still, a reliability of 58% might be worth taking into account somehow, and before rejecting any application of the pattern, one should see if the group of exceptions might possibly be reduced through the application of domain specific rules.

With a reliability of 30%, PR3.2 for the <i> domain stands out as the most unreliable sub-rule. However, many of the 74 words that do not follow PR3.2 here actually follow other easy and reliable domain specific rules. For example, we know that stressed <i> retains pronunciation /ɪ/ despite a VCv environment when it fits the description VCvv, as in *cond**i**tion*, *civ**i**lian*, *cont**i**nue*, *eff**i**cient*, *susp**i**cious*, *w**i**dow*, etc. (Wijk 1966, 20). A total of 41 of the 74 supposedly irregular words are actually subjected to this domain specific rule.

Table 3. Pronunciation Rule 3.2

PR-3.2				
Domain	**It.**	**Reg.**	**Irreg.**	**Ratio**
<a>	193	140	53	72%
<e>	69	31	38	44%
<i>	106	32	74	30%
<y>	3	2	1	66%
<o>	100	51	49	51%
<u>	69	57	12	82%
Totals→	**540**	**313**	**227**	**58%**

Furthermore, 12 of the 74 irregular words are actually predicted by known distant post-nuclear contexts—such as suffixes *–ish*, *–ic*, or *–it*—that tend to fix stress on the previous syllable and to predict the short value of the unigraph (Bozman 1988, 48). Words like *clinic*, *diminish* or *explicit* break PR3.2, but do so in order to follow an overriding rule. Other known rules finally reduce the 74 supposedly irregular words to some 6 words: *casino*, *city*, *consider*, *diplomat*, *prison*, and *sibling*. Not all groups of exceptions will allow such consistent and intensive reduction; however, there can be no doubt that the total reliability of 83.6% would increase if domain specific rules where included.

Another challenge concerning regularity is the confirmation of Wjik's claim that irregularities tend to abound within the most basic vocabulary, (Wijk 1966, 11). A portion of this basic vocabulary is constituted by

function words, which have, as we have seen, an average regularity of 69% at the initial Level. In general, the average regularity of rules among the first 601 words falls below a threshold of 80%. However, on a more detailed inspection we see that PR2.2, 4.1, 4.2, 5.1, 5.2, and 6 obtain a regularity well above 80% even within this first level—table 4.

Table 4. Regularity at Level 1

	Level 1		Level 1
PR1	77%	**PR4.1**	90%
PR2.1	73%	**PR4.2**	94%
PR2.2	83%	**PR5.1**	86%
PR3.1	81%	**PR5.2**	93%
PR3.2	47%	**PR6**	97%

The strong form of some of the function words—articles, prepositions and conjunctions— has proved to be rather regular at both initial and final Levels. Although, function words are most frequently pronounced with their weak form, common items like *there*, *of*, *were*, *to*, *whom*, etc. may add to the perception of a chaotic system. In fact, Wijk's perception of irregularity among common words, functional and lexical, is confirmed by the results. However, his subsequent conclusion that pronunciation rules are not to be taught at initial levels should be reconsidered. With the enhanced discrimination that our procedure permits, we see that there is much within basic vocabulary that remains regular and teachable at Level 1.

In relation to rule presence, stressed-unigraph words are much more frequent in English than stressed-diagraph and stressed-trigraph words. Up to seven out of ten words the students encounter during their English training is bound to be the kind to which one of our ten basic rules is applicable. Of these, however, PR1, 4 and 5 are alarmingly infrequent. In the case of PR1, for example, there are merely 40 words, to be taught over five levels. One actually wonders whether PR1—referred by others as the CV rule—actually exists at all. If we move beyond the 5000 wordlist, we would certainly find more cases, but they would have to be considered relatively infrequent, and their usefulness for non-advanced EFL students is rendered debatable. Furthermore, the reliability of PR1 is only beyond 80% in the <e, y> domains, which instructors might choose to teach as domain-specific rules, if at all—see Appendix.

The situation with PR5.1 and 5.2 is very similar. Frequency here seems insufficient, and if we consider it as a fundamental condition for *teachability*, the convenience of investing effort in the teaching and learning of these rules is questionable, to say the least. But they actually constitute very reliable pronunciation rules. If an EFL instructor decides that these

rules are worth teaching, the best strategy would probably not be waiting for these words to come up, but rather to confront their teaching explicitly, and be ready to work with not so frequent words or even pseudo-words. The same could be said about the PR4 type, where presence, though still limited, is larger than in PR5 and reliability the highest of the entire set.

An interesting aspect that emerges upon analyzing our results—figure 4—is that there is an inverse correlation between presence and level in the case of oxytones, and a direct correlation in the case of paroxytones and proparoxytones. Words like *can*, *cane*, *car* or *care* are in general less frequent than words like *manner*, *vapor*, *party* or *Mary*; but they are more frequent at Level 1. This has to do, in part, with the fact that the strong forms of many frequent functional words are mostly oxytones. Still, figure 4 suggests a possible order in the teaching of pronunciation rules: PR1, 2.1, 3.1, 4.1 and 5.1 might be taught at the first levels. Chances for practicing PR2.2, 3.2, 4.2, and 5.2, by no means scarce at Level 1, will only increase in the following levels.

6. Conclusions and Further Research

The first conclusion that we can draw from our study is that there are, in fact, a reduced number of pronunciation rules that may help our EFL students to interpret the phonemic value of the stressed unigraph in most cases. We should not underestimate the relevance of this fact. Though books on English orthography are usually lengthy and complex, a small set of ten basic rules has proved to be extremely exhaustive. We must assume that the reason why these rules are not usually taught in EFL courses has little to do with their reliability or their applicability. Further domain specific rules would increase our perception of consistency, but they would actually cover a much smaller number of cases.

Orthographic structures like those present in *manner* (PR2.2), *pet* (PR2.1), *enemy* (PR6) and *cone* (PR3.1) are, in this particular order, the most frequent in everyday American English, and they regularly allow the prediction of a particular vowel phoneme in the stressed syllable. On the other hand, orthographic structures like those present in *rely* (PR1), *car* (PR4.1), *person* (PR4.2), *mire* (PR5.1) and *hero* (PR5.2) are much less frequent, but they tend to be even more reliable in the prediction of the phonemic value of any unigraph in their stressed syllable. While research into the best ways of teaching pronunciation rules is still pending, it seems reasonable to think that for rules that have limited presence but high regularity, explicit teaching would be adequate.

Word-type PR3.2 is extremely problematic in grapho-phonemic terms. It stands as the third most frequent structure in English; however, the

predictability of the value of its stressed unigraph is only slightly above 50%: We have *vapor* but *manor*, *Peter* but *second*, *icon* but *idiot*, *odor* but *body*, *Cuban* but *punish*. One of the most frequent structures in English is also the most irregular. Any EFL instructor with a mind to teach pronunciation rules should probably either discard PR3.2 or, preferably, be ready to give it a special treatment—making room, perhaps, for some domain specific rules.

Words like *can*, *cane*, *car* and *care*—oxytone structures—tend to be relatively frequent among the first 601 items of our list. This suggests that the oxytone rules PR1, 2.1, 3.1, 4.1 and 5.1 might better be taught at the initial levels. Paroxytone and proparoxytone rules might be quite profitably dealt with at later stages.

The present study has some limitations. Although some parts of the processing have been made automatically, word-type tagging and information on rule regularity had to be manually completed. This has made it impossible to work with a larger corpus, which would have been very desirable. Nevertheless, the amount of processed words, having been selected with a view to EFL applicability, is neither insufficient for consolidating reliable knowledge, nor particularly small when compared with previous research. Venezky (1970) dealt with 20,000 words and Stanback with 17,602; but Kessler and Treiman (2001, 2003), also using both automated and manual procedures, made important contributions by processing smaller carefully selected corpuses (1329 and 914). Relatively small corpuses, compiled after strict criteria, may lead to strong conclusions about specific aspects of English orthography—monosyllabic words, rimes, vocalic unigraphs, etc. On the other hand, careful selection is not necessarily incompatible with a larger corpus. I am currently working on the design of scripts and automatic protocols that will hopefully increase automatic processing and allow me to explore the teachability of a larger battery of rules (domain specific rules, overriding principles, unstressed syllables, digraphs, etc.) within a much larger corpus.

For the time being, we may hold to the conclusion that the ten basic post-nuclear general-systemic rules constitute, as a whole, teachable material; at least in terms of frequency and regularity. However, ten basic rules might still be a few rules too many for instructors who are legitimately interested in the development of effective communicative skills rather than in the unveiling of peculiar fine-grained aspects of the English language to their students. Some of the research reviewed above points to possible unexpected advantages of grapho-phonemic training. This is a matter that requires further study. Would grapho-phonemic training aid vocabulary memorization and recall? Would it build up confidence in EFL students?

Would it have a positive impact on the assimilation of the English phonological system? Would a development of grapho-phonemic competence correlate with an increase in listening skills? Would it bring about an improvement in oral skills? These and other related questions must be left for future exploration.

WORKS CITED

Aro, Mikko and Heinz Wimmer. 2003. "Learning to Read: English in Comparison to Six More Regular Orthographies." ***Applied Psycholinguistics*** 24: 621-635.

[AUTHOR, A. reference]

Berndt, Rita Sloane, James A. Reggia and Charlotte C. Mitchum. 1987. "Empirically Derived Probabilities for Grapheme-to-Phoneme Correspondences in English." ***Behavior Research Methods, Instruments, & Computers*** 19 (1): 1-9.

Bozman, Timothy. 1988. ***Sound Barriers. A Practice Book for Spanish Students of English Phonetics***. Zaragoza: Universidad de Zaragoza.

Brown, H. Douglas. 1970. "Categories of spelling difficulty in speakers of English as a first and second language." ***Journal of Verbal Learning and Verbal Behavior*** 9: 232–236.

Celce-Murcia, Marianne, Donna M. Brinton and Janet M. Goodwin. 1996. ***Teaching pronunciation: A reference for teachers of English to Speakers of Other Languages***. Cambridge, UK: Cambridge UP.

Cheung, Him, Hsuan-Chih Chen, Chun Yip Lai, On Chi Wong and Melanie Hills. 2001. "The Development of Phonological Awareness: Effects of Spoken Language Experience and Orthography." ***Cognition*** 81: 227-241.

Cummings, Donald Wayne. 1988. ***American English Spelling. An Informal Description***. London: The John Hopkins University Press.

Daelemans, Walter M. P. and Antal P. J. van den Bosch. 1996. "Language-independent and data-oriented grapheme-to-phoneme conversion." In ***Progress in Speech Synthesis***, edited by Jan P. H. Van Santen, Richard Sproat, Joseph Olive and Julia Hirschberg. New York: Springer Verlag: 77-90.

Davies, Mark. ***Word frequency data. Corpus of Contemporary American English***. http://www.wordfrequency.info/. Downloaded on June 2015.

Dickerson, Wayne B. 1984. "The Role of Formal Rules in Pronunciation." In ***On TESOL'83: The Question of Control***, edited by Jean Handscombe, Richard A. Orem and Barry P. Taylor, 135-148. Alexandria, VA: TESOL.

—.1987. “Orthography as a Pronunciation Resource.” ***Word Englishes***, 6 (1): 11-20.

—.1994. “Empowering Students with Predictive Skills.” In ***Pronunciation Pedagogy and Theory: New Directions, New Views***, edited by Joan Morley, 17-35. Alexandria, VA: TESOL.

Doignon-Camus, Nadege and Daniel Zagar. 2014. “The Syllabic Bridge: The First Steps in Learning Spelling to Sound Correspondences.” ***Journal of Child Language*** 41 (5): 1147-1165.

Ediger, Anne M. 2001. “Teaching Children Literacy Skills in a Second Language.” In ***Teaching English as a Second or Foreign Language***, edited by Marianne Celce-Murcia, 153-170. London: Heinle & Heinle.

Escudero, Paola, Rachel Hayes-Harb and Holger Mitterer. 2008. “Novel Second-Language Words and Asymmetric Lexical Access.” ***Journal of Phonetics*** 36: 345-360.

Frederiksen, John R. and Judith F. Kroll. 1976. “Spelling and sound: Approaches to the internal lexicon.” ***Journal of Experimental Psychology: Human Perception and Performance*** 2 (3): 361-379.

Hanna, Paul R., Jean S. Hanna, Richard E. Hodges and Edwin H. Rudorf. 1966. ***Phoneme-grapheme correspondences as cues to spelling improvement***. Washington D. C.: U. S. Government Printing Office.

Kessler, Bret and Rebecca Treiman. 2001. “Relationships between Sounds and Letters in English Monosyllables” ***Journal of Memory and Language*** 44: 592-617.

—.2003. “Is English Spelling Chaotic? Misconceptions Concerning its Irregularity.” ***Reading Psychology*** 24: 267-289.

Kreidler, Charles W. 1972. “Teaching English Spelling and Pronunciation.” ***TESOL Quarterly***, 6(1): 3-12.

Martínez Martínez, Angélica María. 2011. “Explicit and Differentiated Phonics Instruction as a Tool to Improve Literacy Skills for Children Learning English as a Foreign Language.” ***Gist. Education and Learning Research Journal*** 5: 25-49.

Lan, Yizhou and Mengjie Wu. 2013. “Application of Form-Focused Instruction in English Pronunciation: Examples from Mandarin Learners.” ***Creative Education*** 4 (9): 29-34.

Long, Michael H. and Peter Robinson. 1998. “Focus on form: Theory, research and practice.” In ***Focus on Form in Classroom Second Language Acquisition***, edited by Catherine Doughty and Jessica Williams, 15-41. Cambridge: Cambridge UP.

Olshtain, Elite. 2001. “Functional Tasks for Mastering the Mechanics of Writing and Going Just Beyond.” In ***Teaching English as a Second***

or Foreign Language, edited by Marianne Celce-Murcia, 207-218. London: Heinle & Heinle.

Rangriz, Samaneh, Amin Marban. 2015. "The Effect of Letter-Sound Correspondence Instruction on Iranian EFL Learners' English Pronunciation Improvement." ***Journal of Applied Linguistics and Language Research*** 2 (7): 36-44.

Schwartz, Ana I., Judith F. Kroll and Michele Diaz. 2007. "Reading Words in Spanish and English: Mapping Orthography to Phonology in Two Languages." ***Language and Cognitive Processes***, 22(1): 106-129.

Stanback, Margaret L. 1992. "Syllable and rime patterns for teaching reading: Analysis of a frequency-based vocabulary of 17,602 words." ***Annals of Dyslexia***, 42, 196–221

Treiman, Rebecca, John Mullenix, E. Daylene Richmond-Welty and Ranka Bijeljac-Babic. 1995. "The Special Role of Rimes in the Description, Use, and Acquisition of English Orthography." ***Journal of Experimental Psychology: General***, 124 (2): 107-136.

Venezky, Richard L. 1970. ***The Structure of English Orthography***. The Hague: Mouton.

Wells, John C. 1990. ***Longman Pronunciation Dictionary***. London: Pearson Longman.

Wijk, Axel. 1966. ***Rules of Pronunciation for the English Language***. London: Oxford UP.

APPENDIX. FULL DATA CONCERNING REGULARITY AND FREQUENCY

	PR-1					PR-6			
Domain	**It.**	**Reg.**	**Irreg.**	**%**	**Domain**	**It.**	**Reg.**	**Irreg.**	**%**
<a>	2	1	1	50%	<a>	61	54	7	88%
<e>	7	7	0	100%	<e>	102	100	2	98%
<i>	3	2	1	66%	<i>	85	78	7	91%
<y>	17	17	0	100%	<y>	1	1	0	100%
<o>	11	7	4	63%	<o>	66	59	7	89%
<u>	0	…	…	…	<u>	NA	NA	NA	NA
Totals→	**40**	**34**	**6**	**85%**	**Totals→**	**315**	**292**	**23**	**92%**
	PR-2.1					**PR-2.2**			
Domain	**It.**	**Reg.**	**Irreg.**	**%**	**Domain**	**It.**	**Reg.**	**Irreg.**	**%**
<a>	174	144	30	82%	<a>	163	149	14	91%
<e>	163	163	0	100%	<e>	217	214	3	98%
<i>	190	158	32	83%	<i>	148	141	7	95%
<y>	3	3	0	100%	<y>	8	8	0	100%
<o>	106	54	52	50%	<o>	105	80	25	76%
<u>	81	74	7	91%	<u>	83	76	7	91%
Totals→	**717**	**596**	**121**	**83%**	**Totals→**	**724**	**668**	**56**	**92%**
	PR-3.1					**PR-3.2**			
Domain	**It.**	**Reg.**	**Irreg.**	**%**	**Domain**	**It.**	**Reg.**	**Irreg.**	**%**
<a>	92	90	2	97%	<a>	193	140	53	72%
<e>	14	13	1	92%	<e>	69	31	38	44%
<i>	98	85	13	86%	<i>	106	32	74	30%
<y>	2	2	0	100%	<y>	3	2	1	66%
<o>	65	49	16	75%	<o>	100	51	49	51%
<u>	23	23	0	100%	<u>	69	57	12	82%
Totals→	**294**	**262**	**32**	**89%**	**Totals→**	**540**	**313**	**227**	**58%**
	PR-4.1					**PR-4.2**			
Domain	**It.**	**Reg.**	**Irreg.**	**%**	**Domain**	**It.**	**Reg.**	**Irreg.**	**%**
<a>	44	37	7	84%	<a>	34	33	1	97%
<e>	22	22	0	100%	<e>	32	32	0	100%
<i>	12	12	0	100%	<i>	7	7	0	100%
<y>	0	…	…	…	<y>	0	…	…	…
<o>	41	37	4	90%	<o>	43	41	2	95%
<u>	13	13	0	100%	<u>	15	15	0	100%
Totals→	**132**	**121**	**11**	**91%**	**Totals→**	**131**	**128**	**3**	**97%**
	PR-5.1					**PR-5.2**			
Domain	**It.**	**Reg.**	**Irreg.**	**%**	**Domain**	**It.**	**Reg.**	**Irreg.**	**%**
<a>	16	16	0	100%	<a>	9	6	3	66%
<e>	7	5	2	71%	<e>	17	13	4	76%
<i>	11	11	0	100%	<i>	6	5	1	83%
<y>	0	…	…	…	<y>	0	…	…	…
<o>	11	11	0	100%	<o>	18	17	1	94%
<u>	7	7	0	100%	<u>	10	9	1	90%
Totals→	**52**	**50**	**2**	**96%**	**Totals→**	**60**	**50**	**10**	**83%**

It.: Word Items; Reg. Regular; Irreg.: Irregular