The reading aloud of one- and two- syllable words. Different problems that require different solutions? Insights from a quantitative analysis of the print-to-sound relations

Marielle Lange

(University of Edinburgh)

Alain Content

(Université Libre de Bruxelles)

Correspondence:

Marielle Lange

University of Edinburgh, School of Informatics

Institute for Adaptive and Neural Computation

5 Forrest Hill

Edinburgh EH1 2QL (UK)

M.lange@ed.ac.uk

Abstract:

Most models of English word recognition limit their domain of simulation to one syllable words and there is little straightforward empirical data to guide the development of more complex models of reading that would simulate the full set of words a reader is usually exposed to. However, typical reading material consists of polysyllabic words that are influenced by factors which are not present in one syllable words, such as the influence of stress on pronunciation, the influence of context, and the impact of segmentation ambiguity. An issue that arises, therefore is whether the present models, eventually very successful at simulating one-syllable words, in fact present a convincing solution to an inappropriately worded problem.

In this study, we present an attempt to reach a clearer understanding of polysyllabic word reading. As a result of the lack of empirical or modeling data, corpus analysis seems the most appropriate technique to use to try to systematically investigate the role of different possible factors on performance when reading aloud. A quantitative description of grapheme-phoneme associations of monosyllabic and disyllabic English words (with their British English pronunciations) is provided as well as details of the methodology adopted for segmenting semi-automatically the spelling and pronunciation of the words into graphemes and phonemes. The data obtained on the distribution of the pronunciations of the different graphemes of the language are used to proceed to a comparison of the predictability of the pronunciation of monosyllabic and disyllabic words. We argue that these data indicate that current theories of monosyllabic word reading cannot be taken as satisfying theories of reading for the whole range of words a reader is exposed to.

INTRODUCTION

[***Note: Notation conventions***: In this study, graphemes are formatted in bold (e.g., **ai**), phonemes are represented by the set of symbols from the International Phonetic Alphabet and enclosed in slant brackets (/b/, /p/, /a/), a silence is indicated by the symbol /=/, word or part of word exemplars appear in italics (e.g., *every*), a sequence of letters possibly forming a grapheme but not in that context is underlined (e.g., ***ph*** in u*ph*ill), a word's phonetic transcriptions are enclosed in square brackets (e.g., [evrɪ]). When illustrating segmentation into graphemes and phonemes, a “ ' ” is used to separate the grapheme or phoneme units in the strings (e.g., a'b'a'sh')]

Unnoticed to many skilled readers is their ability to make a match between some reference in a conversation to the name of a brand new product and the advert they read the day before in the newspaper, or to make a match between the name of a speaker, as introduced at the start of a talk, and the author name that appeared on the program at a conference.

This ability to derive the pronunciation of a letter string which has never been encountered before, as well as the fact that skilled readers are quite reliable in their reading of not so familiar letter strings, is usually seen as evidence that readers have some knowledge of the stable print-to-sound relations of their language is their cognitive systems. This hypothesis is at least largely accepted for readers of alphabetic languages, in which quite stable relationships exist between print and sound. After all, labelling such orthographies as alphabetic marks the fact that these languages evolved from a Roman alphabet in which each sound in the language was represented by a code of its own. In the course of history changes imposed in the sound of some words without a simultaneous change in their spelling have sometimes altered the initial one-to-one relationship, but all alphabetical languages retain that characteristic of having fairly stable print-to-sound relations: a letter sequence is frequently pronounced with the same sound (the sequence **sh** read as /ʃ/) and a sound transcribed by the same letter or letter sequence (the speech sound /ʃ/ written as **sh**). For instance, even though English has the reputation of being far more chaotic in its pronunciation than other alphabetic languages such as French, German, or Spanish, between 80 and 95 percent of words can be correctly pronounced by the application of letter(s)-to-sound rules [REF].

However, psycholinguists have not reached any consensus yet as to the exact nature of the representations and processes recruited for the conversion of a letter string into speech. This lack of consensus is particularly apparent in the drastically different if not antagonistic hypotheses of the dual-route and connectionist models of readers' performance.

The central dogma of dual-route (DR) models is that an explanation of humans' ability to read requires two distinct routes, or procedures: a global process for the correct pronunciation of words which have an arbitrary pronunciation (i.e., *have*), and an analytic process for deciphering previously unseen words that cannot be processed globally. Both procedures operate in parallel; the global one rapidly retrieves the pronunciation of the words that are already familiar and the analytic one derives the pronunciation by applying print-sound correspondences. The dominant instance of this theory is the dual-route model introduced by Coltheart and colleagues (Coltheart, 1978, 1985, xxx). In the lexical route, words are identified as wholes, by accessing them via their orthographic address, which is connected in the mental lexicon to both their phonological address and their meaning. In the conversion route, print is mapped onto sound at the level of the phoneme, with graphemes defined as the letter or letter combination that represents a single speech sound (for example, **p** in *print* or **ph** in *grapheme*). Coltheart's group assume a knowledge limited to the most frequent pronunciation of each grapheme of the language and conversion operates by identifying of the letter or sequence of letters that operate as a grapheme (for example, the sequence **ph**) and transforming it into its most common pronunciation in the language (*i.e.*, /f/ for **ph**). When phonological representations have been obtained for every grapheme in the string, they are merged into a phonological code. In the first version of this model, conversion operated in two steps: the isolation of the graphemes was followed by their translation. In the most recent version (*DRC-L* with *L* for Letters; Rastle & Coltheart, 1998; Coltheart et al., 2001), graphemes are translated into the corresponding phonemes during a letter-by-letter deciphering of the letter string.

In contrast, the single-route theory rejects the hypothesis of two procedures based on different computational principles (global and analytical print-to-sound translation) and the reliance on two separate sources of knowledge (word specific and infra-lexical print-to-sound relations). As stated by Seidenberg and McClelland (1989, p. 525), "*The key feature of [this model] is the assumption that there is a single, uniform procedure for computing a phonological representation from an orthographic representation that is applicable to irregular words and nonwords as well as regular words*.". That claim however only related to the processes involved in print-to-sound translation; the general theory is that skilled reading more typically requires the combined support of both the semantic and phonological pathways. The most familiar instance of this theory is the parallel distributed processing (PDP) model of *Seidenberg & McClelland (1989)*. In this model the translation of a letter string into its phonology relies on parallel distributed processing (PDP) in a connectionist network made of three fully interconnected adjacent layers of units: an input layer coding the orthographic form, an output layer coding the phonological representations, and an intermediate layer of hidden units. Activation propagates from the units of one layer to the units of the next layer, in a way that depends the value of the weight of the connections that link units from one layer to units from the next layer. All weights are set during a learning phase to encode the quasi-regularities in the spelling-to-sound mapping in a way that reflects the aggregate effect of training on the words that form the vocabulary of the network. Each time a word is presented, an algorithm estimates the discrepancy between the response produced by the network and the response it was expected to produce and uses this estimate to adjust the strength of the connections between the different levels of units (*back propagation learning algorithm*), such that performance improves gradually as the network discovers a set of connection weights that minimizes the error on the training corpus. This ability to learn its representations is an important property of this class of models because it complements an explanation of skill and impaired reading by an explanation of the way the system structures itself in the course of the development.

In the mid-eighties, Humphreys and Evett (1985) concluded that despite a great deal of research in cognitive psychology and the neuropsychology of language, it proved impossible to disentangle these drastically different explanations. It was therefore hoped that the introduction of computational models in the years that followed would have the potential to settle these theoretical disputes, by separating the realistic from the unrealistic as hypotheses of human processing

The use of computational modeling to express a theory in the form of a computer program or a connectionist network presents a number of advantages over a purely verbal model. First, for the model to run, the theory (cognitive processes and representations) needs to be both fully and adequately specified. As a consequence, the simulation has the potential to disclose any incoherence in the theory in the course of the model's development. Conversely, once completed, the computational model offers some support to the theoretical elaborations that the model implements by establishing that the hypothetical processes can be effectively instantiated. At this stage, the computational model elicits any hypotheses about the format of the representations or the properties of the mechanisms and eases the evaluation of the realism of the specific hypotheses. The simulation can also be used to explore the performance of the model as a function of its architecture and representations and in this way makes it easier to identify which aspects of the model are important for adequately simulating performance. For example, Norris (1994) used his model to evaluate the impact of particular computational choices about the associations that were represented (*e.g.*, using grapheme-phoneme, body-rime, or multiple levels of associations; and when the frequency value for the associations was computed either as a function of the number of words comprising the association or as a function of the lexical frequency of these words.

Further, compared to a verbal formulation that at most predicts the direction of the effect, a computational model is capable of producing precise quantitative predictions which can be directly compared with a large set of behavioral and neuropsychological data in every task the model simulates (e.g., model processing time vs. reaction time; model accuracy vs. error scores). The degree of fit between the model's performance and readers' performance can then be used to evaluate each theory individually and reveal the extent to which the theory from which the model was generated accurately describes the processes taking place in the human cognitive system (Coltheart *et al*., 2001; Rastle & Coltheart, 1999; Seidenberg & Plaut, 1998; Spieler & Balota, 1997) or to contrast the accuracy of different models and compare the degrees of fit of different models (Grainger & Jacobs, 1994; also Grainger & Jacobs, 1996).

In the presence of competing theories, the precision required by computational modeling help to reveal the ways in which the models make different predictions about the results of experiments not yet conducted. Carrying out these new experiments may potentially result in the reformulation or even refutation of at least one of the theories in competition. The computational model can then be used to understand how the failure to correctly predict some aspects of a phenomenon relates to specific assumptions of the theory and to encourage the reformulation of specific aspects of the theory in order to better simulate the behavioral data. The possibility of modifying the model prevents the predictions about how changes in the representations or processes would affect the performance of the network from being overly speculative (*observational fragility*).

The initial outcomes of computational modeling were more than encouraging in this respect. The introduction of Seidenberg and McClelland's (1989) connectionist model (henceforth, SM) and the demonstration of its ability to simulate key findings in reading brought considerable credit to ideas introduced by Glushko (1979) ten years earlier. However, the adequacy of the single-route theory was strongly undermined by more extensive comparisons with readers' performance. In its simulation of lexical decision performance, the single mechanism was found to produce about 80% errors when an average subject only produced about 6% (Besner, Twilley, McCann, & Seergobin, 1990; Fera & Besner, 1992); in a simulation of its performance on nonword naming, the SM model was found to produce about 55% of correct responses when skilled readers produced about 90% (Besner et al., 1990). In defense of their model, Seidenberg and McClelland (1990, p. 448) suggested that the poor performance in nonword reading could be due to the (small) size of the training corpus, which was limited to about 3000 words whereas skilled readers are exposed to approximately 10 times that number. They raised the possibility that this set was too limited to give the model the opportunity to learn all the print-to-sound relations necessary for a good generalization on unseen sequences. This was, however, dismissed by the demonstration by Coltheart *et al*. (1993) that its model read non-words as efficiently as human readers with a knowledge of the spelling-sound correspondence extracted by a rule-learning algorithm from the same vocabulary. As a result, the dual-route theory was reinstated. Coltheart and colleagues argued that poor performance in lexical decision was evidence for their view that a knowledge of the letter sequences that make up a word could not be developed without a lexical level of representation and the poor performance on nonword naming supported the argument that adequate generalization could not be achieved without a transcoding process relying on an explicit knowledge of the spelling-sound regularities of the language (for instance a grapheme-phoneme correspondence system).

However, as became apparent in the long run the contribution of computational modeling to the resolution of theoretical debates proved a lot more limited than initially expected.

Contrary to early expectations, the "runnable criterion" was not able to dismiss any of the leading theories. It was possible to produce instantiations of the most irreconcilable theories of reading aloud one can envision, the dual-route theory of reading defended by Coltheart and colleagues and the single-route theory introduced by McClelland and colleagues. In addition, computational modeling gave rise to the introduction of a tremendous variety of computer models of written word pronunciation (Ans, Carbonnel, & Valdois, 1998; Bullinaria, 1994; Coltheart *et al*., 1993, 2001; Harm & Seidenberg, 1999; Jacobs, Rey, Ziegler, & Grainger, 1998; Norris, 1994; Plaut & McClelland, 1993; Plaut *et al.,* 1996; Plaut, 1999; Rastle & Coltheart, 1998; Zorzi et al., 1998). Rather than constraining the explanations of the way spelling is transformed into sound, the introduction of computational models defined an even larger space of multi-facetted hypotheses regarding both the nature of the processes and the nature of the representations. Not all these computational models can be seen as complete models of reading as they do not address the six central facts about reading (Coltheart et al., 1993), namely (1) the reading aloud of exception words by skilled readers, (2) the reading aloud of nonwords by skilled readers, (3) the simulation of a lexical decision task, (4) the cause and source of surface dyslexia; (5) the cause of phonological dyslexia, (6) the emergence of developmental dyslexia. But at least four of them (PSMP, ZHB, DRC, ACV) were presented by their authors as capable of at least broadly simulating the major aspects of readers' performance, in a way that reasonably approximated the behavioral data.

Model	Number of routes	Nature of the knowledge of print-to-sound relations	Nature of the transcoding process
Coltheart et al. (DRC-L)	Two, based on different computational principles (parallel processing vs. sequential algorithm) and different architecture (interactive-activation network vs. rule system).	Explicit knowledge of the most frequent pronunciation of each grapheme of the language.	Sequential process, letter-by-letter rule-based transcoding of the letter string.
Ans et al. (AVC)	Two, based on similar computational principles (analogy) and similar architecture (memory traces).	Analogy to memory traces storing the pronunciation of known syllables.	Syllable-by-syllable deciphering, activation-synthesis of the pronunciations of all orthographic segments sharing any similarity with the sequence to be pronounced.
Zorzi et al. (ZHB)	Two, based on similar computational principles (delta rule) but different architectures (associationist vs PDP network).	Reflected in the connection weights of an associative network connecting letters to phonemes, with connections to the phonemes in the corresponding positions as well as to the adjacent positions.	Parallel and simultaneous translation of all the letters in the string (at least for monosyllabic words).
Plaut et al. (PMSP)	Single uniform process (though they include attractors that can be seen as specialised in irregular word reading).	Reflected in the connection weights of a network connecting graphemes to phonemes and transiting by a layer of hidden units.	Parallel distributed processing.

It soon appeared that the opportunity to generate pronunciations together with indices of the time required to generate such pronunciations proved of little help to indicate the lesser adequacy of any of these four theories of human reading performance. A first problem was the absence of any accepted procedure to assess what constitutes an unacceptable mismatch in performance. Spieler and Balota (1997) revealed that the performance of the *SM* and *PMSP* network was excessively poorly correlated with the pronunciation latencies of skilled readers. However, Seidenberg and Plaut (1998) showed that the comparison human-humans in Spieler and Balota's (1997) data was even worse than the comparison network-humans uncovered by these authors.

A second problem was the absence of any scaling of the different importance of the degree of fit for models defined by different approaches to modeling. In localist networks like DRC or MROM-P, the model is defined a priori by the modeler, and the parameters are set to determine the best fit between the network performance and a large set of behavioral data (*data-fitting approach*); in contrast, in connectionist models like PSMP or ZHB as little as possible is defined by the modeler; a great part of the network configuration is assigned to a learning algorithm. Also, the ranking of the models according of the degree of fit often varies with the data set presented to the model. The models simulate many different aspects of reading and they do not simulate them with an homogeneous degree of fit and it is difficult to relate the closeness of the fit to the realism of the different components of the model.

It also became rapidly apparent that the global degree of fit had little to say about the superior validity of any of the specific theoretical hypothesis introduced in one of these four computer models. For instance, the DRC model of Coltheart *et al*. (2001) is presented as the one with the best global performance over rival computational models in a simulation of skilled reading performance. For the authors, this suggested that the *DRC* model provided the most adequate description of the cognitive processes involved in reading, but there are several aspects of reading performance that concurrent models are reputed to accommodate far better than *DRC*. The main ones are the way that representations and processes are acquired as well as the cause of differences in the pattern of performance between readers of different languages (Besner, 1987, for a review) or of different ability levels (Herdman, LeFevre, & Greenham, 1994).

In sum, what we have learned from the past is that modeling is certainly an essential part of theory evaluation, for theory cannot be evaluated by itself. A model is a theoretical description that can help us understand how a system or process works, or how it might work (Cobuild). However, once a theory has been shown to meet the runnable criterion, the precise quantitative estimates of the degree of fit do not answer all the questions a psychologists has, for what is under evaluation is not the global performance of the model but the fact that every process or representation hypothesized in the model has a good chance to be the same as in human cognitive systems.

The lesson has certainly been taken. As early as 1994, Forster reminded us that the aim in psychology is not to produce some complex machinery that closely approximates the performance of skilled readers. The aim ought to be to produce a close approximation of the cognitive machinery a skilled reader solicits when reading aloud. He illustrated this with the example of his next-door neighbor: "*Suppose I discover that my next-door neighbor can correctly predict the outcome of every word recognition experiment that I do. This would be a surprising discovery, certainly, but it would not have any scientific utility at all until it was discovered how she was able to do it. I could scarcely publish my next-door neighbor as a theory without having explicated the reasoning involved*" (Forster, 1994, p. 1295). Authors of computational models as Plaut and McClelland have insisted on the need to compare the human and network performance in a qualitative way despite the fact that precise measurement are available. Even though a war of computational models has dominated the area for the last ten years, empirical evidence has regularly been added which had in view the evaluation of models' specific hypotheses. For instance, Kawamoto and Kello (1998) found that the pronunciation can be initiated as soon as the first phoneme has been specified which undermined the hypothesis shared by all current models that pronunciation starts only after all phonemes to be pronounced have been identified. Sometimes, empirical evidence has lead to redefinitions of the theoretical frameworks behind the computer model. For instance, when Rastle and Coltheart (1998) found the whammy effect (naming times are longer for nonwords that contain a multiletter grapheme such as “ph”), this proved incompatible with the hypothesis that conversion proceeds in two separate stages, *i.e*., segmentation into graphemes followed by conversion of each grapheme into its most frequent pronunciation. The hypothesis was replaced by one which rather assumes a letter by letter decoding process with activation of graphemes every time a new letter is read.

However, one lesson that still needs to be learned is that there is another central drawback of the use of computational models as a substitute for the theory under evaluation. Because of the urge to test the goodness-of-fit between simulated and performance data, experimentation is almost always conditional on the predictions of the runnable versions of the models. The introduction of computational modeling has dramatically limited the development of better theories of reading because it leads researchers to collect data in the range that is already understood, that is, the range that the models simulate successfully.

In this paper, we try to demonstrate that a deeper understanding of the cognitive processes underlying reading can be obtained when the model is confronted with what it cannot yet explain. Importantly, what is scarcely discussed in theoretical papers is that none of the four dominant models of English word recognition explains how skilled readers process the majority of words they encounter.

At present, computational models of English reading only simulate monosyllabic words. They all have coding constraints that only allow the representation of one-syllable words and thereby exclude a large number of the words a reader is exposed to. In fact, Gimson (1980) estimated that 80% of the words in the oral modality are polysyllabic and Henderson claimed that this proportion is certainly larger in printed texts (Henderson, 1982, p. 462). [can make mini-simulation on this, using this very paper, counting the number of words with non-adjacent vowels, except e#, ed#, es#].). Whatever the degree of match to existing data is, a theory will inevitably be incomplete if it does not provide realistic hypotheses about how the print-to-sound translation processes cope with polysyllables.

In English, polysyllabic and monosyllabic words differ in important ways in their properties. In particular, with polysyllabic words, stress, syllabic structure, and morphology, all come into play. It is largely accepted that these properties make deriving of the pronunciation a much more complex problem than it is for monosyllabic words. In polysyllabic words, there are important changes in pronunciation in the unstressed syllable of polysyllabic words: complex vowels normally associated to a complex vocalic sound (*i.e.*, diphthongs) in stressed syllables are often pronounced with a simple vocalic sound in unstressed syllable (*m*__a__*te* [m**eɪ**t] but *clim*__a__*te* [klaɪm**ɪ**t], *sust*__ai__*n* [səst**eɪ**n] but *mount*__ai__*n* [maʊt**ɪ**n]) and vowels normally associated with a simple vocalic sound (*i.e.*, monophthongs) in stressed syllables are generally reduced to a *schwa* in an unstressed syllable (*t*__e__*ntative* [t**e**ntətɪv] but *pat*__e__*nt* [peɪt**ə**nt] -- manage manure). Also in polysyllabic words, various orthographic structures determine variations in pronunciation. This is the case for morphological or syllabic boundaries (*ph* is pronounced differently in *shepherd* and *morpheme*, *oa* is pronounced differently in *gaol* and *chaos*), morphological structure (*e* is typically pronounced /I/ in the prefixes *re*, *be*, *react*, but /x/ in *reach* or bexx), as well as information about syntactic class (e.g., final *y* is regularly sounded /ɪ/ in adverbs as *amply* and /aɪ/ in verbs as *comply*) or noun etymology (XXX). In addition, polysyllabic words tend to include proportionally more borrowings from French, Greek, and Latin words which are often associated with "non-native" pronunciations.

Currently, little is known about the degree to which each of these factors affect print-to-sound predictability in polysyllabic words. A large number of the explorations of language structure follow the computational models and concern only monosyllabic words (Treiman et al., 1995, Ziegler et al., 1997, Peereman & Content, 1997, for a quantitative description of the body-rime associations in English and French; Stanback, 1992, for a descriptive study of subsyllabic segments). With descriptions of the grapheme-phoneme regularities in English in which polysyllabic words are included (Berndt, Reggia, and Mitchum, 1987; Haas, 1970; Hill & Ure, 1962; Kruisinga, 1957; Venezky, 1970; Wijk, 1966 -- Berndt et al., 1994 for statistics about graphemes), no separate estimates are available for monosyllabic and polysyllabic words.

However, a study by Zhang (1995, p. 81) made it particularly apparent that the performance of a model specialized in the reading of monosyllabic words cannot be taken as a reliable index of its performance when exposed to polysyllabic words. This study, which presented data about grapheme-phoneme associations, their dispersions and strength, based on an algorithm for the automatic derivation of the pronunciation of American English words, revealed that a rule system derived from a set of monosyllabic words predicted about 97% of correct pronunciations for monosyllabic words, but only 19% of correct pronunciations when the same table was used to determine the pronunciation of disyllabic words (cf. 4% on 3 syllable words, 0.4% on 4 syllable words).

Assuredly, given these differences, whatever close the degree of match might be, a computational model should never be considered as a plausible theory of the reading processes if it is inherently unable to process polysyllabic words. Also, because of the important differences in properties between polysyllabic words and monosyllabic words, it is almost certain that more details of the machinery used by a skilled reader could be identified by including polysyllabic words. Notably, by limiting themselves to a simulation of monosyllabic words, many theories of English word recognition leave unexplained the way stress assignment and vowel reduction is handled during conversion when polysyllabic words are read aloud. It is possible, as Rastle & Coltheart (2000) proposed, that stress is non-lexically derived. However, given that stress cannot be accurately predicted on the basis of strictly non-lexical information (*i.e.,* different pronunciations are given to the very close spellings in *nature* vs. *mature*) such a system will never be completely reliable. The wide variation of pronunciations as a function of stress and morphological information might also undermine the view that the pronunciation of polysyllabic words can be obtained by a single uniform procedure, as stated in single-route models.

The difficulty then is to know how to orient experimental studies when it is unclear as to what aspect of polysyllabic words pose problems to polysyllabic word reading and ALSO as to what the prediction of each theory should be on these strings. What needs to be recognized is that there is a theory behind a computational model and that these theoretical hypotheses can be evaluated independently of the availability of quantitative predictions. But still, there needs to be some way to guide the collection of empirical data: intuitions need to be gained about ways in which current theories might be inadequate before initiating empirical studies.

For this, we will adopt the metatheoretical approach introduced by Frauenfelder, Content, & Peereman (1996). This approach, illustrated in *figure xx*, basically proposes to integrate the computational and empirical approach with work describing the properties of a language based on the analyses of lexical databases. The benefit of such an approach has already been made apparent by Treiman and colleagues, who have shown that statistical analyses of the properties of the words in a given language used in conjunction with empirical studies can also help to better understand the domain of simulation and constrain the mechanisms implied in the acquisition, perception, and comprehension of the words of the language (e.g., Treiman et al., 1995). + Jared & Chateau (2003)

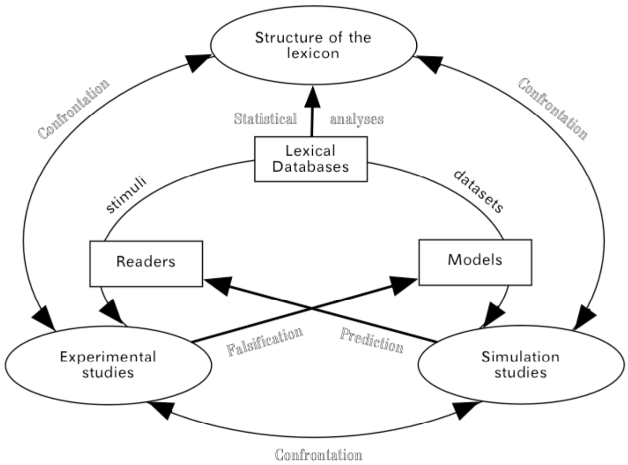


FIG xx. Metatheoretical approach (adapted from Frauenfelder, Content, & Peereman, 1996)

In the present paper, we start from statistical analyses introduced in the form of a linguistic description which provides precise measures of the ambiguity of the pronunciation of the different graphemes of English in both monosyllabic and disyllabic words (based on a corpus of about 20,000 words of 1 or 2 syllables with their British pronunciation). The prelude to the analyses is the segmentation of two large corpora, one of monosyllabic words and another one of disyllabic words, into graphemes and phonemes. The end result is the abstraction of a table that provides a frequency count for each grapheme-phoneme association. The guidelines as well as the procedure followed to achieve that goal are described in details in the Method section.

The statistical data revealing grapheme-phoneme relations in English polysyllabic words are then used to evaluate the efficacy of grapheme-phoneme associations for translating disyllables into their phonology. This evaluation will in turn be used to discuss the realism of the grapheme-phoneme conversion system in *DRC*. As we mentioned earlier, Coltheart and colleagues put forward the higher degree of fit between simulation and human data to suggest the higher realism as a model of reading. However, what was most apparent in that study is that much of the variation in word naming performance is not explained even by this model (Coltheart et al., 2001). Although *DRC-L* obtained a superior performance in nonword naming, explaining 48% of human variance (compared to 3% and 12% for *PMSP* and *ZHB)*, it only explained 5% of human variance on the naming of words, which are the most usual reading material (see *Table 5* in Coltheart *et al*., 2001). Clearly, even though this model is presented as the best guess of what the human activity might be when reading aloud, there is certainly room for refinements of the hypothesized reading processes. In particular, the grapheme-phoneme conversion system does not yet capture human behavior, as strongly argued in a study by Andrews and Scarratt (1998), which compared the pronunciation of nonwords by readers and by the models *DRC* and *PMSP*. Their study showed that even though both *DRC* and *PMSP* were shown to produce a regularity effect, and a regularity by frequency interaction, as well as a realistic number of errors, the pronunciation of the string by the readers differed in important ways from the pronunciation of the strings by the computational models. This finding also strongly questions the conclusion of Seidenberg et al. (1994) that "*[b]oth models [DRC and PMSP models] generate plausible nonword pronunciations and match subject's responses accurately*" (p. 1177).

That said, it may look as if we were trying to disprove that model and implicitly restore the others. Although DRC is the prime target in our discussions, this Is not at all that we believe it is fundamentally wrong. It is rather that it makes the easiest target of our evaluation. Despite the fact that the current computational model does not allow the representation of polysyllabic words, it is possible to extrapolate its performance on polysyllabic words, as the theories behind it makes explicit predictions about the format of rules and representations. In the absence of a model which is able to process polysyllabic words, any discussion of the theoretical implications of our results would prove highly hazardous with connectionist models. In such models, the configuration is the result of the execution of a learning algorithm which encodes emergent regularities between input and output in the words in the corpus.

Clearly, our purpose is not to suggest the lack of validity of the dual-route theory, as a whole. The fact that computational models rely on completely different theoretical hypothesis in our view highlight the fact that we still understand too little of the representations and processes involved in reading to adequately constrain computational models. If DRC was shown to offer a better global degree of fit, three other models were shown to be significantly correlated with human performance as well. It is a lot more probable that all of these models capture part of the truth about cognitive processes and that the part captured by rival models is not necessarily a subset of the part captured by the model characterized by the higher degree of fit. It is more important to try to understand what aspects of performance relate to what aspect of the models than to dismiss all hypotheses of a theory on the ground that one is shown to be inadequate.

Our demarche is therefore to use the current computational models to understand the current space of possible hypotheses about the nature of the reading systems and to conduct empirical evaluation and statistical analyses to demonstrate that some of these hypotheses, though realistic because used to define a runnable model, are not fully coherent with what is known about readers' performance. The goal is not to use such findings to suggest the lack of validity of any theoretical approach. After all, it has already been said elsewhere that none of the current models provides a truly convincing explanation (as discussed by Davis, 2000, Grainger & Whitney (2004), Shillcock *et al.* (1999), or Whitney (2001) assumptions about input are not realistic). Our goal is rather to identify the weaknesses of some theoretical instantiations to facilitate a reduction of the domain of the hypotheses. It is then the job of modelers to find what aspect of their model is the cause of that failure and in what ways the model can be accommodated to correctly simulate the data and from there to establish whether these findings determine the downfall of the theory.

Three theoretical issues will be discussed, on the basis of our statistical analyses of grapheme-phoneme associations:

(1) The properties of polysyllabic words make us predict that a system which has a knowledge limited to grapheme-phoneme relations (i.e., NO context) will be relatively inefficient, as will a system in which stress is determined during print-to-sound conversion. This therefore throws doubt on the sufficiency of a strictly grapheme-phoneme system for the reading of polysyllabic words, as supposed by *DRC*. Furthermore, while low-level – grapheme-phoneme – non-lexical conversion rules have proved powerful enough for the derivation of the pronunciation of a large percentage of monosyllabic words, the important reduction in predictability of the relationship between graphemes and phonemes in disyllabic words suggests that the rule-system presently implemented in the *DRC* model will have enormous difficulties with the processing of polysyllabic words. In addition, the greater influence of orthographic context on pronunciation in polysyllabic words might challenge the idea of a strictly grapheme-phoneme correspondence system. This aspect of English is even more apparent in polysyllabic words than in monosyllabic words.

(2) The properties of polysyllabic words predict that deciphering on the basis of the units presently listed as graphemes in the DRC rule system is an inefficient process (e.g., a lot of digraphs are alternatively pronounced with one or two phonemes in polysyllabic words: *oa* in *goal* vs. *foam*, *ai* in *naive* vs. *waive*). This questions the status of the grapheme, which is defined as the written counterpart of a phoneme, as a coherent unit of description and representation.

(3) The way the description of phonetic phenomena in polysyllabic words demand an output in the form of phonetic features questions the option present in almost all models of a print-to-sound conversion system of an output in the form of a sequence of phonemes. It is acknowledged that letter-sound relationships are governed as much by their environment, or their position in a word, as by simple grapheme-phoneme correspondences rules. G. B. Shaw's renowned spelling *fish* as *ghoti*, with **gh** from the ending sound in *rough*, **o** from the first vowel sound in *women*, and the **ti** from the middle sound of *nation* ignored this altogether. However, it is largely ignored that sometimes, it is the phonetic environment and not the orthographic environment that predicts a change in pronunciation. To refer to the *ghoti* illustration, the sequence **gh** never represents the sound /f/ at the beginning of a word; **ti** for /ʃ/ occurs only in a particular phonological context that causes the palatalisation of the consonant.

METHOD

The aim of the analyses was to gather data on how the pronunciations of the different graphemes of the language are distributed. This required the production of a table providing estimates of the frequency with which each grapheme-phoneme association occurs in the language, which in turn required the segmentation of the spelling and pronunciation of words into graphemes and phonemes.

The methodology used for these analyses proceeded in four steps: (1) a corpus was selected for the analyses. (2) The mappings from letters to sound that will form the grapheme-phoneme associations listed in our table were determined, then the orthographic and phonological representations of each entry of the corpora were segmented according to the graphemes and phonemes listed in the table. (3) The grapheme-phoneme association frequency values were calculated for each entry of the table. And finally, (4) different estimates of grapheme and grapheme's pronunciation predictability were derived from this association frequency count. At every step, we aimed for linguistic and computational coherence of the resulting grapheme-phoneme association system. Chiefly, in agreement with most of the linguistic analyses (e.g., Venezky, 1970, p. 52; Gak, 1976, p. 31; Haas, 1970, p. 42), we treated some letter clusters as separate and mute graphemes (e.g., final *e* in *date*) and we had the option of taking into account the influence of neighboring graphemes and phonemes on the pronunciation of one spelling unit (e.g., in *race*, *e* marks a soft pronunciation of the *c*).

The result of this was a system of about xxx "rules". The complete list of associations, with an illustrative example for each association, is provided in the on-line documents. The grapheme-phoneme association tables associated with the three different corpora can be downloaded from the website of the "University of Edinburgh" (http://homepages.inf.ed.ac.uk/mlange/gpa/). A segmented corpora and a computer program to recover the segmented representations and compute the selected variables is made available. This computer program will also recover the orthographic and phonological representations segmented into graphemes and phonemes for words in the corpus or construct it using a similar segmentation algorithm as the one used in our analysis for unfamiliar words or nonwords.

1. English corpus

Analyses were conducted on three distinct corpora, one of 6,448 monosyllabic words, one of 13,627 disyllabic words, and one of 20,075 words that combined the monosyllabic and disyllabic corpora. Entries were extracted from the *Celex* computerized database for English (CELEX; Baayen, Piepenbrook, & van Rijn, 1993), using the "*lemma corpus*". For each lexical entry, we recovered the data relating to the primary phonemic transcriptions of British English (DISC format, which assigns an ASCII code to each phonological unit in the phonetic system of English), the stress pattern, and the word form frequency (spelling wordform frequency, *Cobuild F/1 million* value).

Discarded entries. A decision was made to reject entries that do not really correspond to English words: (1)

entries corresponding to abbreviations and contractions (*e.g.*, *std* and *hasn't*), as well as compound words (*e.g.*, *space-suit* and *back door*); (2) words with diacritics (*e.g.*, *entrée*) on the grounds that the diacritics provided a cue to the foreign origin of the word, indicating that the spelling-to-sound conversion rules of English are not valid for this word -- foreign words compatible with the character set of English were retained in our corpora, although it could be considered that certain letter sequences reflect the foreign origin of a word (*e.g.*, *chief* and *chef*); (3) entries wrongly categorized as two-syllable words, principally because of an error in the coding of the phonology or an error in the syllabic segmentation of the phonological transcription. For instance, *beseeching* was coded with an incomplete phonological transcription [bɪsiːtʃ] that did not match the transcription of the Cobuild dictionary of English language [bɪsiːtʃɪŋ] and *alias* was segmented into two syllables in the phonological representation but three in the orthographic representation.

Changes to the phonetic transcriptions. Modifications of the phonetic transcriptions were applied at different points in time in the analyses described bellow. Prior to any analysis, the /R/ which, in *Celex*, indicates a silent linking /r/ (*e.g.*, *acto<u>r</u>*) was deleted (*e.g.*, *actor* [æktəʀ] was changed into [æktə]). At a later stage of the analyses, polyphonic phonemic units coded by a single DISC code in the phonetic transcription provided by *Celex* were replaced by the DISC codes of their constituent phonemes in order to facilitate the identification of the nature of the phonetic transformation. This was specifically the case for affricates (/tʃ/ and /dʒ/), and syllabic consonants (/l̩/, /n̩/ and /m̩/ recoded as /əl/, /ən/, /əm/), and for some diphthongs (e.g., /ɪə/).

[Note: Although these code modifications contribute to the legibility of the transcriptions, legibility was not the most important motivation for them. With affricates such as the /dʒ/ sound in *adjoin*, it often appeared more appropriate to consider that the pronunciation resulted from the successive translation of the graphemes **d** and **j** (/d/ followed by /ʒ/), possibly merged into an affricate during (co–)articulation, rather than from the dubious translation of the grapheme **dj** into /dʒ/. For syllabic consonants, this option ensured the homogeneity of the coding in Celex (Originally, the phonology of *briton* was /brɪtn̩/ but the one of *sexton* was /sɛkstən/; in our analyses, /n̩/ was always written as /ən/). Systematic rewriting of these diphthongs as two phonemes were encouraged by the fact that the diphthong codes were sometimes evidently inappropriately assigned to sequences of phonemes. Chiefly, for most trisyllabic words (not included in our study), phonemes belonging to different syllables were inappropriately grouped under a diphthong code (*a-li-as* with [ɪə]represented by [7] in Celex)]

Outside these systematic recodings, only a very small number of phonological representations were altered. In all these cases, the Celex phonetic transcriptions appeared as incorrect and were coded in a different way in the *Cobuild* Dictionary (for instance, the phonology of *glacial* was changed from /gleɪsjəl/ to /gleɪʃəl/).

2. Semi-automatic segmentation of the spelling and pronunciation into graphemes and phonemes

Principal decisions (underpinnings)

The prospect of counting the frequency of the grapheme-phoneme associations requires that the orthographic and phonological representations of each string be segmented into matching spelling and sound units (here graphemes and phonemes). An important difficulty is that although there is a received inventory of the phonemes of English (cf. *International Phonetic Alphabet* which proposes one symbol for each speech sound considered as distinctive for the language), there is no such inventory of graphemes, their orthographic counterparts.

The fact that graphemes are traditionally defined as letters or group of letters that transcribe a single phoneme suggests that there is a mandatory one-to-one match between the grapheme units in the orthographic transcription and the phoneme units in the phonemic transcription. However, segmentation ambiguities may arise when graphemes can be seen as matched onto a silence, not represented in the phonetic strings.

A difficulty that rapidly emerges when attempting to segment a string into graphemes is that there is not necessarily the same number of graphemes as phonemes; sometimes graphemes can be seen as matching a unit that is not represented in the phonetic transcription. The clearest case is the one of letters or letter clusters that appear to be simply unsounded (e.g., **p** in *pneumatic*, **s** in *island*, **b** in doubt, **n** in *hymn*). This is typically the case with functionless scribal insertions (**b** in *debt*, **c** in *indict*, *victual*) added at some point in the history to mark the Latin or French origin of the word (sometimes based on false hypotheses). [Note: for example, Middle English vitaille [vitæil], from Old French vitaille, spelled *victual* because of the Latin root *victualia*; compare *victual* and *perfect*, *island* and *baptism*; *e.g.*, **c** in *indict* vs *depict*, *d* in *debt*; Middle English *yland* transformed into *island*, with a similar spelling to the unrelated *isle*).]. This is also the case with letters that at some point in the history of the language had their sound become unsounded in the spoken form because of the phonetic environment (*e.g.*, **e** in *ev***e***ry* [evrɪ], **b** in *subponea* [suponea]; the post-dental plosive **t** which was regularly lost in *often* in the sixteenth century, Potter, 77 [REF]), but were retained in the spelling because they give a morphological or syntactic information (**b** in *bomb* and *bomber*, *bombard*; Catach, 1984, Henderson & Chard, 1980, Vallins, 1965, Venezky, 1999 vs sounded **l** in *fault*, sounded **d** in *adventure*). This is also the case with letters that cannot be matched onto a specific phoneme because their realisation has merged with the one of an adjacent letter. An example of this is the palatalised pronunciation of consonants that takes place in words as *special* and *nation* when the following phoneme /j/ (yod) merges into fricative or labio-dental phonemes /s, z, t, and d/ to give the pronunciations /ʃ, ʒ, tʃ, dʒ/).

Ambiguities in the segmentation arise every time the number of units you are ready to see as a grapheme does not match the number of phonemes in the transcription. This is illustrated by the word *half*, which has been segmented in three diffferent ways in published works: as **h:a:l:f** (Venezky, 1970), **h:al:f** (Coltheart *et al*., 1993), and

h:a:lf (Berndt *et al*., 1987; this study also lists the graphemes **ld** in *would*, **lk** in *walk*, **lm** in *palm*, and **lv** in *calves)*.

Therefore, prior to any counting, guidelines need to be established to guide the segmentation of letter sequences into graphemes as well as to determine how to account for the contexts that modify the grapheme's pronunciation in a systematic way. The optimal guidelines are not easy to determine. Large units hide ambiguities in the pronunciation with ad-hoc segmentations of the letter sequence as a function of the pronunciation of the sequence; small units miss numerous subregularities in the pronunciation. An illustration of the former is the case of **ough** at the end of words. It may be tempting to treat it as one unit when pronounced with a single phoneme, as in *through* /θruː/. But then there is the problem that the same unit is matched onto two phonemes in other words, as in *enough* /ɪnʌf/, without clear contextual constraint on the segmentation. In *slough*, for example, *ough* is pronounced with one phoneme, /aʊ/, when used as a noun -- [slaʊ], "swampy ground" -- but with two when used as a verb *slough* -- [slʌf], "to cast off"). An illustration of the latter is the case of silent **l** in **a+l+Cons** (for instance, *al* in *balk* /bɔːk/). To treat it as a separate unit, would cause the loss of any information about the context in which it is silent (a **l** followed by a consonant is often silent in *al* but never in *el*, *il*, or *ul*).

The following solution was proposed by Venezky (1970, 1995) and Haas (1970): to be ready to recognize some letters as separate units associated to a silence (*e.g*., *b* in *de<u>b</u>t*) and to add marking codes to some graphemes to signal contexts reliably associated with a change in pronunciation (*e.g*., *c* has is soft pronunciation before the vowels *e*, *i*, *y* and its hard pronunciation in other contexts; it is is pronounced /s/ in cell but /k/ in car).

<u>1. Graphemes can be silent</u>

The recognition that silent graphemes can be silent prevents the identifying of dubious grapheme units (***bt*** in *de**bt***) associated with non-systematic segmentations (*e.g*., compare t in *de**bt*** and *o**<u>bt</u>**ain*).

<u>2. Graphemes can be adjoined marks of contextual influences Constraints upon the translation</u>

Adding marking codes captures the fact that some letters help disclose the pronunciation of an adjacent letter. 'Mute' **e** is the clearest case of this. Mute **e** has many special functions which alter the pronunciation of one of the preceding letters. After a single vowel and a single consonant, a final **e** distinguishes a word that has its vowel pronounced with a long sound from a word that has this vowel pronounced with a short sound (contrast *mate*, *hide* and *mad*, *had*); after **c** and **g**, an **e** distinguishes the words with the soft (/s/ and /dʒ/) and hard variants (/k/ or /g/) of these consonants in front of the vowels *a*, *o*, *u* or at the end of words (*e.g*., *practicable* vs. *notic<u>e</u>able*, *flag* vs. *page*, or *arc* vs. *farce*); after final **s**, a final *e* distinguishes a word that ends in voiceless *s* from a plural *s* that is pronounced /z/ (contrast *dense*, *dens*). After final **th**, *e* may distinguish a verb with voiced *th* /ð/ from a noun with voiceless *th* /θ/ (*sheath*/*sheathe*, *teeth*/*teethe*, *wreath*/*wreathe*).

It may be tempting to see an unsounded letter that specifies the pronunciation of polyvalent letters as acting as a diacritic mark, like marks added on the top or bottom of characters in some languages to identify the precise realization of the letter (e.g., *é* and *â* in French). As a matter of fact, some of these diacritics signs were in the remote past represented with adjoined letters. For instance, what is now a cedilla beneath the letter **c** (which, in French, marks a soft realization of the **c** phonemic associate), was originally represented by **cz** [Note: it was replaced by **c** with a subscripted **z**, then replaced by a cedilla borrowed from Spanish by Geofroy Tory during the 16th century].

However, there are two main problems with treating mute **e** as a diacritic, attached to the letter it marks. Firstly, it would be rather inefficient as *mute e* sometimes marks the value of different preceding letters simultaneously. For instance, the final **e** in *pace* marks the long pronunciation of the preceding vowel **a** and the soft pronunciation of the **c** at the same time. If **e** had to be glued to one particular letter, it could be to only one of them, which would mean the loss of any information about its influence on the other grapheme(s). Clearly, a better description is obtained when the characters that disclose the pronunciation of an adjacent phoneme are considered distinct from the letter they mark. Secondly, the marking function is sometimes completely independent of the phonological realisation of the marking letter. Although **e** softens the **c** in the same way in the words *dance, cell, dancer*, it is unsounded only in the first word, and therefore cannot be viewed as a diacritic in this case.

In addition, orthographic alternation patterns provide a further demonstration that the marker is independent (Venezky, 1970, 1995). When the final **e** functions as a marker for the preceding **c** or **g**, it is dropped before a suffix which begins with a letter which constrains their pronunciation in a similar way. Hence, in English, the **e** of *notice* disappears before na *-ing* suffix in which the **i** already marks the association of **c** with /s/ (e.g., *noticing*), but it is maintained before an *-able* suffix (because *noticable* would associate **c** with /k/). Similarly, the **e** added after an otherwise final **u** is dropped before any suffix because the sole function of the **e** is to avoid the presence of a final **u** (*argue*, *arguing*). This is really quite systematic. In inflected forms, **e** is only retained to avoid ambiguity (contrast *singing*/*singeing*) or, exceptionally, in *ageing* (although *aging* also occurs, especially in American English, in which it is the preferred form).

3. Marks code linguistic phenomena rather than the adjacent letter

It remains to be decided what is the most appropriate way to add marks of context. The obvious option is to use the identity of the adjacent letter as the mark of context. A difficulty with this option, however, is that similar transformations occur in different contexts. For instance, the **c** followed by a mute *e* case is just a particular instance of the general rule that says the soft sound of **c** tends to occur before *i*, *e*, or *y*, while the hard sound occurs before *a*, *o*, and *u*. If the adjacent letter was used as a mark of context, a different coding would sometimes have to be used in words as closely related as *dan*:$\mathbf{c_e}$:*er*, and *dan*:$\mathbf{c_i}$:*ing*, with different frequency estimates collected for $\mathbf{c_e}$ and $\mathbf{c_i}$, even though they are pronounced the same.

Therefore, for clarity and coherence in the linguistic description, we took the option of grouping under a common code the various contexts which condition the pronunciation of a grapheme in an identical way. For example, the label *soft* was used to mark the soft pronunciation of the letters **c** and **g** (*i.e.*, /s/ and /ʒ/) which generally occurs when these consonants are followed by one of the letters **e**, **i**, **y** ($\mathbf{c}_{\mathbf{soft}}$*ell*, $\mathbf{c}_{\mathbf{soft}}$*ity*, but **c***an*). [Note: all realized as front vowels].

An advantage of using meaningful labels rather than the adjacent letter to mark contextual influences is that different labels can be introduced for the different ways the same grapheme influences the pronunciation of adjacent phonemes. The change in pronunciation indicated by the letter **e** is not the same in *noticeable* and *ocean*. The **c** is pronounced /s/ in the former case, but /ʃ/ in the second. Importantly, to understand these two pronunciations, it is useful to consider a distinction between changes in pronunciation induced by an orthographic context and those induced by phonological phenomena (such as palatalisation). With the sequence *ce*, for instance, the cue to the silent value of *e* in *noticeable* is clearly in the orthography as the *e* is dropped in *noticing*. However, in *ocean*, the silent value of *e* seems rather result from a phonological phenomenon. In that context, the grapheme **c** takes an unusual /ʃ/ pronunciation which exactly parallels the one of **c** in *precious*, with a prevocalic and unstressed vowel (e or *i*) being softened to a yod /j/ and the yod is assimilated to the preceding consonant. Considering phonological phenomena also helps to understand differences in pronunciation in circumstances where the orthographic context is strictly identical (e.g., *ti* is is pronounced /ti/ in *chris-t-i-an* but /ʃ/ in *gen-ti-an*).

The grouping and labelling option again importantly contributes to clarity in the way context influences pronunciation. First, it forces us to capture most of the linguistic phenomena in a complete and transparent manner as shown by the example of the palatalised consonants. The pronunciations of **ss** in *issue* [ɪ ʃ uː] or ***z*** in *azure* [æ ʒ ə r]), (which exactly parallel **ss** in *mission* or **s** in *Asian*) can be nicely elicited with the following marking options. The grapheme **i** would receive the mark *yod* because /i/ before a vowel is reduced to a schwa and transformed into /j/. The grapheme **u** would receive the mark *yod* in the appropriate phonetic contexts to indicate the presence of an inserted yod (that is, in *venue* [v e n juː] and *tenure* [t e n jə], but not in *blue* [b l uː]). Some of these **u** graphemes would be further marked as *ass* (assimilated) when the inserted yod is merged into the preceding sound (also in *measure* or *adjure* [a d ʒ ʊə]). The consonants would receive the mark *pal* indicating they are palatalised when followed by an assimilated yod. Second, we are able to provide a precise account of successive transformations in the presence of a phonological phenomenon which would otherwise be impossible to achieve. For example, the ***i*** in *fanion* and *indian* would be marked as *yod* and the unsounded ***i*** in *precious* and *cautious* would be marked as *yod+ass* (assimilated yod) because it is viewed as translated by a /j/ which is then assimilated to the previous consonants (e.g., $prec_{pal}i_{yod+ass}ous$).

[Note: The terms of “graphotactics” and “phonotactics” have sometimes been used to refer to modifications which orthographic or phonetic contexts impose on a grapheme's pronunciation. Haas (1970, p. 59), for example, proposes: "*If we say (with certain qualifications) 'Grapheme **c** sounds /k/ if it occurs before <a, o, u>, but /s/ before <e, i, y>', we are referring to purely graphemic conditions of the occurrence of **c**; we are stating a <u>graphotacti</u>c rule for the reader*." The term of *phonotaxis* is also sometimes invoked in the description of the constraints of phonemic context on pronunciation: restrictions on the way in which phonemes can be sequentially arranged in a syllable have a direct impact on pronunciation. Venezky (1970) made this clear using the *subponea* example: certain sequences of stop consonants such as the sequences /pb/ and /bp/ do not occur in English, so when their normal graphemic correspondences occur the first grapheme is silent. However, these terms obscure the difference between atypical translations (e.g. of *ce* as /k/) and prohibited letter or phoneme sequences of English (e.g., initial *ck*) even though the second aspect is of no relevance for print-to-sound translation. As a matter of fact, knowning that **ck** can only appear at the end of a word has no predictive value for the pronunciation. Furthermore, although the illegality of the sound sequences /pb/, /bp/, /wr/, /mn/, and /mb/ determines a predictable alteration of the phonemic associates in *write, hymn* and *lamb*, it does not reliably predict which pronunciation will occur (for instance, **wh** is translated into /w/ in “whale” but into /h/ in “whole”). Moreover, we cannot be sure that a knowledge of the possible phoneme sequences in the language actually takes part in the grapheme-phoneme transcoding process (it is certainly more efficient to directly code the fact that **w** is silent before **r** or that the **wr** sequence is always pronounced /r/). We prefer therefore to use the terms orthographic or phonetic environment to make it clear that we refer exclusively to information that is used when translating spelling into sound.

Description of the semi-automatic procedure

In our system, a grapheme can be a sounded letter or a mute letter or a group of letters. Which letters have to be associated to a silence and which of them have to be considered as forming a multiletter grapheme? A segmentation process is necessary to decide this. The aim is to determine how to isolate the units in spelling and sound which match onto grapheme-phoneme associations. This has to be done in a way that uses every letter or phoneme of the string and that never uses the same letter or phoneme in different associations (no overlap). Furthermore, associations have to be selected in a way that determines a coherent segmentation of every string in the corpus.

This difficult task was approached in two stages: First, we isolated a set of grapheme-phoneme associations which were realistic (consistent with linguistic descriptions), coherent (the same group of letters was segmented the same way in every word it is part of) and comprehensive (capable of segmenting every word in the corpus). Second, we proceeded to reconsider some of the grapheme units to better account for contextual influences.

1. First table of grapheme-phoneme associations

The initial set of grapheme-phoneme associations was isolated through different passes through the corpora using a semi-automatic approach: additions or deletions of grapheme and phoneme units from the association table were performed manually but numerous computer algorithms were used to, on the one hand, help isolate grapheme-phoneme correspondences, and, on the other hand, guarantee a systematic segmentation of the orthographic and phonological transcriptions according to the associations listed in the table.

Analyses were conducted on the corpus of monosyllabic words first and they were extended to the corpus of polysyllabic words only after complete identification of the constituent associations in the monosyllabic words were completely identified. About nine runs were necessary to determine a set of associations which could to segment each word in the corpus.

Initially, only words with a straightforward mapping of graphemes onto phonemes were considered. These were words with as many letters as phonemes, excluding strings which contained potential multi-phoneme units (*i.e.*, groups of phonemes that are sometimes associated with one letter such as /ks/ in *axis* [æksɪs] or /ju/ in *cube* [kjub]). Using these strings, a preliminary table of letter-sound associations was set up based on a one-to-one mapping of each letter onto the phoneme in the same position.

With strings which included more letters than phonemes, identifying the orthographic and phonological units that correspond to multiletter graphemes (*e.g.*, **air**, **ch** in *chair*), to mute graphemes (**s** in *isle*), or multi-phoneme units (/ks/ in *taxi*) was a more complex problem. It was treated, through successive passes, firstly applying a segmentation algorithm, and secondly by applying an association identification algorithm. The *segmentation algorithm* operated to attempt to produce for each entry in the corpus a version of the orthographic and phonological representations which was segmented into graphemes and phonemes. The goal was to produce a segmentation for which the number of graphemes was identical to the number of phonemes and in which each one-to-one mapping of graphemes onto phonemes matched the associations listed in a provisional grapheme-phoneme association table (*e.g.*, *cheep* mapped onto its phonological transcription [tʃiːp] with the grapheme-phoneme associations, **ch** to /tʃ/, **ee** to /iː/, **p** to /p/).

Whenever it was not possible to obtain a complete segmentation of the string with this algorithm, an *association identification algorithm* was given control instead. Grounded on the *rule learning algorithm* implemented in the *DRC* model of Coltheart *et al*. (1993), this algorithm aimed at signalling a single new mapping between an unmatched letter or letter sequence and either a phoneme or a silence, such that the addition of this association in the table would lead to the successful segmentation of that string [Note: a thorough description of the operation of this algorithm is provided in the online documents].

At the end of each pass, each association isolated by the *identification algorithm* was evaluated, each one in turn, and additions or deletions from the provisional association table were performed manually. Some of these decisions were obvious, as when for *card* [kaːd] and *lord* [lɔːd], the algorithm proposed the association **rd** to /d/ instead of the associations **ar** to /aː/ and **or** to /ɔː/. Others were not. In that case, previous works (Berndt *et al*., 1987; Collins & Mees, 1984; Haas, 1970; Roach, 1995; Venezky, 1970; Wijk, 1966-- add new refs) were used to suggest alternative segmentation solutions and lexical exemplars were generated for each one of the tables instantiating the different options. Decisions about what association to include in the table were made only after careful examination of these exemplars.

At this stage, any letter viewed as a functionless scribal insertion was isolated as a grapheme of its own (e.g., **b** in *debt*) and mute phonemes were added into the phonemic string when appropriate (*e.g.*, *isle* [aɪl] became **ai**.**s**.**l**.**e** [aɪ.=.l.=] ‘, where "=" stands for a silence). Letter or letter sequences found to be alternatively silent or sounded in analogous contexts were treated as distinct graphemes (for instance, **gh** was treated as a grapheme because it is silent in *plough* but pronounced /f/ in *tou*gh, *enough*, *cough*). Decisions about the segmentation of letter or letter sequences found to be systematically silent in a specific context were delayed till the analysis of contextual influences (for instance, at this stage **kn** was treated as a grapheme though it is well known that **k** is always mute in initial **kn**). It is only at the next stage that some of these segmentations were modified and codes were added to some graphemes to mark context-dependent associations.

These different steps were repeated until an association table was produced which we considered to be satisfactory and able to successfully segment every word in the lexicon [Note: For about 10 mappings that involved a particularly complex phonemic unit, segmentations were done by hand (**p.ur.e** [p.jʊə.◊])].

2. Identification of variations of pronunciations predicted by context, with grapheme disjunctions when appropriate

The objective of the next stage was to analyse the contextual influences. First, the phonetic codes corresponding to complex phonemic units were replaced by their constituent phonemes in order to facilitate the identification of the nature of the phonetic transformation. Then, the words exemplifying the associations of our initial table were scrutinised in order to identify and classify the contextual influences which take place in the English words. Some segmentations were modified and codes were added to some graphemes to mark context-dependent associations. Again, a semi-automatic approach was used where the role of algorithms was to ensure a systematic marking of the relevant contexts.

Identification of the relevant orthographic contexts was guided by the conditions listed in English studies (*e.g.*, Aronoff and Koch, 1996; Venezky, 1970). Most of them involved the surrounding letters: one letter to the right and/or one letter to the left and one letter on the right with one intervening letter (*e.g.*, soft pronunciation of **c** in chan**ce**, /z/ pronunciation of **s** between two vowels as in ga**s**oil; long pronunciation of a vowel followed by a final **e** with a single intermediary consonant, as in ch**i**me). Position was considered as a special case of orthographic context, with a space on the left or right marking an initial or final position, respectively.

Identification of the relevant phonological contexts was informed by numerous phonetic books and references (Chomsky and Halle, 1968, also site on the internet). They largely corresponded to the broad categories of phonological phenomena: *(i)* assimilative changes (when the voice, place or manner of articulation of a sound is transformed by the sounds that surround them in the word) with, for instance, assimiliation of place when /s/ occurs in front of a jod such as in “action” /ækʃn̩/) or assimilation of voicing when **z** occurs after **t** in *quartz* /kwɔːrts/, *(ii)* elision or complete assimilation (when speech sounds are fully assimilated to an adjacent sound, as in the loss of jod after palatalisation), *(iii)* transformation of sonorants into syllabic consonants (*e.g.*, when the high sonority consonants **n**, **l** and **r** occur in the sequences *on*, *an*, *al*, *ol*, *el*, and *il* in final position and in an unstressed syllable, they are associated with the syllabic consonants /n̩/, and /l̩/), *(iv)* epenthesis or insertion of a speech sound (e.g., schwa is inserted between **s** and **m** in *prism* /pr̩ɪzəm/). Association exemplars sorted by sound or pattern were produced to ensure that these phenomena were correctly coded as well as to assess whether other principles were at work in the words.

This analysis of contextual influences was then used to decide whether to divide multiletter graphemes.

Multiletter graphemes were disjoined every time one part of the grapheme could be viewed as silently marking an orthographic or phonemic influence on the pronunciation. In this case, the mute marking element was isolated as a separate grapheme with no specific sound value and the remainder of the cluster was left on its own, ready to receive label(s) marking contextual influences.

Multiletter graphemes were also disjoined when pronunciation could be seen as the result of the translation of each component letter with one of these components *having its realisation merged into an adjacent sound*, independently of the question of whether the sound transformation is synchronic (happens during processing) or diachronic (happened at some period during history). It appeared that a very large number of clusters often treated as graphemes in the literature could undergo disjunction according to this guideline. This was the case for (a) ***fully assimilated letters*** whose phonemic realisation becomes non distinguishable from a preceding grapheme with which it shares a single phonetic feature (*e.g.*, with *bp* in *subponea*, the voiced bilabial /b/ becomes the unvoiced bilabial /p/; with *mn* in autumn the alveolar nasal /n/ becomes bilabial nasal /m/), or (b) ***partially assimilated letters***, where two realisations differing by more than one feature influence each other in a way that creates a realisation intermediate between the ones of the letters (in *ti*, the /t/ and the /j/ amalgamate into /ʃ/, in *ng* or *nk*, the /n/ and the /k/ or /g/ amalgamate into /ŋ/; in non-rhotic dialects of British English in the sequence V+r followed by a consonant or at the end of a word, /r/ is merged into the realisation of the vowel to create centring diphthongs -- that is a vowel sound ending in /ə/ -- compare *far* and *fire* [faɪə]). This was also the case for (c) ***geminate and pseudo-geminate letters*** (*i.e.*, *cc*, *gg*, *sc*, *xc*, etc.), including those for which both component letters have the same realisation (**sc** translated by /s/ in *scene*; **cc** translated by /k/ in *accord*), those for which the second letter has a realisation different from that of the first (*e.g.*, **cc** is pronounced /ks/ when followed by *e*, *i*, *y*), and those for which the letters both have a different realisation (**sc** translated by /sk/ in *scandal*; **cc** translated by /ks/ in *accent*). Finally, (d) any ***geminate letter*** (*bb*, *dd*, *ff)* was disjoined on the ground that it changes next to nothing in the consistency estimates (the cluster *bb* is rewritten b_{ass}:*b*) but allows us to view as irregular the words where a geminate cluster is pronounced with two sounds (*e.g.*, when clusters are not levelled such as /nn/ realized as /n/ in compounds, as in *greenness*).

After the above disjunctions were applied, 26 multiletter sequences remained, involving vowel diphthongs as **ai**, **ou**, (oa, ea; ae, ee, oe; ai, ie; ei, oi, ui; oo; au, eau, eu, ou; ay, oy, ey; ew, ow, aw) and **consonant + h** sequences (ch, gh, sh, th, ph). We decided not to disjoin these sequences as they all had a phonemic value that was distinct from that of any component letter and that, therefore, couldn't be understood as a phonological phenomenon involving the separate constituents (*e.g,* the pronunciation /eɪ/ for **ai** cannot be predicted from general rules based upon *a* and *i* separately).

3. Adding of marks of context

Finally, labels were added to some graphemes to mark the contexts which determine a predictable variant in the pronunciation of the grapheme.

We considered three kinds of contexts associated with predictable variations in grapheme's pronunciation: (1) variant pronunciations ruled by neighboring letters (immediate or disconnected; *e.g.*, **c** has a soft variant when followed by **[e, i, y]**, as in *circus*; **s** is generally pronounced /z/ when it occurs between two vowels, as in *cosy*; a final **e** after a single consonant is a marker of the *long* variant of a preceding vowel, *rate*). The initial or final positions were seen as a special case of contextual influence, initial or final position being indicated by the presence of a space at the left or right of the grapheme (*e.g.*, in English, *y* is pronounced in different ways before a vowel and before a consonant or at word ending -- *yellow* vs. *byte* and *by*). (2), variant pronunciations ruled by phonological phenomena such as assimilation, which alter the pronunciation of sounds in sequences (i.e., the identity of surrounding phonemes). (3) variant pronunciation of vowels induced by stress (final *e* tends to be silent when not bearing primary stress – *the* but *race*; unstressed i in front of an unstressed vowel tends to have a /j/ pronunciation – *lion* but *billion*).

Since these analyses are meant to provide an evaluation of the power of strictly non-lexical rules, the only regularities we considered were those based on locally derivable information. Notably, morphological or syllabic boundaries were never considered in the marking (e.g., though *h* is marked for initial position, the **h** in *guildhall* was left unmarked). For the orthographic contexts, the marks simply considered the adjacent letters (or word boundary) independently of the phonemic realisation of these adjacent letters (**c** marked as *soft* in both *cell* and *mace*), as if transcoding proceeded in a moving window of no more than three letters. For the phonetic contexts, coding functions were more difficult to instantiate because we had no access to the underlying representation. For instance, in *nation*, there is no trace of the /j/ realisation in the phonemic representation so it had to be inferred from the orthographic and phonological contexts. It was acceptable to have recourse to intricate coding functions which take into consideration a large variety of information to allow a best guess of the underlying pronunciation of the grapheme or surrounding graphemes. For instance, /j/ realisation of **i** was marked by taking into consideration both the following vowel context and the word stress pattern. Nevertheless, we were careful to avoid completely circular definitions of the kind “if the pronunciation is a schwa, mark the grapheme as realised as a schwa”.

All marking codes were applied by a computer program that erased and then rewrote all marking codes in one pass. As described above, instead of marking the adjacent letters, labels give an indication of the expected pronunciation of the grapheme. For instance, for single vowels and multiletter vowels, a mark “`long`” is introduced to indicate a context in which an extended vowel is expected (i.e., one of /aɪ/, /aʊ/, /ɔɪ/, /eɪ/, /əʊ/). Different contexts can be considered, provided they cause similar changes in pronunciation. For instance, the mark “`long`” is added to single vowels based on the identity of the following characters: either a space indicating a final position or a sequence of a single consonant and a vowel. In both cases, the "`long`" mark is only applied when the vowel carries primary stress. The program is therefore organised into a succession of functions applying the marks for specific orthographic contexts or phonological phenomena (*i.e.*, transformation of consonants, deletion of consonants, tense value for vowels, glide values, reduction of vowels), ordered in such a way that a function takes into consideration the marks previously applied when appropriate. As apparent from the examples provided above, it is possible for multiple markings to be attributed to a single grapheme (e.g., the **i** in *nation* is marked as `yod+ass`).

3. Estimates of grapheme-phoneme reliability

The final step was the one of the counting. For each word in the corpus, the orthographic form segmented into graphemes and the phonological form segmented into phonemes were processed from left to right and each time a given grapheme-phoneme association was encountered, its frequency value was incremented, by one in the type count and by the word frequency value (*Cobuild* F/1 million values from *Celex* database) in the token count. From the basic association frequency values, several estimates of the predictability of the graphemes and of their pronunciation were derived. The ones provided in our table are: (1) grapheme-phoneme association frequency; (2) categorization of each association as regular or irregular; (3) grapheme-phoneme association probability; (4) number of phonemic variants of a grapheme; (5) dominance classification, with each grapheme-phoneme association of a given grapheme ranked by decreasing frequency values; (6) grapheme frequency; (7) uncertainty of the pronunciation of each grapheme as measured by grapheme's entropy. These estimates are described in full detail below.

Measures related to grapheme-phoneme associations

Association's frequency. Number of times each grapheme-phoneme association appears in the corpus.

Association's regularity categorisation. A value of 1 is given to the regular association, that is the most frequent pronunciation of a grapheme, and a value of 0 is given to any other pronunciation of the grapheme.

Association's dominancy ranking. The most common (or frequent) pronunciation of a grapheme receives a value of 1, the other phonemic variants are numbered from 2 to the number of phonemic variants of the grapheme, by decreasing association frequency values.

Association's probability/consistency. The probability that a given grapheme will be associated to a specific phoneme. This is computed as the association frequency value divided by the grapheme frequency value, that is, the sum of the frequency values of the different phonemic variants of a grapheme.

$$p(GPA) = \frac{GP\ association\ frequency}{Grapheme\ \ frequency}$$

This exactly parallels the ratio of summed friends to summed enemies used to estimate body-rime consistency (Jared, McRae, & Seidenberg, 1990). It takes the ratio of the number of friends relative to the total number of friends plus enemies, where a friend is a word with the same grapheme and the same pronunciation and an enemy is a word with the same grapheme and a different pronunciation.

Measures related to graphemes

Grapheme frequency. Frequency of occurrence of a grapheme.

Number of phonemic variants. The number of phonemic variants of a grapheme registers the number of ways a graphemes can be pronounced. The grapheme **au**, for instance, has 5 phonemic variants in English, /ɔː/ as in August, /ɑː/ as in aunt, /ɒ/ as in aussie, /əʊ/ as in mauve, /aʊ/ as in Frau.

Grapheme Entropy (H_g). The information statistics *H* introduced by Shannon (1948*a*, *b*)[1] (and exploited by Treiman *et al*. (1995) in a study of the consistency of the associations between orthographic and phonological rimes) provides the richest measure of the predictability of the pronunciation of a grapheme. The computed entropy value reflects both the number of phonemic variants of the grapheme and the probability values of these different variants. By contrast with the association probability value that is the property of isolated associations, *H* provides a measure of the uncertainty of a the pronunciation of the grapheme which reflects the distribution of probability of association for the all set of the phonemic variants of a grapheme. Its value is computed as follows:

$$H_g = \sum_{i=1}^{n} p_i \log_2 1 / p_i$$

where p_i is the probability of the i^{th} phonemic variant for each of the n phonemic variants of the grapheme. The value of H is minimal and equals 0 when there is a single phonemic alternative with association probability of 1, as for **j** to /ʒ/; it is maximal and equals $log\ n$ when the n phonemic variants of a grapheme are equally probable. To give an idea of the entropy scale on which the values fit, the maximal entropy value, corresponding to a situation where each grapheme could be associated with the same probability to any of the 44 phonemes of English is 5.46 (that is $log_2\ 44$). Its value is $log_2\ 3$ or 1.58 if the grapheme is associated to three equally probable phonemic variants, and $log_2\ 2$ or 1 when there are two variants of same probability and $log_2\ 1$ or 0 when there is a single variant. If there are two variants with respective probability values of .95, .05, H would be .29. Note that the entropy scale is reverse compared to other estimates since entropy is not a measure of the predictability of the pronunciation but a measure of its uncertainty: a value of 0 corresponds to a lack of uncertainty in the case of a unique (and thus completely predictable) pronunciation of probability value 1.

---

[1] For an introduction, see: Attneave (1959), Schneider (1995), "Information Theory Primer" (http://www.lecb.ncifcrf.gov/~toms/paper/primer/latex/index.html) and MacKay ,"A Short Course in Information Theory" (http://www.cs.toronto.edu/~mackay/info-theory/course.html).

Table xx. Illustration of the estimates provided in our tables with the data related to the graphemes AU

MEASURES RELATED TO THE GRAPHEMES AU						
Measures related to GP associations						
Phonemic variants	ɔː	ɑː	ɒ	əʊ	aʊ	
Example	<u>Au</u>gust	<u>au</u>nt	<u>au</u>ssie	m<u>au</u>ve	fr<u>au</u>	
Association frequency	105	6	6	3	1	
Regularity categorization	Reg.	Irr.	Irr.	Irr.	Irr.	
Dominancy ranking	1	2	3	4	5	
Association probability	.87	.05	.05	.02	.02	
Measures related to graphemes						
Number of phonemic variants	5					
Grapheme frequency	120					Σ assoc. frequency values
Grapheme entropy	.80					Σ [-p(assoc) * $\log_2$p(assoc.)]

N*ote*. By types estimates, extracted from the corpus of English monosyllabic words.

4. Measures related to a system of correspondences

For estimation of both graphemes and associations estimates, average values across the table cannot be directly compared, since they depend in part on the number of associations listed in the table (which differs for monosyllabic and disyllabic words) and on the redundancy of the coding (since a higher average of probability values will be obtained with redundant codings). Therefore, system estimates have to be produced which make it possible to compare different systems on an identical scale, despite possible differences in the number of associations. Such estimates can be obtained by weighting the entropy or probability value for each unit in the system (grapheme for the entropy measure and association for the consistency measure) by the probability of the association itself. As both measures are redundant and consistency estimates are easier to make sense of, only system consistency estimates will be provided in our summary statistics.

5. Estimates of a Word's grapheme-phoneme regularity and consistency

Estimates of the regularity or consistency of a word are inferred from the regularity or consistency values of the grapheme-phoneme associations which it consists of: A word is classified as regular if all its graphemes are associated with the most frequent (dominant) pronunciation of the grapheme, and as irregular if one (or more) grapheme is associated with a pronunciation that is not the most frequent one for that grapheme. An estimate of word consistency value is obtained by averaging the probability values of its constituent grapheme-phoneme associations across the string.

RESULTS

[Personal note: Make sure I use the following convention: Use system consistency when using association probability * association frequency. Use Average association consistency when average pb values across the table.]

Distribution of grapheme-phoneme ambiguities

The predictability of the grapheme-phoneme association system produced is summarized with different statistics: the number of associations in the system, the number of graphemes, the average number of pronunciations for each grapheme, and the average probability of the dominant pronunciation. At first statistics of the dominant correspondences (i.e., predicted pronunciation of a grapheme) are provided, with monosyllabic and disyllabic words grouped together.

Table xx. Statistics for the grapheme-phoneme associations in our tables

ALL ASSOCIATIONS	N° of graph.	N° of marks	N° of pronunciations		Dominant corresp. probability (type)		Dominant corresp. probability (token)	
			M	SD	M	SD	M	SD
All words	54	19	2.35	(1.64)	.90	(.14)	.90	(.16)
Vowels	28	11	2.56	(1.83)	.87	(.15)	.88	(.17)
Consonants	26	9	1.87	(.93)	.96	(.08)	.95	(.12)

Note. Statistics relating to the full set of associations in our tables: number of graphemes in the system; average number of different pronunciations of a grapheme; average of grapheme-phoneme associations probability for the dominant association (i.e., the regular correspondence) of each of the graphemes, with type and token estimates respectively, for vowels and consonants (38[2] phonemes, 22 phonemes for vowels et 20 for the consonants) and posyllabic words (43 phonemes in total, 26 for vowels and 24 for consnants) in our corpus. M indicates average values and SD standard deviation values.

Both type and token counts are provided as it has been claimed that high frequency ones are more likely to have irregular pronunciations than low frequency ones (possibly because low frequency words end up being pronounced regularly, over time; Wijk (1966), for instance, had reported that of the commonest 3000 words in his corpus as many as 21% violate the rules he had derived). The type count results from counting the number of words in the corpus that contain each grapheme-phoneme association. The token count is an estimate of the number of times the segment is found in running text (for instance printed newspapers). In practice, it is obtained by weighting each word in the corpus by the frequency of its occurrence in running text, as indexed by the written word frequency value provided in Celex.. no important difference was found between type and token counts in our analyses

Similarly, estimates are given for separately for vowels and consonants as, in English, pronunciation is reputed

[2] Some phonemes such as silence, /j/ (yod), /ə/ (schwa), and /w/ (semi-vowel) can be associated with both vowel and consonant graphemes (for instance, **re** to /ə^r/ in *cadre* and **y** to /j/ in *yellow*).

to vary substantially between graphemes of the two classes. In practical terms, a *vowel grapheme* is defined as a grapheme that begins with a vowel and a *consonant grapheme* as one that begins with a consonant. Our analyses conform to the usual picture of lesser predictability for vowels than for consonants. Vowels have a larger number of possible pronunciations and a lesser probability of dominant grapheme-phoneme association (i.e., if the grapheme is pronounced with its predicted pronunciation) than consonants. Indeed, it is common for vowels to have more than 5 possible pronunciations, both in stressed and unstressed syllables. For instance, the unstressed grapheme **o** has as many as 9 possible pronunciations in our system: /ə/ in *abbot*, /ɒ/ in *bloodshot*, /ɔː/ in *vendor* (but /ə/ in *splendor*), /əʊ/ in *rotate*, /ɛː/ in *framework*, silence in *colonel*, /ʌ/ in *somehow*, /ɪ/ in *pigeon*, /ʊ/ as in *woman* [Note: The letter *e* was also found to live up its reputation of a grapheme with an ever changing pronunciation; it is typically pronounced /iː/ in a prefix as in ***re**act*, alternatively /ɪ/ or /ə/ or ø in the past-tense marker *ed* as in *mov**e**d* or *want**e**d*. Yet, due to the use of marking codes, the variation in pronunciation for **e** was found to be less arbitrary than for the **o**]. For consonants, the maximum number of pronunciation is four. The grapheme **s**, for instance, is most commonly pronounced /s/ as in *adjust*, *bus* or *soap* but can also be pronounced /z/ as in *boys*, /ʃ/ as in *sugar*, or silence as in *fracas*. Only a very small number of consonants have multiple sound values. Most of them only have a sounded and silent version (the grapheme **p**, for instance, is rarely translated by a silence, as in *corps*).

Table xxx presented statistics about ambiguities in pronunciation for the regular associations only. Table xx presents the average of the probability values for all associations listed in the table.

Table xx. Statistics for the grapheme-phoneme associations in our tables

ALL ASSOCIATIONS	N° assoc.	Av. Assoc. probability (type)		Av. Assoc. probability (token)	
		M	SD	M	SD
All words	533	.89	(.42)	.86	(.43)
Vowels	402	.75	(.40)	.78	(.41)
Consonants	131	.97	(.46)	.92	(.47)

Note. Statistics relating to the full set of associations in our tables: number of grapheme-phoneme associations in the system; average of grapheme-phoneme associations consistency for all assoociations, with type and token estimates respectively, for vowels and consonants. M indicates average values and SD standard deviation values.

The consideration of irregular pronunciations causes a decrease in probability values for vowels only. These numbers could be obtained in two situations, one in which a few vowels have a high frequent irregular pronunciation and one in which most vowels have their most frequent one that clearly dominates over a large number of possible pronunciations. Figure xxx. plots the distribution of probability values in our system. It shows that for both vowels and consonants (but even more so for consonants), alternative pronunciations are in general of low probability, with most of the grapheme-phoneme associations having either very high probability (values above .95) or very low probability (values under .15), both for consonants and vowels.

DISTRIBUTION OF THE GRAPHEME-PHONEME ASSOCIATION PROBABILITY VALUES IN THE TABLES

A) MONOSYLLABIC WORDS **B) POLYSYLLABIC WORDS**

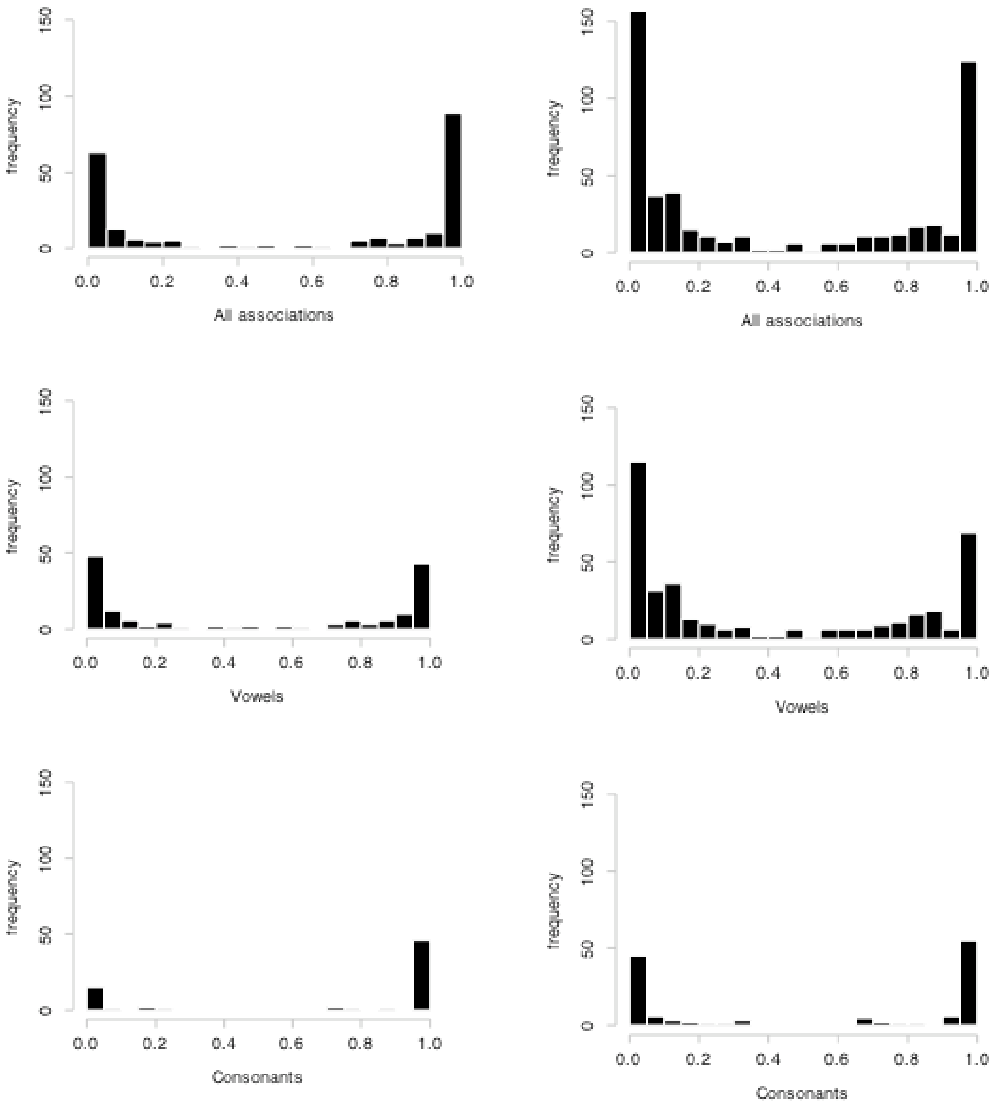


FIG. XX. Distribution of the grapheme-phoneme probability values (type values). Monosyllabic (a) and disyllabic (b) words

Mono vs disyllabic

Because of the important differences in the characteristics of one- and two-syllable words, it is important to evaluate differences in the predictability of pronunciation for these two sizes of words. To allow for the comparison between systems that count different number of associations, a system consitency estimate is used rather than the average association probability illustrated before. This association consistency estimates provides a more appropriate estimate of the ambiguity in grapheme-phoneme association by having, for each association the association probability value multiplied by its frequency of occurrence and the sum of these values divided by the sum of the association frequency values for each system being considered.

Table xx. Statistics about the association systems

ASSOCIATION SYSTEMS	Type		Token	
	All	V & C	All	V & C
Materials				
One-syllable words	.95	.88, .98	.88	.84, .90
Two-syllable words	.87	.71, .97	.86	.69,.96

Note. system grapheme-phoneme consistency, with type and token estimates respectively, for monosyllabic words and disyllabic words in our corpus.

For monosyllabic words, the pronunciation is fairly predictable. The probability of guessing a grapheme's correct pronunciation was, on average, .95, with average grapheme-phoneme consistency values of .88 for vowels and .98 for consonants). For two syllable words, substantially more inconsistencies are displayed in the correspondences between print and sound for two-syllable. It is the pronunciation of vowels, in particular, that is found to be far more uncertain in two syllable words than in one-syllable ones.

We mentioned earlier that the two factors of stress and morphological structure could be part of the reason for the reduced predictability of these pronunciations. In unstressed syllables, graphemes can receive a pronunciation that is different from any of the ones found in stressed syllables. For instance, in the word *patient*, both the *en* and the *t* receive a pronunciation not found in one syllable words. Another factor is morphological structure. In the course of history, when changes in pronunciation have occurred there was reluctance to change the spelling to reflect this change when this would have concealed important information about morphological relationships. As a result, discrepant pronunciations are found more frequently in polysyllabic words than monosyllabic words (for example, *sign* /saɪn/ and *signal* /sɪgnəl/ are pronounced with very different vowels but share the same spelling, which mark their common root).

Efficiency of grapheme-phoneme rules for deriving a word's pronunciation

Coltheart's (1978) intuition was that graphemes had a relatively restricted number of possible pronunciations, compared to the syllable and that this made it both the smallest and most efficient unit of conversion to keep track of for correctly translating a written word into its pronunciation. On the basis of the linguistic work available at the time, Coltheart (1978) proposed that the application of grapheme-phoneme rules would correctly predict the pronunciation of between 80 and 95% of (monosyllabic) words.

The statistics we have introduced indeed all suggest that the grapheme-phoneme associations are highly predictable. If the association consistency estimates closely mirror the proportion of regular words in the corpus, then we are well within the range defined by Coltheart. A problem however is that average grapheme-phoneme consistency of the association system is a reliable estimate of the percent of words correctly read by rules only for situations where only one segment is highly ambiguous in pronunciation, like one-syllable words only. Two-syllable words count two rather ambiguous vowel grapheme and this predict the percent of words correctly pronounced by rules to fall by a power of two of the average consistency of the vowel graphemes. With the consistency values we have, and with all consonant graphemes having association consistency values close to one, the following percent of regularly pronounced words would be expected: about 88% for one-syllable words, which have an average association consistency of .88 for vowel graphemes and close to 1 for each consonant; a low 51% (or.71^2) for two syllable words, which have an average association consistency of .71 for vowel graphemes.

To demonstrate that our average grapheme-phoneme association consistency estimate of .90 (all words, all types graphemes) effectively translates into more than 80% of words correctly pronounced by rules, it is necessary to provide word regularity and consistency estimates. *Word regularity* refers to whether or not all the correspondences in a word are regular. *Word grapheme-phoneme consistency* refers to a continuous estimate of the probability with which the graphemes in the word are pronounced with their respective phonemes, computed as the sum of the probability values associated with every association in the word divided by the number of graphemes.

Word grapheme-phoneme regularity and consistency

The number of words correctly pronounced using the grapheme-phoneme rules of our system are well within this range for one syllable words. In monosyllabic words, only about 12% of the words are not pronounced correctly by grapheme-phoneme rules. As traditionally found, irregular words tend to be of high frequency, causing an important drop in the number of words accurately read in a running text. Given that the association system contain substantially more inconsistencies in the correspondences between print and sound for two-syllable words than for one-syllable words, we can can expect a drop in the number of two-syllable words that can be pronounced correctly by grapheme-phoneme rules. The question then is, How much of a drop will be found?

Table xx. Statistics for the word's grapheme-phoneme predictability in pronunciation

ASSOCIATION SYSTEMS	Naming accuracy	*Word's av. grapheme-phoneme consistency*		*System consistency*	
	%words	*Type*	*Token*	*V+C*	*V, C*
Full corpus	66%	0.89	0.79	.89	.75, .97
One-syllable words	88%	0.94	0.79	.95	.88, .98
Two-syllable words	56%	0.87	0.84	.88	.71, .97

We indicated above a dramatic drop in predictability from one to two syllable words could be observed as two syllable words include twice as much highly ambiguous graphemes (vowels). Here, the predictability falls by more than 30% from one- to two-syllable words, *i.e.*, from 88 to 56%. This shows that the lesser predictability of vowels make the pronunciation of bisyllabic words more ambiguous than conventional models would expect.

Coltheart (1978), in particular, had suggested that the grapheme was an efficient unit for speech to sound translation. The estimates in Table xxx suggest that it is not that so once polysyllabic words are included. Even more so that these estimates are based on a grapheme-phoneme association system to which various kinds of information are known that wouldn't necessarily be readily available to the grapheme-phoneme conversion route described in *DRC*.

Firstly, word stress information is accurately encoded, on the basis of the phonological representation found in Celex, and it is used to identify variant vowel pronunciations in stressed and unstressed syllables. In *DRC*, the conversion system is supposed to be hermetic to any lexical influence, including stress. This means that the system has to use strictly non-lexical rules for stress derivation. To evaluate the drop in predictability associated with a lack of accurate information about stress, we provide estimates of grapheme-phoneme consistency for different systems in which stress is non-lexically derived.

Secondly, our study offers no direct support to the hypothesis of *DRC* that "grapheme-phoneme correspondences" capture a large number of print-to-sound regularities because we adopted a segmentation into graphemes and phonemes that is different from the one adopted for one syllable words in the DRC model. In our analyses, many regularities are captured in marks of context adjoined to the grapheme rather than in graphemes themselves because of our option to split as many multiletter graphemes as possible into a marked grapheme followed by a silent grapheme. In a *DRC*-like system, more regularities are captured in graphemes themselves because of the option to only have the clusters which can be matched onto a sounded phoneme treated as graphemes. In many cases, the identity of the unsounded letter come to replace the mark that was previously attached to the preceding grapheme. The impact of these differences between our classification into graphemes and phonemes and the one adopted in the DRC is difficult to evaluate, a priori. To evaluate the power of a *DRC*-like grapheme-phoneme system, we produced an alternative segmentation of the orthographic string, with silent letters being systematically part of a multiletter grapheme.

Finally, the estimates we present do not capture inconsistencies in segmentation caused by difficulties in deciding how to segment a word into graphemes (*i.e.,* deciding whether the sequence “tio” in “action” should be parsed as three graphemes, “t”, “i”, and “o”, as two with either “ti.o” or “t.io“, or as the single grapheme “tio” [Note: each one of these segmentations can be found in published studies]). Inconsistencies in segmentation are expected to be fairly rare in our material, as the number of multiletter graphemes is very small (for consonants, the clusters ending with “h”, and for vowels, the clusters associated with dipthongs). However, it is possible that in a DRC-like system, a large number of segmentation ambiguities will occur even before translation begins. In such a system, many multiletter graphemes need to be added to the grapheme table in order to obtain a perfect one-to-one match between the number of graphemes and phonemes (e.g., “ti” in “action”) and this create the potential for errors as this will force the processing of that cluster as a grapheme in any word it occurs. This is simply because it is impossible for the print-to-sound conversion system to know, on the basis of spelling alone, that in a word like “action” the cluster “ti” is pronounced with one phoneme and in a word like “tin” it is pronounced with two; information about phonemes is not available before the conversion system has successfully processed the word. To evaluate the way segmentation ambiguities impact on the accuracy of grapheme-phoneme conversion, an alternative segmentation of the string was produced, corresponding to a guess of the segmentation using a DRC-like segmentation algorithm. The percent of words correctly translated by rules was computed anew using the guessed segmentation.

Efficiency of a DRC-like system

Importance of knowing the exact stress value of each grapheme

To evaluate the extent to which knowing the lexical stress value has an impact on the proportion of words read correctly by grapheme-phoneme rules, estimates were computed anew, using systems of associations in which the exact information about stress had been erased and replaced by information obtained on the basis of strictly non-lexical, local, rules of stress assignment. Two new systems are described. The first one has for default rule that, in a two-syllable word, the first vowel is stressed and the second is unstressed. The second has no rule for non-lexical stress assignment and treats every vowel in the word as stressed.

It may come as a surprise that we do not put to the test any other option. However, this is not because we did not try to identify non-lexical rules for stress. Numerous resources were consulted (notably, Chomsky and Halle, 1968, was analysed) and in-depth analyses were conducted on our association system to try to define efficient rules for stress assignment on the basis of spelling alone. This was without success. The rules became very complex and in most cases they did not lead to any significant increase in predictability compared to the default rule of stress on the first vowel.

Table xx. Non-lexical rules for stress assignment

TWO-SYLLABLE WORDS	Naming accuracy	# Irr. Segm.	System consistency	
	%words	M	V + C	V, C
Non-lexical rules of stress:				
First vowel is treated as stressed	50%	2.11	.85	.64, .97
Any vowel is treated as stressed	35%	2.65	.78	.55, .91

Estimates are provided for two-syllable words only, using type values.

An important decrease in the consistency of grapheme-phoneme associations is observed when lexical stress is not available. When stress is assigned on the basis of strict left-to-right position, with the first vowel in the word being treated as stressed and any of the following ones as unstressed, about 50% of the words are pronounced correctly. In the total absence of any guess on stress information, this value fell to 35%. A large part of this decrease can be attributed to difficulties in reducing vowels appropriately.

Contribution of strictly context-free generalizations

In our system, many regularities are captured by marks of context rather than by grapheme-phoneme associations. To evaluate the power of strictly grapheme-phoneme rules, a distinction was introduced between *context-free* generalizations – which apply regardless of what surrounds each grapheme (*e.g.*, the grapheme **ss** has /s/ for its regular pronunciation, as in *la***ss***o*) – and *context-sensitive* generalizations – which depend on adjacent elements and only apply in a certain context. Context-free and context-sensitive generalizations are estimated using system consistency values and word naming accuracy using two different orthographic reprsenttations, the original one, with marks of context, and a new one with marks of context stripped off.

Table xx. Impact of marks of context on naming accuracy and system consistency

TWO-SYLLABLE WORDS (NON-LEXICAL STRESS)	Naming accuracy	# Irr. Segm.	System consistency	
	%words	M	V + C	V, C
Our system				
Marks of context used	50%	2.11	.85	.64, .97
Marks of context dropped	8%	4.12	.57	.40, .67

Estimates are provided for two-syllable words only, using type values.

Given that we opted for a system that accepts mute graphemes and introduces complex marks of context (orthographic and phonological contexts, position in word), it does not come as a surprise that correct pronunciation is found to depend heavily on contextual information. When marks of contexts are stripped from the words, pronunciation accuracy is as low as 8%. For both vowels and consonants, a knowledge of the letters and and/or phonemes that surround the grapheme contributes to a disambiguation of the pronunciation. Figure xxx shows that context helps disambiguate the pronunciation of a large number of graphemes for both categories of graphemes. There is a decrease in average association consistency of more than .20 for 8 of 25 consonant graphemes and for 11 of 27 of vowel graphemes. The difference is that for consonants, a knowledge of the adjacent elements is usually sufficient to completely disambiguate the pronunciation: for vowels, many pronunciations remain downright capricious.

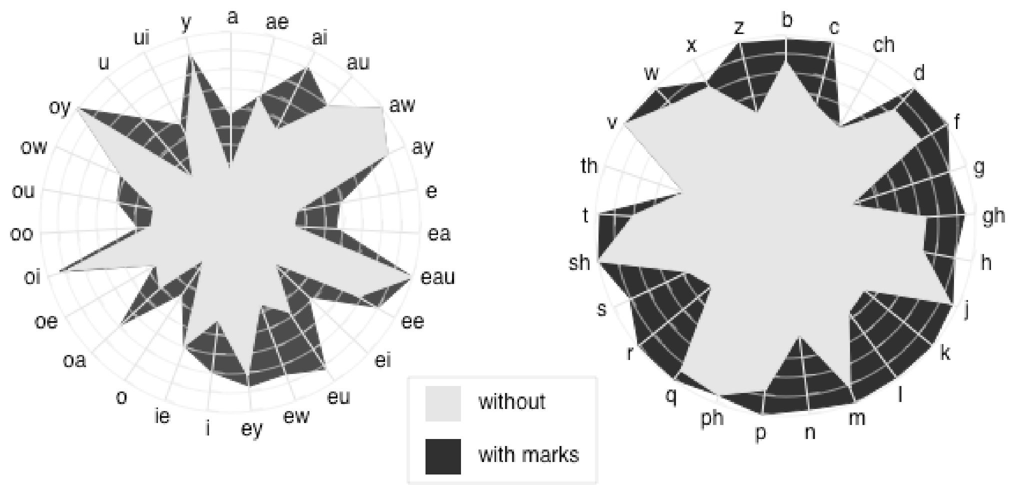


FIG XXX. A spider graph is used to represent the difference in average association consistency values (weighted by their frequency of occurrence) for vowels and consonants. The length of each axis is set to 1 and the contribution of marks of context is represented by the difference between the area in black (■) and the one in light grey (■).

These statistics cannot be used to make any strong claim about the generalization power of strict grapheme-phoneme rules in a DRC-like conversion system. We already made it clear that context is encoded very differently in our system compared to the grapheme-phoneme conversion system found in DRC. Because our marking codes capture many of the contextual influences that would be captured by relying on bigger multiletter units in a DRC system, we also generated a DRC-like association system. For each word in our corpus, a new segmentation of the words was produced by the following guideline found in Coltheart and colleagues (2001): the fundamental unit of pronunciation is the phoneme and graphemes are units that map onto them; mute graphemes are forbidden. Accordingly, our parsed representations were modified so to have any unsounded letter inserted into the grapheme that immediately precedes it in the word (or into the grapheme that follows for an unsounded letter in word initial position). For a word like “through” /θru:/, this caused a change of segmentation from “th.r.ou.gh” to “th.r.ough”. [Note: In DRC, contextual influences are coded with the identity of the adjacent letter, rather than a label for a category of change, attached to the grapheme. For the results reported here, our marking system was used to examine difference between marked and unmarked systems].

Table xx. Impact of marks of context on system consistency

TWO-SYLLABLE WORDS	Naming accuracy	# Irr. Segm.	System consistency	
	%words	M	V + C	V, C
DRC-like system				
Marks of context used	52%	1.72	.83	.63, .95
Marks of context dropped	43%	2.01	.79	.57, .92

Estimates are provided for two-syllable words only, using type values.

Expectedly, marks of context are found to be have less of an impact in a DRC-like system than it had in our association system. Still, it is found that a futher 7% decrease in naming accuracy occurs between a system that capture emergent regularities with marks of context to capture and a DRC-like one that relies on large multiletter units. Then, an important issue, in our view, is that this 7% drop occurs with an association system that is made less economical because of the preference that is given to larger units. Any system which does not allow a grapheme to be mapped onto a silence is deemed to count more graphemes than one that does allow this. This is because in the former system, what would have been multiletter graphemes are recoded as composed of a sounded grapheme and a silence. With the letter "e", for instance, the 10 clusters **ce**, **ed**, **de**, **el**, **le, en**, **ne**, **es**, **se**, **gue** (as in *ocean*, *paced*, *made*, *easel*, *able*, *shorten*, *cane*, *races*, *danse*, *rogue*) can easily be coded with only 7 graphemes (**c**, **e**, **d**, **l**, **n**, **s**, and **gu**). The exact difference in number of units and associations that need to be represented is summarized in Table xx. What these statistics reveal is that allowing for a grapheme to be mapped onto a silence results in a more efficient system, one that captures the greatest part of the print-to-sound relations of the language with the smallest number of units, that is, one with the minimum redundancy.

Table xx. Efficacy of different association systems

TWO SYLLABLE WORDS (NON-LEXICAL STRESS)	WORD NAMING ACCURACY	#GRAPH	#MARKS	#GRAPH+M	#ASSOC.
Method					
Our system	50%	54	19	223	519
DRC-like system	43%	231	0	285	540

Efficacy of association systems based on different coding options, with a comparison of our system (mute graphemes and marks of context) or a DRC-like system (large multigrapheme units).

Importantly, it was also found that the regularity categorization based on our table compared better with the regularity classification of strings typically used in regularity studies. Using our original association system, words such as *gross*, *pint*, *lose*, *sew*, *steak*, *broad*, *gauge*, *wool*, *aisle*, *doll*, *break*, *sword*, *touch*, *brooch*, *deaf*, *foot*, *says*, *some* and *soot* (which are all classified as irregular in their grapheme-phoneme correspondences in more than ten experiments) are classified as irregular by our association tables. Words generally listed as regular on the grapheme-phoneme correspondences but inconsistent on the body-rime segments such as *doll*, *boot*, *brow*, *cave*, *dash*, *hive*, *hoot*, *jowl*, and *lass* are all categorized as regular following our table. [Consulted references: Andrews (1982), Baron and Strawson (1976), Barron (1980, 1981), Bauer and Stanovich (1980), Berent and Perfetti (1995), Coltheart (1979), Glushko (1979), Jared *et al.* (1990), Kay and Lesser (1985), Kay and Marcel (1981), Paap and Noel (1991), Parkin (1982, 1984), Parkin and Underwood (1983), Parkin, McMullen, and Graystone (1986), Rosson (1985), Seidenberg (1985), Seidenberg *et al.* (1984), Stanhope and Parkin (1987), Waters, Seidenberg, and Bruck (1984), Waters and Seidenberg (1985).] In contrast, the DRC-like system provides a regularity categorization that produce a categorization of regular and irregular materials different from the one found in the literature. The *DRC*-like system classified as regular any string that is regular on either the grapheme-phoneme or body-rime associations (how come?). For instance, *pint* is erroneously categorized as regular and every regular but inconsistent [need to explain this] word is erroneously classified as irregular.

Impact of segmentation difficulties

A last issue when estimating the generalization power of grapheme-phoneme rules is that a system that starts with the word presented as a sequence of graphemes, eventual ambiguities in deciding how to split the sequence into graphemes are de facto hidden. When reading a real word, however, an incorrect pronunciation could be produced whenever a cluster is wrongly identified as a grapheme (or, more specifically, when a cluster matching one of the grapheme units found in the grapheme-phoneme association table). This can happen, for instance, when a morpheme or syllable boundary forbids the clustering of letters. For instance, both *react* [rɪeɪkt] and *dream* [drɪm] would be correctly pronounced using grapheme-phoneme rules that are applied on strings that have been previously segmented into graphemes. However, pronunciation errors can be expected for *react* [rɪeɪkt] if a segmentation algorithm had to be applied first to split the sequence into graphemes. This is because the sequence “ea” is usually treated as as single grapheme, as in *dream* or *reach*. Importantly, as the contrast between *react* and *reach* shows, orthographic context is often of little help to decide of the correct segmentation.

These segmentation ambiguities cannot be expected to have much of an impact when only one-syllable words are analysed, simply because no syllable (by definition) or morphological boundaries (except for “ed” or “s”, which mark past-tense or plural) can be found in such words. However, they can certainly be expected to impact on the pronunciation of two-syllable words, as 10 of the multiletter graphemes listed of our association system are found to have inconsistent segmentations (*e.g.*, **ai** in 'archaic', **ea** in 'create', **ei** in 'deity', **ie** in 'client' or ‘acquiesce', **oa** in ‘coarticulation’; **oe** in 'poet' or ‘poetical', **oi** in 'heroic', **oo** in 'cooperate', **ui** in 'ruin'; ***ph*** in *shepherd* and *morpheme*). They can be expected to have an even greater impact in a DRC-like system which counts many more multiletter graphemes with discrepant segmentations.

Two-syllable words	***%words incorrect segmentation***	***Multiletter graphemes in the rule system***
Original system	2%	**26** (ae, ai, au, aw, ay, ch, ea, ee, ei, eu, ew, ey, gh, ie, oa, oe, oi, oo, ou, ow, oy, ph, sh, th, ui, eau)

DRC-like system	68%	**223** multiletter units (**two letters long:** ae, ah, ai, al, an, ar, as, au, aw, ay, bb, be, bt, cc, ce, ch, ci, ck, cq, ct, cz, db, dd, de, dg, dh, di, dj, ea, ed, ee, ei, el, en, eo, er, es, et, eu, ew, ey, fe, ff, ft, ge, gg, gh, gi, gm, gn, gu, he, hi, ho, ie, il, in, ir, is, ke, kh, kk, kn, ld, le, ll, mb, me, mm, mn, mp, nc, nd, ne, ng, nk, nn, nw, oa, oe, oh, oi, ol, on, oo, or, ot, ou, ow, oy, pb, pe, ph, pp, ps, pt, qu, re, rh, rr, sc, se, sh, si, ss, st, sw, te, th, ti, ts, tt, tw, tz, ue, ui, ul, uo, ur, ut, uy, ve, vv, wh, wr, xc, xh, xi, xs, ye, yr, ze, zz; **three letters long**, ach, ain, air, ais, ait, and, are, awe, ces, che, cqu, ddh, dge, ear, eau, eer, eir, ere, eue, eur, ewe, eye, ffe, ggi, gne, gue, hei, her, hou, i-e, igh, lle, mne, ngh, nne, oar, oer, olo, oor, ore, ort, oul, oup, our, owe, ppe, pti, que, rri, sch, sci, seh, shi, sse, ssi, sth, tch, the, tte, uoy, ure; **four letters long**: aigh, augh, dgeh, eigh, heir, ngue, orps, ough; **vowel or vowel cluster + final e**: a-e, e-e, o-e, u-e, y-e; ai-e, ar-e, au-e, ea-e, ee-e, ei-e, er-e, ie-e, ir-e, oa-e, oi-e, oo-e, or-e, ou-e, ui-e, ur-e)

Of course, this does not necessarily predict that many errors in pronunciation. With silent letters, it does not matter much whether the letter ended being attached onto the previous grapheme or onto the next one. To evaluate the performance of the DRC-like system on real words. Only 7% of the words had their predicted pronunciation, obtained using the most frequent association for each grapheme, matching the correct pronunciation. This indicates that segmentation ambiguities clearly impact on pronunciation performance.

DISCUSSION

[rewrite these two paragraphs in a softer tone]

In their recent papers, Coltheart and colleagues (2001) used the fact that their has a degree of fit with readers' performance higher than competing models that the DRC hypotheses are more valid than those of the alternative models. They also used the lack of unambiguous empirical evidence in favour of more complex hypotheses to support the rather basic (if not simplistic) hypotheses made in the DRC model. For instance, "*One refinement of dual-route modeling that goes beyond DRC in its current form is the idea that different GPC [Grapheme-Phoneme Correspondences] rules might have different strengths, with the strength of the correspondence being a function of, for example, the proportion of words in which the correspondence occurs. Although simple to implement, we have not explored the notion of rule strength in the DRC model because we are not aware of any work which demonstrates that any kind of rule-strength variable has effects on naming latencies when other variables known to affect such latencies such as neighborhood size (e.g., Andrews, 1992) and string length (e.g., Weekes, 1997) are controlled.*"

An issue we would like to raise in this discussion, however, is what value a high degree of fit has when aspects of the models are not clearly motivated by experimental data. Here, we will use both the results of our analyses and the intuitions we gained when conducting them to evaluate whether there is evidence that unambiguously supports any of the theoretical choices made in DRC. We will discuss in turn the nature of the knowledge of print-to-sound regularities and the nature of the print-to-sound conversion process.

1. On representations

In support of the efficacy of grapheme-phoneme rules, Coltheart (1978) put forward a picture of English which was characterised by considerable regularity, with an extremely large proportion of words which could be read correctly with a knowledge of a grapheme's most frequent pronunciation. [Note: inside a range of 80-95%: 80% for Hanna and Hanna, 1959; 87% for Forbes, 1964; 90% for Hanna and Hanna, 1965; and 90-95% for Wijk, 1966]. In contrast, advocates of spelling reform, such as Dewey or Shaw have always insisted on the large number of alternative grapheme-sound or sound- grapheme correspondences that exist in the language; they have provided a picture of written English that is one of extreme complexity that does much to explain the difficulties that learners and users of every kind experience in decoding and encoding the alphabetic forms of English words.

Both pictures were reflected in our analyses of the grapheme-phoneme associations in English monosyllabic and disyllabic words. A large number of English words can be translated into speech using only rules of print-to-sound translation, without the use of any information from a mental dictionary or lexicon. However, the system of rules that is needed to achieve that level of performance is a lot more complex that one of strictly grapheme-phoneme rules. This is true even on the apparently very predictable monosyllabic words. When strict grapheme-phoneme correspondences were used, about xx% of monosyllabic words were translated correctly. It is only when information about the grapheme's position and surrounding context was taken into account that the percentage rose above the lower limit of 80%, with xx% of the words correctly translated on the basis of the correspondence rules. Further, the descriptive analyses revealed that with a grapheme-phoneme rule system, performance is far from being equivalent for monosyllabic and disyllabic words. When disyllabic words were introduced, the percentage of words correctly translated was well under the 80% figure when using strict grapheme-phoneme rules, even when marks of context were included. And these results are with a system of correspondences chosen to capture a maximum number of regularities: they do not necessarily reflect the performance of a DRC-like system. When strictly grapheme-phoneme rules were applied by referring to a different rule system, one whose segmentation captured contextual influences by systematically adjoining mute markers to the appropriate letter clusters, only about xx% [DRC system no marks of context] of the words were pronounced correctly.

[Importance of marks of context for correct pronunciation – not PRIMARILY graphemes]

The necessity of considering information outwith a grapheme correspondence to decide of its pronunciation was acknowledged by Coltheart and colleagues. In the description of the runnable version of their model (Coltheart et al., 1993; Coltheart et al., 2001), they indicated that "[s]ome of the GPC rules are context-sensitive -- for example, *c* is translated to /s/ when the following letter is *e*, *i*, or *y*, and is otherwise translated to /k/. Some of the GPC rules are position-sensitive -- for example, there are three rules for the grapheme *y*. In initial position, it is given the phoneme /j/, in medial position it is given the phoneme /ɪ/; in final position it is given the phoneme /aɪ/".

What was not acknowledged however is that such contextual rules are central to DRC's the level of performance. Without contextual rules, strict grapheme-phoneme correspondence rules are found to have poor generalization power. This seriously undermines the claim made in DRC that grapheme-phoneme correspondences are all that is needed to convert a large proportion of words into their phonology. At the least, it questions the theoretical coherence of the idea of a rule system based only on grapheme-phoneme correspondences, especially since the assumed high degree of predictability of grapheme-phoneme relations was the crucial argument for the proposition that conversion is based on translating graphemes into phonemes. This could be questioned even for monosyllabic words, and it is clearly not the case for disyllabic words. This forces us to consider what arguments remain for the hypotheses introduced by Coltheart and colleagues. Several aspects will be considered in turn. (1) The adequacy of proposing that no unit other than the grapheme is encoded in the system, when that unit alone does not lead to the expected level of performance. (2) The value of proposing that grapheme-phoneme correspondence are central to conversion, given that context needs to be encoded anyway. Coding letters in context would produce nearly the same results as coding grapheme units, so is there any empirical evidence that we actually activate grapheme-size units when converting print into phonology? (3) xxx

No argument against units larger than the grapheme

Why should context be limited to one letter to the right or left?

In the preceding discussion, we have focused on grapheme-phoneme correspondences and regularities in the grapheme and its immediate context. However, what our analysis made clear was that the reality of language is one of variations in print-to-sound relations within the breadth of general constraints (cf. Dell *et al*., 2000). Some of these constraints are imposed by one adjacent letter, as described above. However, it is well known that numerous other constraints, outside the scope of one letter to the right or left (or an adjacent blank space) used by Coltheart *et al.* (2001) in their rule-system for monosyllabic words, also play a role.

Empirical studies have revealed an independent influence of contextual constraints across a number of constraints (Shallice *et al*., 1983, in English, and Peereman, 1991, in French for local context). An even more imposing body of work has argued for an influence of the regularity (or rather consistency) of the body-rime relation on the readers performance. This includes the finding that the pronunciation patterns are far more stable within rime units (which refers the vowel peak of the syllable plus any sounds following the peak within the syllable) than they are within grapheme units (Adams, 1990; Stahl, 1992). This is the case for thirty-seven rimes (e.g., *at*, *ack*, *ap*, *ash*, *eat*, *op*, *ing*) appearing in over 500 different words that children commonly see in early grades (Adams, 1990; Wylie & Durrell, 1970; see also Kessler and Treiman, 1995, Kessler and Treiman, 1997, for information about the way that final consonants influence the pronunciation of the vowel in monosyllabic words). A large body of experiments have demonstrated that pronunciation of the rime unit is highly accurate, along with numerous pieces of evidence in favour of an influence of body-rime regularity and consistency on reader's performance.

Unexpectedly, Coltheart and colleagues decided not to consider contexts associated with body-rime regularities in their latest computational model (Coltheart et al. 2001) . They justified this decision on the basis of the discovery of a confound between the manipulation of consistency and the use of strings with multiletter graphemes of opaque pronunciation (*e.g*., *ph* where the pronunciation of the grapheme is different from the one suggested by the letters of which it is composed and which, because of this, activates the wrong phoneme during letter-by-letter translation). Cotlheart *et al*. (2001) questioned the necessity to introduce body-rime correspondences; the *DRC* model reproduced the results which had been presented by Jared (1997) as evidence for the participation of body units in the conversion. As a result, they asserted that "[…] the body of experiments showing effects of consistency on reading aloud are compatible with the DRC model despite the fact that this model contains no level of representation specific to orthographic bodies" (Coltheart et al., 2001, p. xxx). Their conviction was seemingly so firm that they even avoided coding body-rime regularities that could be captured with context-sensitive rules in monosyllabic words (e.g., the sequences *oot*, *ead*, *ook*, *ood* are described as quite regular by Aronoff and Koch, 1996, and could all be coded with contextual rules; *ook*, for instance, is nearly always pronounced as /ʊ/ when followed by ‘k’ -- *book, cook, hook* – while **oo** is normally pronounced as /uː/ -- *boot, broom, cool*).

In our view, it is to take quite a shortcut to use the results of a single study to assert that body-rime consistency does not influence reading performance. Jared's study is only one of a large number of studies which reported body-rime consistency effects. Even if it was proved that there was a confound in the material of that study, this is not proof of an absence of consistency effect when such a confound is avoided. In fact, in a more recent study comparing consistency and length effect in German and English, Perry *et al*. (2000) reported evidence for a consistency effect that cannot be interpreted as a grapheme complexity effect, and which is not reflected in the performance of *DRC*. Given the still lower predictability of strictly grapheme-phoneme correspondences in disyllabic words, it is certain that regularities of larger segments need to be considered in order to achieve high levels of performance.

Why reject the coding of units larger than the grapheme?

What would be the best way to encode constraints on larger segments?

In the eighties, evidence for the influence of body-rime regularity and consistency on readers' performance was interpreted as evidence for the psychological reality of the body and forced the hypothesis that sublexical conversion was based on segments of different sizes (these models were grouped by Carr and Pollatsek under the label of *parallel coding systems*; *e.g*., Coltheart, 1985; Harris & Coltheart, 1986; Kay & Bishop, 1987; Norris & Brown, 1985; Paap, Noel, & Johansen, 1992; Patterson & Morton, 1985; Shallice, Warrington, & McCarthy, 1983; Shallice & McCarthy, 1985). In these *parallel coding systems* of multiple levels of association (which constituted either close or remote extensions of the classical dual-route theory), both the grapheme and the body were systematically listed as conversion units. Some studies even posited conversion units of any size, from letters to subsyllabic units, to morphemes, to words (e.g., Shallice, Warrington, & McCarthy, 1983; Shallice & McCarthy, 1985).

Interestingly, the need to capture regularities of segments larger than the grapheme does not necessarily imply the need to recognise units of varying sizes. Nothing prevents the breadth of constraint or the continuum from being captured within a grapheme-phoneme rule system that includes context-sensitive rules (like in *DRC*). It is only necessary to assume that context-sensitive rules can, sometimes, take account of the two, three, or even four subsequent consonant letters till the end of word. Context-sensitive rules could also be extended to incorporate a knowledge of the pronunciation of the rimes of monosyllabic words or the one of sequences of consonant letters that are highly regular (such as *tch*, always pronounced as in *match*), in the form of wide contextual rules.

However, there is little reason to avoid introducing units which code clusters larger than the grapheme in a model that elsewhere contains detectors for stress assignment affixes. Incontrovertibly, Rastle and Coltheart’s (2000) algorithm for non-lexical grapheme-phoneme translation, stress assignment, and vowel reduction imposes the identification of orthographic components within words which regularly serve as stress-placing affixes (suffixes *—ness*, *—ess*; prefixes *ab—*, *pre—*, etc.). In this study, Rastle and Cotlheart (2000) investigated the interaction of stress regularity and frequency in disyllabic letter strings in readers and in the DRC model, where stress regularity was defined by various versions of a non-lexical stress assignment algorithm. With the constraint of a non-lexical stress assignment system, the predicted interaction of frequency with stress regularity could be obtained only when affix detectors were included; it was not found when the vowel of the first syllable was treated as stressed, without stress detectors. They concluded that affix detectors must be introduced if the hypothesis of non-lexically assigned stress was to be maintained. Strangely enough, the authors declared that "*this system falls entirely within the principles of the current DRC model. Currently, the nonlexical route of the DRC model relies on rules which translate graphemes to phonemes. Thus the nonlexical route already contains a store of graphemes, a store of instances in which letters combine to form graphemes. The nonlexical system describe here contains this store, and also a store of instances in which letters combine to form affixes*".

It would certainly be more coherent if the hypothesis of strictly grapheme-phoneme rules was to be sustained without the introduction of affix detectors for stress assignment. It has to be made clear, however, that this cannot be done. Certainly, the fact that English speakers generally agree on where stress should be placed in individual words was what led Chomsky and Halle (1968) to foster the belief that the stress system has some method, and to search for a set of procedures which reliably determine stress in English. However, the procedures they came up with largely depends on information which does not necessarily show up in spelling, namely the number of syllables in the word (in disyllabic words, it is generally the case that the first vowel is stressed and the second unstressed) and the syllable's phonetic characteristics (as described by Chomsky and Halle, 1968, p. 29, stress generally falls on syllables that contain phonetically strong clusters); the grammatical category to which the word belongs (noun, verb, adjective, etc.; *e.g.*, *Import* as a noun, vs. *imPORT* as a verb); and morphological structure (a large number of suffixes act as stress attractors, such as *ain*, *esce* in verbs, or *esque*, *ique*, *ette*, *ee* in nouns; see Kelly, Morris, & Vereikia (1998). Our own analysis confirmed Rastle and Coltheart's (2000) study and made it clear that a system for (non-lexical) stress assignment cannot be efficient without introducing segments larger than the grapheme. The identification of affixes, notably, is crucial for correct stress assignment as the stress is typically displaced if the first syllable is made of a prefix (in syllabic structures which requires stress on the first syllable; e.g., *proPEL*). [Note: a study by Baker and Smith (1976) strongly suggest a knowledge of each word's stress pattern and a knowledge of rules to guide them in how to assign stress correctly. They found that participants appeared to draw both on analogy with a similar real word and on pronunciation by phonological rule when reading aloud polysyllabic nonsense words.]

But why would a system have "detectors" for orthographic segments as large as four letters for identification of potential affixes during stress assignement, but "detectors" only for graphemes for print-to-sound conversion? Two arguments were introduced in the seminal paper of 1978. A first argument was the impossibility, alleged by Coltheart, of carrying out syllabic segmentation with no information about the pronunciation of the string (compare, for example, *lem-on* and *de-mon*). This was, however, in a context where only two candidates were considered as units of representation intermediate between the grapheme and the syllable. This argument is of dubious relevance in the context of subsyllabic units such as the rime. It does not militate against the idea that the system should also include detectors for segments intermediate between the grapheme and the syllable (the rime is easy to isolate on the basis of orthographic information alone). The second argument was that although the alternative (syllabic) unit formed a relatively efficient translation unit, it was particularly uneconomical, being associated with an extraordinarily high cost in terms of the number of associations that had to be memorised compared to the grapheme. However, economy in the number of associations cannot be the principal criterion in a symbolic model which assumes, in the lexical route, one unit per word, or nearly 16,000 units (in *DRC-L*) to represent words of one to eight letters, either under their orthographical form or their phonological form, and about 70.000 word units (35.000 x 2?) if it was to implement the lexicon of the average reader.

The belief that phonology is coded as a sequence of phonemes

In sum, given what is known about how we learn to read as well as what is known about the distribution of print-to-sound regularities in English monosyllabic and disyllabic words, there is no reason to believe that the whole range of print-to-sound regularities is captured by a system which is strictly limited to grapheme-size units. We have shown that the idea that a system limited to grapheme-size units is fit to translate correctly any significant proportion of words must be dismissed. For monosyllabic words, this would be possible only if contextual rules are introduced. With polysyllabic words, it is impossible sytems such as DRC. Still, one argument in favour of limiting segments to the size of graphemes still needs to be discussed. It is the argument that "the spelling-to-sound rules which characterize the nonlexical reading route operate at only one level of phonology, the phoneme" (Coltheart et al., p. xxx).

Obviously, if we are to represent speech with clearly dissociated slots corresponding to single phonemes, it is necessary to use units in spelling that can be mapped onto the speech units in these slots. It should be noted, however, that a system in which the spoken counterparts of graphemes were limited to single phonemes would be unable to process all words. As mentionned in the Method section, the single-letter grapheme "x" is translated by the two phonemes sequence /ks/. When disyllabic words are introduced, even more polyphonic units have to be included, further undermining the theoretical coherence of the grapheme-phoneme hypothesis.

In addition, as we will argue below, the developmental evidence gives no principled reason to suppose that the spoken word is exclusively encoded as a sequence of phonemes [check Gaskins et al., 1996/1997; Foorman et al., 1998; Bowey & Hansen, 1994; Burgess & Lonigan, 1998; Johnston, Anderson, & Holligan, 1996; Stahl & Murray, 1994; Wagner et al., 1994, 1997]. Certainly, in alphabetic languages, letters are named individually and this appears to play an influential role in how phonemic awareness develops in the course of reading instruction (Perfetti, Beck, Bell, & Hughes, 1987), that is, the development of the awareness that the continuous flow of speech can be segmented into units of phonemic size which match units in the written word (e.g., words that start with the /b/ sound, like bat, ball, and beach, begin with the letter "b"). Children's ability to detect and manipulate phonemes (e.g., to classify words based on their initial sounds, like with *bat*, *ball*, *beach*, *bell*, and *bill* which all start with the /b/ sound; Bradley, & Bryant, 1985--1983?) is reinforced by abundant training. Nevertheless, some studies (Goswami & Bryant, 1990; Treiman *et al*., 1995b) indicate that -- even when there is phonological awareness -- children find easier to segment syllables into their initial consonant grouping plus their rhyme (the vowel and the consonants which follow it), rather than into any other possible segmentation (e.g., the one suggested by the rules of grapheme-phoneme correspondence). This is because in spoken words rimes are easier to perceive than phonemes, for a phoneme is pronounced as we are still articulating the ones that came before it. Later on, children's phonemic awareness skills benefit from teaching that highlights some of the regular grapheme-phoneme correspondences. However, again, the teaching of print-to-sound correspondences is rarely limited to grapheme-phoneme correspondences, especially in English. Children are often taught the individual rime spellings which are indeed regular in their pronunciations (they are encouraged to memorize that *ook* at the end of a syllable is pronounced /ʊk/) rather than being encouraged to memorize separately each word that ends in *ook*. With consonants, they are often trained on blends instead of isolated graphemes (*e.g*., *bl e nd* rather than *b l e n d*). Even the phonic method, which puts spelling-sound correspondences at the center of its teaching strategy, is based not on graphemes but on *phonograms*, that includes any word ending of high frequency and fairly consistent pronunciation (as for instance *ill* and *ack*). [ref?]

No argument for the presence of units coding graphemes

There is no doubt that grapheme-phoneme associations need to be introduced to capture part of the print-to-sound relations in English. Nevertheless, the fact that a print-to-sound system does not attain a satisfactory level of generalisation without a knowledge of some of the most stable grapheme-phoneme associations is not a particularly appealing argument for the hypothesis that the system is exclusively made of grapheme-sized units.

We mentioned the parsimony argument introduced, in Coltheart (1978) to reject the introduction of units larger than the grapheme. Could this argument also be used to evaluate whether a level of representation associated with the grapheme is required at all?

As a matter of fact, the fact that the notion of grapheme is useful in linguistic description does not necessarily imply that it has any psychological reality. The concept of grapheme is certainly convenient for describing spelling-sound correspondences that are not exclusively of the letter-sound type; we found it useful in our own analyses when describing spelling-sound relations that imply invariant relations between certain groups of letters and their pronunciation. However, in our analyses, we found that a much more efficient description was obtained when we limited the number of graphemes we used. Notably, we found that we were better able to capture the grapheme-phoneme regularities in polysyllabic words, minimize irregularities in segmentation in polysyllabic words, and also account for phonetic phenomena, when we spliced many of the letter clusters that were listed as graphemes in the DRC table, treating their second letter as a silent grapheme which marked *an alternative association for the previous grapheme*.

Certainly, we did not reject the idea of the grapheme altogether. We decided not to split graphemes for which the second letter could not be seen as having a null value in all circumstances. For instance, despite the fact that *h* in *ch* can be seen as silent in *chrome* or *chaos*, we did not isolate *h* as a mute marker because in words like *chief* the addition of *h* to *c* determines a pronunciation that occurs in no other context (*c* alone never has the pronunciation /tch/). But the main motivation for this was the fact that a system with graphemes is more legible; the regularities could just as easily have been captured by marking individual letters for the immediate (and eventually extended) context in which they occur. In other words, we decided to invoke graphemes only for practical reasons. This decision could be challenged if the second letter of a multiletter grapheme was seen as an auxiliary value marker: rather than contributing directly to the representation of a sound, it allows another letter to have a phonetic value which it could not otherwise have had in this position. Then, the letter *h* would be treated as a separate orthographic unit which marks an alternative sound for *t*, *c*, *s*, and *p* in *think*, *chip*, *shy*, and *phase*. This is not as awkward as it may appear at first. Yaguello indicated that moset of the xxx sounds come from (h in latin)….. This would then eliminate all multiletter consonant graphemes from our tables. A similar reanalysis could be performed on multiletter vowel graphemes by following phonics teaching guidelines, which typically treat most of the vowel digraphs as separate units: for example, "When *e* directly follows another vowel letter, that letter has its long value: after *a* as in *maelstrom*, after *e* as in the digraph *ee* of *wheel*, after *i* as in *tie* or *fiery*, after *o* as in *toe*, after *u* as in *Tuesday*. [However] when there are two vowels together, 'the first one is long and the second one is silent' is true only 45% of the time; Clymer, 1963/1996].

Interestingly, when we derived a system that codes the associations between bigrams (single letters with information about the letter on the right) and single phonemes, this system was found to be no less efficient than the grapheme-phoneme system [More efficient than DRC-like?]. -- This raises the possibility that an associative system coding single letters in context has generalization capacities at least as good as the system of grapheme-phoneme rules proposed by DRC. Such a system would then be very close (if not equivalent) to the conversion system introduced by Zorzi and colleagues (1998). In that model, letters activate phonemes in parallel across the letters, and each letter modifies the level of activation of the phonemes in the corresponding positions as well as in the adjacent positions. The level of activation of the phonemes augments progressively and when that level reaches a critical threshold, a response is produced.

Once it has been demonstrated that it the addition of graphemes (i.e., , intermediate units smaller than the body and larger than the letter) was not critical to the efficiency of a conversion system, it is important to provide unambiguous empirical evidence for the representation of graphemes in the reading system. The traditional view is that the (low-level) regularity effect that are observed on naming performances reflects the association between graphemic and phonemic units explicitly represented. However, it cannot be taken for granted that these regularity effects cannot be explained by a knowledge of the pronunciation of letters in context rather than the one of grapheme-phoneme associations. Some data is needed that more directly test the hypothesis of the representation of graphemes.

To the best of our knowledge, only two studies present such data. These are Pring (1981) and Rey et al. (2000). In a lexical decision task, Pring found that altering the format of the letter inside a grapheme eliminated the pseudohomophony effect: the ‘no’ decisions to pseudowords like *grait* (homophone of great) were slower and less accurate than the ‘no’ decisions to pseudowords like *brait* (which is not a homophone of any English word) when the letter case was altered outside the multiletter graphemes (*e.g.*, *GRaiT*). No such difference was found when the disruption occured inside a multiletter grapheme (*e.g.*, *GraIT*). Pring concluded that graphemic units participate in print-to-sound conversion. In a letter detection task, Rey and colleagues showed that participants were slower to detect letters which were part of a multiletter grapheme (e.g., A in BEACH) than when they corresponded to a single-letter grapheme (i.e., A in PLACE), both in English and French [but conclude that activation of E, A and EA]. Again, however, the fact that these results fit the predictions of a system which assumes a representation of the graphemes of the language does not guarantee that the effect emerges because readers actually represent graphemes. In Pring's study, the disruption of a graphemic unit is confounded with the disruption of the onset, nucleus, and coda units *GR ai T vs GR aI T* (grapheme disruption and ONC (onset-nucleus-coda) disruption was confounded for every item in the list, as confirmed by personal communication). It is not guaranteed that grapheme rather than these other linguistic units was the source of the observed reduction of the pseudohomohony effect. In Rey et al.'s study, the difference observed in the letter detection task is explained by the hypothesis that graphemes are treated as wholes, which makes them more difficult to process at the letter level. (This is analogous to the results of Drewnowski & Healy, 1977 of detection errors on THE and AND because words like *the* are processed as a whole, it is harder to process them at the letter level). However, once more, it is not clear, however, that the grapheme level is actually required by these data. Since Rastle and Coltheart's (1998) finding of longer naming times for nonwords like *fooph* (which contain multiletter graphemes) than for nonwords like *frusp* (that do not), it is recognized that with multiletter graphemes activate there is activation of the pronunciation of the component letters as well as the pronunciation of the grapheme. A whammy effect emerges because clusters like **oo** or **ph** cause the activation of different phonemes. In this view, it might be that the longer detection times for *a* in **ea** are caused by the activation of three phonemes, the regular pronunciation of *e*, the regular pronunciation of *a*, and the regular pronunciation of **ea**, partially activated by both the letter *e* and the letter *a*. Notably, no difference in letter detection performance is found between detection of *o* in *cloud* and in *prove*. It was as difficult to identify the letter when an inconsistent pronunciation was activated by a multiletter grapheme (**ou**, for which the pronunciation of the grapheme cannot be predicted from the pronunciation of the letters) as when an inconsistent pronunciation was activated by a single letter grapheme (**o**, for which the pronunciation is irregular),. [Diff in target absent trials?].

[do not really know about neuropsychological evidence -> send to Caramazza]

In short, there is no clear empirical evidence to tell us what orthographic units are represented in systems which posit the non-lexical translation of print into sound. There is very little support for the idea that only grapheme-phoneme correspondences are encoded in print-to-sound conversion systems. There is a lack of conclusive evidence in support of the hypothesis of a grapheme-phoneme rule system. This is certainly an area that needs to be investigated further.

2. On the processes

In the previous section, we highlighted the lack of theoretical and empirical evidence for the general hypothesis that the print-to-sound conversion system contains exclusively grapheme-phoneme correspondences. In this section, we will argue that our analysis of the grapheme-phoneme regularities in disyllabic words as well as the empirical studies inspired by these analyses cast serious doubt on the ability of the rule system assumed by DRC to explain how readers encode and use the print-to-sound regularities.

Firstly, we claim that the hypothesis of a rule system is not realistic. (1) Studies that exploit this analysis of grapheme-phoneme associations in English indicate that naming speed differs for strings with correspondences of high or low predictability, which challenges the hypothesis that no information about the probability of the association in the language is represented. (2) Empirical evidence reveals an interference from the less common pronunciations of the graphemes in a letter detection task, which undermines the hypothesis that only the most common pronunciation of a grapheme is stored. [Even if the localist assumption could be retained, the one-to-one connection would then be questioned].

Secondly, we argue that the significant difficulties encountered in attempting to isolate a set of non-overlapping grapheme-phoneme rules cast serious doubts on the possibility that such a system could be set up in the course of development.

Thirdly, we claim that the processes assumed in DRC of a letter-by-letter grapheme-phoneme conversion is very likely to be inadequate when disyllabic words are introduced. (1) With the large grapheme units that the DRC system contains, many (disyllabic) words cannot be correctly segmented into graphemes without prior knowledge of the pronunciation. (2) Several empirical studies based on our grapheme-phoneme analyses contest the idea that only the first letter and then the grapheme are translated in a letter-by-letter conversion; they rather support the idea that there is activation of the pronunciation of the component letters of a grapheme in parallel to activation of the pronunciation of the grapheme.

Finally, we highlight the fact that in order to account for phonological phenomena, it is necessary to assume a system in which the output is coded with features rather than phonemes. [?? goes to representations?]

Not a rule system

Association rather than rule system

In the mid eighties, the finding by Kay and Marcel (1981) that the pronunciation of the body of a nonword can be biased towards the pronunciation of a previously presented regular or irregular word (*e.g.*, *-ove* se traduit-il par

/ʔʊv/, /ʔv/ ou /ʔv/) engendered a general abandonment of the hypothesis of a system of all-or-none correspondence rules. Nevertheless, in their computational model, Coltheart et al. (2001) decided to reconsider the idea of a rule system they had adopted in the late seventies. They opted to represent print-to-sound relations by a directory of rules in which a grapheme is only connected to its most frequent phonemic counterpart. In this system, there is no knowledge of alternative pronunciations and thus no influence of the inconsistency of the pronunciation of the grapheme.

The rationale for such a rule system was given in a paper by Rastle and Coltheart's (1998) in which they presented related experimental work. We already referred to their statement that : "*Although simple to implement, we have not explored the notion of rule strength in the DRC model because we are not aware of any work which demonstrates that any kind of rule-strength variable has effects on naming latencies when other variables known to affect such latencies such as neighborhood size (e.g., Andrews, 1992) and string length (e.g., Weekes, 1997) are controlled*." Their claim is not that no information about strength is represented in the conversion system. It is rather that they prefer not to introduce in their system an additional hypothesis that is not clearly imposed by empirical evidence.

In view of our description of the grapheme-phoneme regularities, this option is far from reflecting the reality of the language: print-to-sound regularities vary in terms of continuum, that is, of degree of consistency of association, rather than in term of dichotomy (Frauenfelder *et al*., 1993 and Treiman *et al*., 1995, for a similar argument). This is for sure not an argument against the concept of rules. There is nothing to prevent gradual variations from being captured by rules. Still, the notion of an all-or-none rule system sounds counter-intuitive in this context.

First, it is difficult to conceive how such a rule system could be set up during learning without any encoding of the relative strength of the associations. In order to know that /s/ is the most common pronunciation for the grapheme "s" it seems mandatory to keep track of the different pronunciations of this grapheme and their frequencies. This is in fact the solution which Coltheart adopted in the initial DRC model (Coltheart, 1978). In that paper, a rule-extraction algorithm was introduced to prove that it was possible to isolate a system of grapheme-phoneme correspondences rules for phonological transcoding on the basis of mere exposure to a fairly limited corpus of words. The algorithm was abandoned in later versions since it was not meant to be a satisfactory explanation of what happened in the course of reading development. In these later versions, the set of rules was selected manually (Coltheart *et al*., 2001) with no explanation of the way in which regular correspondences are identified during development.

Second, different studies have provided data that support the idea that representations are influenced by their frequency of exposure. For instance, a study by Treiman *et al*. (1995) revealed that subjects varied in their sensitivity to lead (i.e., initial consonants and vowel cluster) and rime units (i.e., vowel cluster and final consonants) in proportion to the structural properties of the language and of differences in the predictability of the pronunciation of these units in the language. In this study, naming times tended to be faster for vowels in the rime unit than in the lead unit since, in English, the ambiguity of pronunciation of a grapheme is often reduced by considering of the consonant which follow it.

As mentioned earlier, such findings were dismissed because of the confound between the manipulation of print-to-sound consistency and the presence of complex multiletter graphemes in the string in one of the consistency studies. Though many works have been published which demonstrate that sublexical (body or coda) rule-strength affects naming, Coltheart and colleagues treat them as dubious because other variables which are known to affect such latencies (such as neighborhood size and string length) were not properly controlled.

The absence of solid evidence for such representation is not, however, solid evidence for its absence of representation. A more valid approach would be to exploit statistical analyses of grapheme-phoneme associations like those presented in this paper and evaluate whether readers' performance is sensitive to variations in grapheme-phoneme correspondence consistency. This is exactly what Lange & Content (1999) did. They proceeded to a factorial manipulation of an entropy variable which gives an estimate of the uncertainty of the pronunciation of a grapheme [Note: Because French is almost entirely regular when contextual regularities are considered, entropy values derived from a system of strictly acontextual grapheme-phoneme correspondences were used]. What they found was that pseudowords containing a grapheme with an uncertain pronunciation (such as **g** in *gatte*) were named less rapidly than pseudowords comprising a grapheme with a regular and systematic pronunciation (such as **v** in *vatte*); the pseudowords in the two lists of pseudowords were chosen so to be equivalent with respect to a long list of orthographic variables.

This effect of grapheme's pronunciation predictability cannot be explained without at least a hypothesis of simple associations, where the speed or the effectiveness of the transcoding process reflects the predictability of the associations in the language. In other words, the notion of all-or-none rules, which implies that the pronunciation of a grapheme is retrieved by consulting a list of correspondences, has to be replaced by one of associations, which supposes that a grapheme and a phoneme are united by a connection whose strength is a function of the predictability of the association.

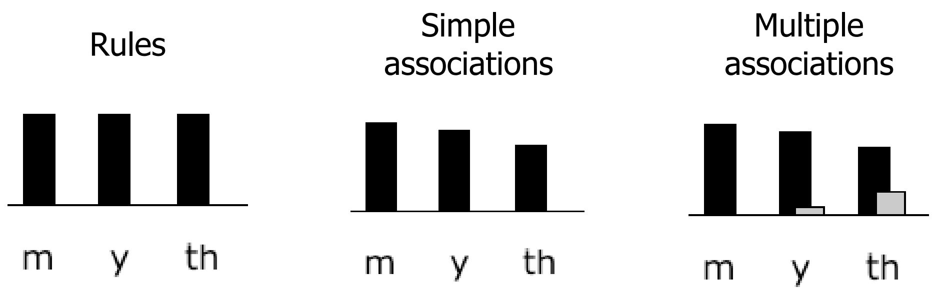
Multiple associations, with dominant as well as irregular pronunciations represented

Interestingly, in Lange & Content's (1999) experiment, 30% of the naming errors were correspondence assignment errors in which a less common pronunciation was given in place of the most common one. This suggests that the less frequent pronunciation of a grapheme was sometimes also activated during conversion and may indicate that, associations are multiple rather than simple, as opposed to the DRC hypothesis that only the most frequent pronunciation of a grapheme is represented in the conversion system.

After all, in disyllabic words, it is often the case that the pronunciation of single letter graphemes is strongly dependent on lexical information (brief discussions of *morphological influences* on English orthographic translation have been offered by Venezky, 1970; Klima, 1972; Henderson, 1982; and Smith, Meredith, Pattison, & Sterling, 1983). Marckwardt (1957, p. 34) mentions that in a number of words ending in fricative sounds, the fricative is voiced or voiceless depending of the syntactic class (noun, adjective, or verb). For instance, "*abuse*" is pronounced [əbius] as a noun and [əbiuz] as a verb, "*close* is pronounced [klos] as an adjective and [kloz] as a verb. So, it is sometimes necessary for a grapheme to map onto different pronunciations. In a reading model like DRC, which assumes that conversion is done without any information about the lexical properties of the string, fewer errors would be made if, when reading aloud, both phonemes were activated but one was selected when information about the role of the word was known. If this can be done for some graphemes such as **s**, whose pronunciation changes with the syntactic category of the string, it might as well happen for all graphemes which have multiple pronunciations. With consonants such as *c* and *g*, for instance, the activation of multiple pronunciations would minimise the delay due to the complete change in the activation pattern when the information about the next letter is added. Share (1995) suggested something exactly along these lines. He proposed that the encoding of all phonemic associates (*e.g.*, *ead* in *bead* and *head*) get activated in the course of learning would be advantageous since it allows the child to make the connection between a word with irregular correspondences (e.g., *head*) and the same word already in his or her spoken vocabulary.

In the light of these findings, Lange (2002) undertook to find empirical evidence for the multiple activation hypothesis. A letter detection task in French was used to reveal traces of the activation of the less frequent phonemic associates of a grapheme (*e.g.*, both **g** to /g/ and **g** to /ʒ/ in a French nonword like *bongour*). Based on Peereman's (1991) finding of a *bonjour* pronunciation of *bongour* (a nonword that would regularly be pronounced [bõgur] in French), this study introduced a distinction between pseudowords that are homophone-by-rule and pseudowords that are homophone-by-association. Homophone-on-rules pseudowords are homophone of a French word when translated by the dominant grapheme-phoneme correspondences of French (e.g., *geudi*) and homophones-on-associations pseudowords are homophones of a French word when translated by any possible phonemic associate of a grapheme (e.g., *bongour*). When participants had to decide whether the letter *j* (substituted by *g* in the pseudoword) was in the string, a homophony disadvantage was observed and this homophony disadvantage occurred for the items *homophone-on-rules* as well as for the ones *homophone-on-associations*. Participants needed more time to decide that *j* is absent in *bongour* compared to a closely matched control pseudoword, *bondour*. Because by construction *bongour* and *bondour* are equivalent in terms of orthographic similarity to *bonjour*, the difference observed in their decision times can only be explained by the partial activation of /ʒ/ from *g* during the transcoding process, reinforcing the uncertainty about the absence of *j* in the string.

Together, these results contest the proposal of Coltheart et al. (2001) that nothing more than a directory of rules is needed to explain human data. They indicate that the knowledge of the print-to-sound relations used by readers for the translation of a letter strings into its phonology must be conceived at least as a network of multiple associations with a representation of every phonemic associate of the grapheme, in which minor associations of a graphemes are activated alongside with the most frequent ones, with a strength that is a function of the predictability of the association in the language. Tat is, they indicate a system closer to the mid-eighties versions of the dual-route model (for instance, the one published by Reggia, Marsland, & Berndt, 1988) than to DRC.



Impossible to acquire the units listed in DRC during development

Coltheart *et al*. (2000) established that they "think of this set of rules as a set of hypotheses about what rules people know." However, there is currently no explanation of how people get to know these rules. It is sometimes mentioned that a central caveat of the DRC model is that it does not provide any explanation of the way the rule system is acquired during development. In view of our analysis, this caveat is even bigger than ever heretofore envisioned. Indeed, the issue of learning in our view poses problem to the very idea that a directory of grapheme-phoneme rules of the kind assumed by Coltheart and colleagues could ever be isolated in the course of learning to read.

Undoubtedly, the development of phonemic awareness may encourage the identification of those letters or groups of letters in the string which have a sound of their own. However, the proposition that phonemic awareness is what permits the development of awareness of the graphemes listed in the *DRC* rule system and then of a set of grapheme-phoneme rules whose main characteristic is that they are exhaustive and non-overlapping is not straightforward. Well installed phonological awareness does not necessarily automatically make the grapheme, the written counterpart of the phoneme, a salient unit for the child to pick up merely by being exposed to the written language.

In the analyses presented here, it proved a difficult task to isolate a set of grapheme-phoneme correspondence rules, especially when disyllabic words were included. There were some items for which the isolation of the grapheme-phoneme associations was quite evident, as for the word *cheep* (“a faint, shrill sound like that of a young bird; a chirp”) which can be mapped onto its phonological transcription [tʃiːp] with the grapheme-phoneme associations, **ch** to /tʃ/, **ee** to /iː /, **p** to /p/. However, for a large number of items isolating the grapheme-phoneme associations was a complex problem. Mostly because of the importance of context, it was not an easy task to segment every letter string of the language in order to isolate a set of non-overlapping units (note in passing that context-sensitive rules in some way impose a breach of the non-overlapping hypothesis attached to the idea of a rule system). Althgough we could identify grapheme-phoneme associations of English using a computer algorithm which extracted a set of correspondence rules from monosyllabic words, the algorithm proved completely unsuccessful with disyllabic words. [example of crazy parsing]. In addition, the identification of such a set of associations by manual selection proved far from trivial given the multiplicity of possible solutions. This multiplicity of possible solutions was even further augmented when we considered silent letters output rules for transforming phonetic codes. In view of the difficulties encountered in our analysis, we doubt that it could e possible that a complete set of rules of the kind assumed in the *DRC* model could be extracted during development. It is therefore unfortunate that the mechanism which would allow the abstraction of these rules, partly on the basis of exposure to the language, and partly on the basis of explicit training of the rules, is presently left unspecified in the theoretical framework proposed by Coltheart and colleagues.

It is possible to conceive that many grapheme-phoneme correspondences are taught explicitly during the teaching of reading. However, we should not put too much emphasis on the explicit teaching of grapheme-phoneme associations, as the teaching of reading often focuses on the most regular correspondences; the full set of correspondences cannot be acquired this way. Furthermore, there is little evidence that the learning of reading orients the novice reader to the units which are currently listed in the current *DRC* rule system. We have already noted that the *DRC* rule system does not allow for silent letters, but in books which teach the “basic skills of phonic reading and spelling”, there is often a section on silent letters [ref]. In one of them [ref], for instance, the child is asked to insert the silent letter “k, g, w, b, or l” in “…nome”, “…nife”, “thum…”, “sa…mon”, “s…ord”. If explicit teaching constrains the way grapheme-phoneme regularities are encoded then we should expect some of these letters to be treated as graphemes on their own, in some contexts associated with a silence. If Coltheart and colleagues want to maintain that it is only units associated to plain sounds that are encoded in the system, then they need to make clear the process by which a child learns that in the word “*salmon*” she does not need to code “*l”* as a grapheme, because it is not sounded in that particular word, in spite of the fact she was told in school that “*l”* <u>is</u> a grapheme, generally sounded, and sometimes silent. Given these discrepancies with what is known about what kids are taught, it would be desirable to provide some empirical evidence that all the clusters listed as graphemes in the *DRC* rule system are in fact treated as graphemes by readers.

Since this is certainly a difficult task, arguments for general options need to be provided. To come back to the issue of silent letters, they are absent from the *DRC* system because the authors decided to consider as graphemes only units that are the written counterparts of a plain sound (a single phoneme or elementary speech sound, silent letters simply do not correspond to a specific unit in speech). It is far from clear that this option is motivated in any way other than the decision to define a grapheme as such.

On the one hand, we do not know of any empirical evidence that we know about that can tell us whether a silence is sometimes listed as one of the alternative pronunciation of the graphemes. On the other hand, in linguistic descriptions, it is customary to take the position that some letters are best viewed as simply unsounded. Even though Venezky (1970) defined a grapheme as "*[a] functional unit […]a string of one or more letters that acts as a unit in predicting a sound*" (for example, for **ch** read /ʃ/ as in *chef*), he made clear (as described in our Method section) that, for many graphemes, a far more sensible linguistic description is obtained if we invoke a mute phoneme and splice sequences of letters into a sounded grapheme and a mute grapheme. [Note: Considering *debt* and *indict*; it seems more appropriate to propose that some letters are silent before final *t* than to propose that *ct* or *bt* are units functional for pronunciation]. In support of this, we showed, in the Result sections, that a system which allows for graphemes to be associated with a silence captures more regularities in the table than one in which any unsounded letter is necessarily systematically attached to an adjacent letter or letter sequence. This result is obtained even though the former system produces a smaller number of units in the table.

This is not limited to the issue of silent letters. Our analyses decided to take a different stand from DRC on what units were to be listed in the grapheme table. In particular, we imposed grapheme disjunctions for a large number of clusters (e.g., geminates, rhotic vowels and syllabic consonants) and added codes to mark contextual constraints in a systematic way (*e.g.*, **ch|a|n|c**$_e$**|e**). In contrast, DRC made the decision to list a large number of graphemes and to use identity rather than the role of the letter as a mark for contextual influence. In the Method section, we argued that our decision lead to better linguistic coherence and in the Results section, we showed that greater generalisation was achieved with our segmentation options.

Given these findings, *DRC* needs to be supported by better arguments. In general, if the hypothesis of a system of symbolic rules and representations is to be insisted on, an explanation needs to be provided for how such a rule system could form. Is it or not related to things that happen in the course of the development? If so, how? So far, these questions cannot be satisfactorily answered.

Numerous segmentation (hence pronunciation) errors in a DRC-like system

A further issue for the DRC model in its current formulation is that its assumption about what constitutes a grapheme causes specific difficulties for identifying the segments that match the grapheme units stored in the system, quite apart from the problem of the generation of phonology. This is because certain sequences of letters happen to be pronounced with a different number of phonemes, sometimes causing a same letter cluster to be segmented in different ways in different strings *(ui* in *suite* /swiːt/ and in *suiting* /suting/).

These difficulties were recognized by Coltheart (1978) when he stated that “*One of the ways in which English spelling is irregular is at the level of the relationship between letters and functional spelling units; this irregularity is sufficiently widespread to suggest that no parsing procedure exists which correctly analyses every English word into its functional spelling units without using lexical knowledge*" (p. 155). However, since then, segmentation ambiguities have scarcely been evoked in theoretical discussion (one exception is Shallice & McCarthy, 1985). They were left unmentioned in the discussion of the *DRC* model of naming, most likely because its simulation was limited to the domain of English monosyllabic words, where the problem does not occur. The analysis of our corpus of monosyllabic words only revealed one cluster whose segmentation was truly ambiguous, that is **ie** (e.g., *ie* in ***nie**ce* /i̞/ and *v<u>ie</u>w* /ju/). With polysyllabic words, however, it is not rare for an orthographic segment to be associated with different segmentations according to the pronunciation of the letter string (e.g., *ph* will correspond to one functional unit of pronunciation in *mor**ph**eme* but to two in *she<u>ph</u>erd*). In vowels, for instance, a following *e* or *i*, often indicates an inconsistent digraph value, with the letters sometimes pronounced separately (e.g., *diet*, *poet*, *duet*) and sometimes not (e.g., *friend*, *shoe*, *sue*). In fact, almost every digraph and trigraph listed as a grapheme unit in the *DRC* table is subject to inconsistent segmentations (*e.g*., ***ai*** in *n<u>aï</u>ve* or ***rai**se*; ***ao*** in *ch<u>ao</u>s* or ***gaol***; ***ea*** in *r<u>ea</u>ct* or ***gear***; ***dd*** in *go<u>dd</u>amned* or *mi<u>dd</u>ay*, ***gn*** in *si<u>gn</u>et* or ***sign***; ***mb*** in *bo<u>mb</u>ard* or ***bomb***).

Sometimes the position or the local context can help to resolve the issue. For instance, in *fanc**<u>ie</u>**s* or *p**<u>ie</u>**s* (in contrast to **ie** in *bel**ie**ve*), the *e* can be identified as a final mute marker of long vowel (here the "i"). However, in most circumstances, the uncertainty about the appropriate way to segment a grapheme string can be reduced only by taking into account syllabic or morpheme boundaries (*f**oa**m* vs *ch<u>ao</u>s, mor**ph**eme* vs *u<u>ph</u>ill*; ***rea**ch* vs *r<u>ea</u>ct*; Haas, 1970; Venezky, 1970). [Note: As indicated by Henderson (1985) Support for the idea that it is in fact the morpheme boundary that determines the splitting comes from the sequence "mb". It is sounded fully when it straddles a syllabic boundary (*ambit*) but the "b" is silent when the cluster closes a word (*bomb*). The "b" is also silent before inflexional affixes (*bombing*), but is pronounced in other circumstances (*bombard*).].

Unfortunately, information about syllabic or morphological boundaries is denied to the process of direct translation by virtue of the fact that the rule system is, in the DRC set-up, hermetic to influences from information represented in the mental lexicon. In other words, most of the time it is impossible to identify a letter cluster as made of one grapheme or many, without any information about the pronunciation. And, obviously, the pronunciation cannot be used as a guide for segmentation when the goal of the processing is precisely to determine the pronunciation of the letter string.

In the absence of information about syllabic or morphological boundaries, any system that invokes graphemes can be expected to make segmentation errors (and as a consequence, pronunciation errors) once polysyllabic and polymorphemic words are considered since in a system of that kind, any sequence of letters that matches a grapheme listed in the table will always be treated as a grapheme. This means that if the string contains a sequence of letters that matches one of the units listed in the grapheme table, then theletter sequence will be identified as that grapheme and pronounced with that grapheme's regular pronunciation in every string that contain the sequence. In many occasions, the segmentation error will cause a pronunciation error. The more units of 2, 3, 4 letters in the grapheme table that are pronounced either with one or with many phonemes (*ea* in *reach* and *react*), the more the system sill be prone to segmentation and thus pronunciation errors.

In the Method section, we mentioned that our decision to allow for silent graphemes in the system was motivated by the desire to minimize the number of units with variant segmentations. In the Results section, we showed that a system that avoids the introduction of silent graphemes produced far fewer segmentation errors than DRC. Only 2% of segmentation errors were found with our segmentation options compare to 71% of words incorrectly segmented when a *DRC*-like correspondence system attempted a letter-by-letter segmentation of the string into graphemes.

Activation of all the phonemes associated with the component letters of a grapheme

The liaison phenomenon suggests that the phonic value is not completely extinguished + Glaswegian accent with guttural R in car, bar (ask Scobbie, QMU).

Another issue is what pronunciations get partially or fully activated during the conversion process. DRC assumes that in reading aloud, a letter string is deciphered letter-by-letter. At first, the leftmost letter will be isolated, a match with a grapheme in the rule system sought, and then its phoneme counterpart selected. Each time a new letter is processed, it is first joined to the unit previously translated (generally the previous letter; for instance **h** added to **p** in **ph**). If the combination matches a grapheme listed in the table then it will activate the phoneme associated with the grapheme in the same phonemic position (the phoneme /f/ replaces the /p/ previously active). If the combination does not match a grapheme, a switch in both grapheme and phoneme positions will occur (**r** in **cr**).

In the conversion process, only the first letter(s) of the grapheme and then the full grapheme receive activation. The last letter never receives any. To some extent, this predicts systematic difficulties in the presence of the phenomenon of liaison, which is widespread in French and also occurs in some specific contexts in English. This is the phenomenon by which a final consonant that is usually mute is pronounced when the first letter of the following word is a vowel (*e.g*., in French, *sans-abri* vs *sans-papiers [use IPA to show what you mean]*). This happens with linking-r in non-rhotic varieties of English. Final "r", usually mute, is pronounced when the next word begins with a vowel (*e.g*., XXX). In a DRC-like system, the last letter will be the last letter of a multiletter grapheme because it is often found mute in that sequence. This means that the action of the left-to-right conversion process will cause pronunciation errors every time the final grapheme is followed by a vowel. The only way to prevent to prevent errors would be for the conversion system to backtrack every time the linking consonant is followed by a vowel. In early work, for instance, Meyer *et al*. (1974) proposed that if the output of lexical and non-lexical processes were found to be different, the letter string would be reparsed, and new, less frequent grapheme-phoneme correspondences tried out. However, a backtracking mechanism of this kind would almost certainly result in long delays in performance. We know of no empirical evidence to indicate whether there is or is not a time delay in performance with linking consonants. [something of the like that does not seem right].

According to Coltheart and colleagues, the hypothesis of a letter-by-letter mechanism is strongly supported by the finding of a grapheme complexity effect (called whammy effect; Rastle & Coltheart, 1998): naming times are slower for a nonword like *fooph*, which contains the multiletter graphemes **oo** and **ph**, compard to a nonword like *frosp*, which contains no multiletter grapheme. This result is understood to indicate the successive conversion of the first letter and then the grapheme which takes place during the operation of the left-to-right conversion process assumed in DRC. With the grapheme **ph**, the letter **p** and then the cluster **ph** will activate an entry in the graphemes' system. The processing of the first letter, "p", causes the activation of the phoneme most frequently associated with it (/p/ for **p**). When the second letter is added to the one previously processed, the processing of the cluster "ph" causes the successful activation of an entry in the graphemes' system and then the activation its pronunciation (/f/ for **ph**). Due to inhibitory connections between units in the same phonemic position, the level of activation of the phoneme /f/ will tend to increase and the one of the phoneme /p/ decrease, with the rise of the correct pronunciation being slowed down by the temporary activation of an incorrect phoneme.

This position was however undermined by a series of studies (Lange & Content, 2000) manipulating grapheme complexity (the presence of either a two-letter grapheme or a two-letter cluster: *ph* vs *pr*) and convergence of pronunciation (the component letters have or do not have the same pronunciation as the grapheme: *ph* vs *ff*), as well as length in graphemes and length in letters. The motivation for these studies was the observation that grapheme complexity was associated with differences in mean grapheme frequency between the items in the two conditions: the frequency value of multiletter graphemes such as *ph* is lower than the average of the frequency values of the graphemes *p* or *r* in *pr*. To assess whether variations in grapheme frequency played a role in the reported effect, Lange & Content (2000) conducted an experiment in which they showed that grapheme frequency had a significant influence on naming times. French nonwords such as *nuze*, with an rare grapheme, are named more slowly than nonwords such as *nuse*, which contains a frequent grapheme. Applying a covariance analysis on Rastle and Coltheart's (1998) data, they revealed that the reported effect was reversed when mean grapheme frequency was entered as a covariant factor. In an experimental study, they showed a similar facilitatory effect for multiletter graphemes when mean grapheme frequency values were stable across conditions.

<u>Because of the omnipresence of phonetic phenomena in polysyllabic words, the output is unlikely to be a set of phonemes</u>

A final issue is the coding of phonetic phenomenon. In our analyses, we introduced markers of influence of surrounding sounds on the pronunciation of a grapheme, so as to take into consideration the most common phonetic phenomena that are known to occur in English. For instance, we marked the /ʒ/ pronunciation of **s** in *Asian* with a "*pal"* label that suggested that this pronunciation is the result of the assimilation of the consonant to the yod (*i.e.*, palatalisation). A reason for this was that even though these changes in pronunciation could be captured with an orthographic context (the letter **s** followed by **i, a vowel, and a "n"** is to be pronounced /ʒ/), regularities are better captured when the phonological rather than orthographic context is considered (i.e., the sound /ɪ/ becomes /j/ when in front of a vowel in an unstressed syllable and the sound /z/ becomes /ʒ/ when a /j/ is assimilated to it).

On the surface, these marks parallel Coltheart *et al.*'s (2001) hypothesis of explicit "*output phonotactic rules*". In this model, the representation in the phonological buffer obtained after grapheme-phoneme conversion is modified in order to account for a set of phonological constraints listed in another set of rules, For instance, the voice assimilation of **tz** in *quartz* [kwɔːrts] or *kibbutz* [kɪbʊts] is accounted for by a phonotactic rule that transforms the /z/ obtained from the assignment of the regular correspondence **z** to /z/ into /s/ after an unvoiced consonant (and also the **s** of the plurals form, pronounced /s/ instead of /z/ in *antics*, *effects*, *ranks*).

This description implies that phonological constraints are applied just before the articulatory plan pre-programmed from phonemic units is shipped off for motor execution. To our knowledge, there is no psycholinguistics study to support this hypothesis and linguistic evidence rather support the view that transformations of the phonetic codes occur between transcoding and production.

This is best suggested by the /s/ or /z/ pronunciation for the plural marker “s” and the /t/ or /d/ pronunciation of the past-tense marker “d”. In both cases, the pronunciation depends on the last sound of the word these affixes are attached to. The plural marker “s” is pronounced /s/ after voiceless phonemes (as in *cats*) but /z/ after voiced phonemes (as in *hills* or *bees*). The past tense marker “ed” is pronounced /t/ in verbs that end in another voiceless phoneme (*stopped* [stopt]), and /d/ in verbs that end in another voiced phoneme (*stubbed* [stʌbd] or *cried* [kraɪd]). Importantly, the pronunciation of an important “e” is affected by the value taken by the surrounding segments (“es” is pronounced /ɪz/ in *horses* and “ed” is pronounced /ɪd/ in verbs that ends in t or d, like *lifted* [lɪftɪd]),

The view that pronunciations are linked to phonetic phenomenon is reinforced by consideration of how pronunciations have changed over time. Originally, words such as *special* and *nation* were trisyllabic, with the primary stress on the last syllable: [spɛsɪ’ɑl, nɑʃɪ’uːn], as they had been pronounced in French, the language from which they were borrowed. In the course of time, as with other French loan words, stress shifted to the first syllable, conforming to the English pattern, and resulting the final vowel being neutralised [’spɛsɪəl, ’nɑʃɪən]. Next, [ɪ] before an unstressed vowel became the glide [j], causing the collapsing of the original three syllables into two, after which the [sj] combination became [ʃ] by palatalisation. [Note: The reason for this is simply that the tongue position for [j] is just behind that for [ʃ], that is to say, the blade and front are loosely palatal. In transferring from [s] to [j], the tongue-tip remains at the alveolar ridge and acts as a pivot for the blade to swing upward. In so doing, it forms the lengthened channel which is the particular configuration necessary for the production of [ʃ]].

It is reinforced even further by the fact that such transformations are also observed in the spoken form, between words in running speech. Kreidler (1990, p. 11), for example mentions that a speaker may pronounce /tj/ as /ʃ/ in the word sequence *Disappoint you* or /sj/ as /ʃ/ in the sequence *miss you*. Marckwardt (1957, p. 33) also indicates that, if pronounced rapidly, it is almost impossible to avoid saying [fɔəʳʃu] to the word sequence *force you* or saying [kæntʃu] for the word sequence *can't you*). [CD: I may be able to provide a reference, Scobbie]

Most co-articulations stop at a syllable or word boundary

However, it does not seem to be the case that these transformations of the phonemic codes occur right after conversion. Instead, such transformations seem to reflect the operation of a later process, when phonemes are blended for articulation. Evidence of this comes from the fact that phonetic phenomena generally do not hold across syllabic or morpheme boundaries. Geminate consonant clusters are pronounced as single consonants in *add, letter,* and *canned* but as two in *greenness* [gri:nnɪs] and *headdress* [hɛddrɛs]; sequences as *nk* and *ng* (i.e., *n* before letters pronounced /g/ or /k/) are pronounced /ŋ/ only when the letters appear in a same morpheme (/ŋ/ in *singer* but not in *danger*). The letter "s" does not have its variant sound when between two vowels that straddle a morpheme boundary (e.g., **s** is pronounced /s/ rather than /z/ in *asocial*). It is as if morpheme and syllable boundaries were phonetically marked (Krakow, 1999) and these marks prevent the transformation from occurring. In other words, the exact pronunciation cannot be fully determined until information from the lexicon has become available.

Other evidence that co-articulation does not occur on-line, by rewriting the output of the conversion process, but must be planned before the execution of articulation comes from Whalen (1990). He found evidence for anticipatory co-articulation of the second vowel of VCV sequences like *abi*. He further established that anticipatory co-articulation was present only when the identity of the second vowel was known prior to articulation of the first vowel had begun. In this experiment devised by Whalen (1990), subjects had to produce VCV sequences in which one letter was initially missing (*a*_i or *ab*_). Either the medial consonant or the final vowel were supplied when voicing for the first vowel was detected . Because anticipatory co-articulation of the second vowel was found in the *a_i* condition but not in the *ab_* condition, he concluded that co-articulation must be planned rather than produced on-line.]

But co-articulation, not phoneme identity transformation

We argue that in fact a system would be more efficient if contextual rules had access to information about commonalities between, features (/bt/ = stop + stop, /mp/ same articulation, stop; /pm/, /tn/, /kŋ/). Syllabification, for example, takes sonority into account. For that, we need to include information about features (i.e., the properties of phonemes).

Another important drawback for the *DRC* theory is that its phonotactic rules are written in a way that has little to do with how transformations actually take place.

We indicated that pronunciation is described in DRC as the result of phoneme to phoneme rewrite rules operating on the output of the transcoding process. However, describing the process as a change in the identity of a phoneme does not reflect the reality of the transformation. The reality of these transformations is that a change of pronunciation occurs because as a phoneme is articulated it borrows the phonetic features of the adjacent phonemes in specific contexts. For example, one very common process is *final devoicing*, in which the value of one feature is changed (i.e., [+voice] becomes [-voice]; e.g., /s/ to /z/ or /t/ to /d/). Moreover, phoneme transformations do not occur in any context. They only occur when the two sounds are highly similar. To represent these similarity relations it is necessary to go beyond a representation at the level of the phonemes and look into the detailed characteristics of each unit in terms of their phonetic features.

Features are assumed to stand for a set of instructions to the physical articulatory system and include voicing, nasality, tongue height, etc. Each language has a set of distinctive features which distinguishes each phoneme of the language from every other phoneme of the language. For example, the segment /b/ has the features [+voiced], [+labial], [-nasal], and so forth because its articulation involves vibration of the vocal cord, lip closure and lowering of the velum). [Note: suprasegmental phonology, in particular, insists on the need to refer to features for adequate analysis of a language].

The details of pronunciation then are the result of "the overlapping of adjacent articulations" and of co-articulation (Ladefoged, 1993). Speech sounds occurring in succession are not pronounced as single units but they are co-articulated. The pronunciation of a phoneme is influenced by the articulation of the neighbouring phonemes in the utterance, with adjustments in the vocal tract shape made by anticipation of subsequent motion [Note: to be more exact, both articulatory and acoustic aspects matter. Muscular adjustments and movements are the source of the change of /n/ to /ŋ/ before /k/ (as in *sink*). But one tends to be perceived as the other because acoustically, /n/ is only slightly different from /ŋ/]. For example, in the word 'Mum' the vowel phoneme is one that is normally pronounced with the soft palate raised to prevent the escape of air through the nose, while the two /m/ phonemes must have the soft palate lowered. The soft palate cannot be raised very quickly, so the vowel is pronounced with the soft palate still lowered, giving the vowel a nasalized quality. The nasalization is a co-articulation effect caused by the nasal consonant environment. Another example is the lip-rounding of a consonant in the environment of rounded vowels: in the phrase 'you too', the /t/ occurs between two rounded vowels, and there is not enough time in normal speech for the lips to move from rounded to unrounded and back again in a few hundredths of a second; consequently the /t/ is pronounced with lip-rounding. Co-articulation can be both anticipatory (e.g., the lips are rounded in the production of /s/ of 'soon' in anticipation of the following rounded /u/ vowel) and preservative (the rounding from /u/ persists to the following /n/). It is largely acknowledged that actual pronunciation is always influenced by co-articulation (Fowler & Rosenblum, 1989; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985; Whalen, 1990).

Changes in pronunciation that lead to a change of phoneme identity are only a special case of co-articulation. Although the output rules in the DRC system accounts for cases of assimilation between two strictly adjacent segments (such as Nasal Place assimilation), they leave out cases of vowel harmony in which vowel interactions are long-distance (such as xxxx).

Drop articulation, but then not a model of reading aloud anymore

It could be claimed that DRC is a model of reading aloud, that it only explains the cognitive processes that take place prior to articulation, leaving models of word production to determine how the output of the conversion system is transformed in the course of articulation. But then it would not make be model of reading aloud anymore, only a model of print-to-sound conversion. Also, it would make any prediction of the model's performance highly speculative, since the output representation would be impossible to match with subject responses; it would need to be matched onto the output representation prior to articulation, which is unavailable. So, if the claim that DRC forms a complete model of reading aloud is to be maintained, it will be necessary to produce a correct description of what happens during articulation. And, unfortunately, representing the sequence to be uttered as a succession of independent sounds or phonological units is a far cry from what speakers actually produce. Speech is not a sequence of phonemes; to successively utter [b], [a], then [g] is not to speak.

What is often overlooked by researchers running experiments on single word recognition is that people often vary the realisation of the graphemes in a word according to sentence context. According to *Greenberg and Fosler-Lusser (xxx, p. xxx)*, the 50 most common words (accounting for over 50% of the lexical tokens in their corpus) are phonetically realized in dozens of different ways. Speaking rate, in particular, has a dramatic impact on pronunciation. In particular, very fast speaking rates create deviations from canonical pronunciations, commonly by causing the reduction of vocalic nuclei and deletion of segments (mostly codas). In such circumstances, not only are segments shortened, but the overlap between adjacent articulations also increases (Browman & Goldstein, 1990; Gay, 1981).

Also, to come back to the issue of mute phonemes, sometimes it is not that there is no phoneme to match a letter onto. It is rather that the associated phoneme has been completely assimilated by other sounds. What seems to be the disappearance of a phoneme in the transcription is sometimes the result of two successive sounds which largely or completely overlap in features (/ʃʃ/ to /ʃ/ in *fuschia* compared to *fish shop;* /bp/, which share all but one of their features, to /p/ in *subpoena*). The transcription may suggest that the phoneme has dropped out altogether. However, detailed examination of speech shows that such effects are more gradual: in slow speech the /t/ in 'acts' may be fully pronounced, with an audible transition from the preceding /k/ and to the following /s/, while in a more rapid style it may be articulated but not given any audible realisation, and in very rapid speech it may be observable, if at all, only as a rather early movement of the tongue blade towards the /s/ position. [CD: Refs: Browman & Godstein, 1990b, another example perfect memory]

It seems impossible to account for these variations in production with speech rate and to explain the actual realisation, without a representation of the features.

Arguments have been presented to defend the idea that features are represented in the cognitive systems (Browman & Goldstein, 1990b, p. 420; Roelofs, 1999). *Roelofs (1999), in particular, discussed problems encountered in explaining some speech error phenomenon*. That features are directly represented is supported by Lukatela et al. (2001)

What is the evidence for phonemes? Are phonemes necessary, before features?

Conversely, it is far from certain that a reader activates an abstract representation at the level of the phonemes when reading aloud. It is certainly indisputable that a relatively fixed set of sound units (phonemes) can be isolated for each language and that these units permit the speech stream to be represented in terms of a sequence of phonemes. With proper training phoneticians are able to transcribe the sounds they hear with these phonemic symbols with a large degree of agreement. It is well documented that children consciously acquire this type of representation during their very first school year when learning an alphabetic script and this has largely contributed to making the sound segment and even the phoneme a psychologically real thing for literate language users. In short, segmentation appears to be possible in most cases, and speakers seem to be aware of segments in their speech.

But none of this necessarily implies that the output of conversion is a sequence of discrete phonemic units of the kind used to represent the spoken form of words in dictionaries or lexical databases. After all, the phonemes in this sequence are only convenient labels for a set of distinctive articulatory gestures, completely redundant with the set they represent. Conversely, features set do not need to be used as an equivalent of the phoneme; it is possible for them to be activated asynchronously. Although speech is conveniently represented by a sequence of phonemes, it does not happen as discrete units. There is no way to cut a spectrogram into non-overlapping segments in order to isolate the parts corresponding to *b*, *a*, *t* in *bat*. It can even prove extremely difficult to identify separate sound units (segments) that correspond to phonemes, since many of the articulatory movements that create sounds are continuous rather than sharply switched. For example, according to Roach (2002), pre-consonantal /n/ sounds in English (e.g. 'kind' /kaɪnd/) are often almost undetectable except in the form of nasalisation of the vowel preceding them; sequences of fricatives often overlap, so that it is difficult or impossible to split the sequence /fs/ in 'fish soup', or /fθs/ in 'fifths'.

But that does not necessarily imply that phonemes need to be represented before features. As a matter of fact, the need for the phoneme to be represented independently has already been disputed by featural theories of speech processing, both in the domain of speech perception (Marslen-Wilson & Warren, 1994; Warren & Marslen-Wilson, 1988) and speech production (e.g., Mowrey & MacKay, 1990; Dell, Juliano, & Govindjee, 1993; Dell & Juliano, 1996). According to these theories, phonological segments are represented by nothing but their features. In the domain of speech comprehension, Marslen-Wilson & Warren (1994) have argued that “there is no intrinsic computational reason why feature nodes should not communicate directly with the lexical level” (*Marslen-Wilson & Warren, 1994, p. 654*). They dismissed the intervention of a layer of phonemic units between feature analysis and lexical representation and proposed that speech is produced as a continuous sequence of articulatory gestures, which results in the sequence being modulated continuously. They had previously presented evidence that listeners restrict the set of possible analyses on the basis of partial information about the incoming segment when it is still not possible to fully identify this segment (Warren & Marslen-Wilson, 1988). For instance cues to voicing can be heard before definite cues to place of articulation with a vowel preceding a voiced labial stop. Features are activated as soon as they are detected. The idea that the phoneme serves as the fundamental unit of linguistic organization in the production of speech has similarly been questioned. Dell et al. (1993), notably, introduced a speech production system in the form of a PDP model in which lexical items are connected to features via a layer of hidden units [the feature layer is the output layer, which contains one unit for each feature in the language].

In sum, all current models of speech perception and production invoke features, DRC does not represent them. Features are found to be a lot more crucial than phonemes to capture regularities in print-to-sound conversion. DRC decided to code phonemes but not features. DRC theorists have in the past justified their refusal to add an hypotheses largely accepted in the psycholinguistics community (i.e., readers' knowledge of the body-rime regularities) by the absence of unambiguous evidence that forces this assumptions. In view of the review we have provided, it is only fair to invite them to come up with some detailed justification of their coding options at the output level.

CONCLUSIONS

The DRC theorists (Coltheart et al., 1993; Coltheart & Rastle, 1994, 1998; Coltheart et al., 2001) had presented the finding a computer model instantiating the *DRC* theory has a better overall fit with human performance, compared to the connectionist models of Zorzi et al. (1998) or Plaut et al. (1996) as supporting for the view that the DRC model has an higher validity than other models. In support of that view, they also presented the fact that, compared to connectionist models, symbolic models better capture our understanding of the reading processes because they provide specific, clearly identifiable and well motivated hypotheses (this is the classic argument that connectionist models propose a simulation, not an explanation).

The truth, however, is that a high degree of fit or clearly specified hypothesis will never guarantee that the model accurately captures the representations and processes that are involved in human performance. It is not contested that a computer model is to have a critical contribution in establishing the realism of the model and in making explicit the predictions of theories that assume various levels of treatment. However, once the theory has passed the ‘runnable’ test, being able to demonstrate that every single hypothesis made by the model is supported by empirical data will always be more critical than a good overall degree of fit with human performance. In the context of a discussion of the origin of the consistency effects, Rastle and Coltheart (1998), made this point that a good degree of fit offered no support to hypotheses that where not supported by unamibuously empirical evidence. In this study, we presented linguistic and empirical evidence that similarly questions the validity of the hypotheses that were . More specifically, evidence was found that questions hypotheses of the DRC model that never had, to our knowledge, their psychological reality unambiguously established by empirical, developmental or linguistic evidence.

Of particular concern was the fact that the current version of the DRC model has a domain of simulation limited to monosyllabic words when monosyllabic and disyllabic words appear to impose very different constraints on the processing mechanisms. Our analysis of the pronouncing dictionaries provided additional insight into the relationship between graphemes and their pronunciations in a corpus of reading that included polysyllabic words. Importantly, the linguistic description suggested a different set of hypotheses about how regularities might be represented in humans. Some concerned minor aspects, to do with the exact nature of the knowledge of grapheme-phoneme relations, for instance, challenging the hypothesis that orthographic units are defined as a function of the spoken unit, the phoneme, with the constraint that grapheme units must always correspond to sounded speech sound. It was proposed instead that it makes more sense to isolate clusters not associated to a speech sound as units functional for pronunciation. This is because the speech sound can be viewed as having been assimilated by a neighbouring one, or because the letter can be seen as a functionless scribal insertion (**s** in island), or else because the learning of reading encourages the recognition of some letters as mute. We showed that a system based on the DRC hypothesis that the translation of a grapheme must always be a sounded phoneme is not particularly efficient once disyllabic words are introduced. First, it necessitates the use of a far larger number of orthographic units with no corresponding gain in the proportion of regularities captured by the system. Second, these numerous multiletter units are often associated with inconsistent segmentations in polysyllabic words, leading to numerous errors in the identification of graphemes since it operates on the basis of a letter-by-letter deciphering mechanism which has no knowledge beyond the letter immediately on the left or on the right. We also proposed that, to be efficient, a system has to include more than the traditional definition of grapheme-phoneme rules, with the addition of both rules in context and multi-letter to multi-phoneme correspondences.

Other parts of the study concerned major aspects of the way a knowledge of the most stable grapheme-phoneme relations participate in print-to-sound conversion. For these, empirical support was sought after. Importantly, empirically evaluation supported our set of hypotheses, rather than the ones proposed by DRC. For instance, Coltheart (1978) introduced a parsimony argument to validate the idea of an all-or-none, one-to-one rule system, in which each grapheme is connected to its most frequent pronunciation. The reality of language, however, is a system of quasi-regular rather than regular relations. Challenging DRC hypothesis of an all-or-none rule system which exclude any information about the strength of the correspondences in the rule system, Lange and Content (1999) revealed an effect of grapheme-phoneme consistency. Strings which contain a grapheme with a dominant but ambiguous pronunciation were read aloud more slowly than strings which contained only graphemes with non-ambiguous pronunciations. In addition, challenging the other hypothesis of a one-to-one system, i.e., only the most frequent pronunciation is represented, Lange and Content (2001) revealed that a less frequent alternative pronunciation of a grapheme influences performance in a letter detection task. Together, these data showed the need for a system of association, in which the activation of the multiple associates of a grapheme is modulated by their predictability in the language, rather than a system of correspondence. Further, challenging the hypothesis that conversion is a letter-by-letter process, it was found that the whammy effect reported by Rastle and Coltheart (1998) and presented as central evidence for a sequential conversion process was invalidated by a confound with grapheme frequency. In the absence this confound, a pattern of data opposite to the one predicted by the model was observed.

These results are not compatible with the predictions of the current DRC model. We should carefully avoid to suggest that these findings indicate that there is no truth in the DRC model and that this model should be dismissed outright. Certainly, a good degree of fit has never been the only argument in favour of the DRC model. First, Coltheart and colleagues demonstrated that the *Dual-Route Cascaded network* (*DRC*) correctly simulates a large number of empirical data for both normal participants (non-word naming in a partial model; Coltheart *et al.*, 1993; naming latencies and errors in a complete model, Coltheart & Rastle, 1994; Coltheart *et al.*, 1999*a*) and patients with specific naming deficits (Coltheart, Langdon, & Haller, 1996). Second, though our results questioned some specific hypotheses of the dual-route model, they were found to be globally coherent with the dual-route framework and the idea that there is in the reading system a route that operates on the basis of a knowledge of the most regular print-to-sound correspondences: at least, there is evidence to reinforce the claim that skilled readers use a knowledge of the stable relations between graphemes and phonemes. In particular, an effect of the consistency of the grapheme-phoneme relations has been found in French (Lange, 1999) as well as in English (Lange, 2005 – in preparation) when the regularity of larger units was constant and the orthographic properties of the items were strictly controlled so as to make any lexical origin of the observed effect implausible. These findings establish the reality of a pure grapheme-phoneme regularity effect (previously questioned by Parkin, 1984), along with the data of Andrews and Scarratt (1998) providing evidence of an influence of low level regularity on the performance of skilled readers.

However, our evidence to highlight the need for the hypotheses underpinning this model to be better motivated. Just because the *DRC* model presents an explanation particularly compatible with specific empirical (length effects) or neuropsychological (dyslexia) data does not guarantee that the model provides an optimal explanation of human performance or that the processes in the models are the same as the ones which cause the effect in the readers. Jacobs & Grainger (1999) made a similar point in their comments of Levelt's BBS paper. They pointed out that "*finding that one model fits data better than competing models does not establish the best-fitting model as the probable source of the data (Collyer, 1985)*." Importantly, showing that skilled readers use some knowledge of the most frequent pronunciations of letters or multiletter clusters is not the same as proving that print-to-sound conversion operates as a letter-by-letter deciphering of the string with look-up in a grapheme-to-phoneme rule system and activation of the most frequent pronunciation of the clusters recognized as graphemes.

Such results at most offer support to the idea behind a separate non-lexical system, that is, the idea that pronunciation can be obtained from a sequence of letters and its orthographic environment. In the absence of any clarification of what limits there are to acceptable modifications of the model, it is extremely difficult to say whether these results prove the DRC theory to be an inadequate theory of reading. For instance, Coltheart and colleagues could argue that evidence for a grapheme-phoneme consistency effect does not necessarily dismiss their theory. The citation by Rastle and Coltheart that we already mentioned stated that: ***« we have not explored the notion of rule strength in the DRC model because we are not aware of any work which demonstrates that any kind of rule-strength variable has effects on naming latencies when [appropriate variables] are controlled***." The association system presented in this study could be presented as an excellent candidate for a revision of DRC' conversion system. We do not encourage this. What our results indicate is that reading performance is affected by a graded measure of grapheme-phoneme association consistency derived from our analyses. It may be that our analyses may captured some part of the regularity the readers code but we haven't unambiguously demonstrated that our statistics exactly capture the way that readers encode their knowledge of print-to-sound relations (precisely because, as to the model's degree of fit, precise statistics are no guarantee that you have captured exactly what you intended to capture).

Then, the possible theoretical implications of such a change need to be discussed. The very name of **rule system** as well as the citation given above suggest that it would not be admissible to modify the system further than a fuzzy rule system, which allows continuous value logic. As a consequence, evidence for the activation of the less common pronunciation of a grapheme demands a theoretical shift. A new theory would be required in which the lexical and non-lexical routes are organized on the same computational principles, such as interactive networks. But the value of a theory based on these principles would need to be thoroughly justified. In particular, it would be necessary to address the fact that such changes may lead to an extremely weak formulation of the dual-route hypothesis. On the one hand, it would be difficult to establish whether the activation of lexical and sublexical segments takes place inside two separate and non-interfering systems. On the other hand, it would be very difficult to identify how it would differ from the non-lexical system could be set up (Zorzi et al., 1998), for instance, introduced a dual-route connectionist model where a non-lexical system in the form of an associative network) or the less known MROM-P model introduced by Jacobs and colleagues (1998).

DRC is not the only model to be criticized

Do we therefore suggest that connectionist models offer a far better alternative? After all, any consistency effect or competition with multiple pronunciations can be easily explained in any model which, like connectionist models, assimilate knowledge of the print-to-sound regularities with an associative knowledge which reflects diverse levels of regularities and consistency in pronunciation captured after exposure to a representative set of words in the language. “*Written English is ... a quasiregular system... [in which] the relations among entities are statistical rather than categorical*” (Seidenberg & McClelland, 1989, p. 525). Connectionist models can also offer some explanation of how the (emergent) representations develop. But that does not necessarily make them clear winners. At this stage, we cannot be sure that the truth does not lie somewhere in between the current DRC formulation and the alternative connectionist models.

For instance, has not been demonstrated that the *PMSP* model will correctly simulate the effects of grapheme-phoneme consistency and multiplicity of grapheme pronunciations reported here. In the most recent versions, coding is symbolic at the input and output level, representing graphemes and phonemes. Yet, the detour through an intermediate layer and the operation of the backpropagation learning algorithm means that the distinctive print-to-sound relations of English words have to be represented in a **sub-symbolic** mode. Pronunciations are synthesized from the multitude of associations between a single grapheme and the hidden units it is connected to, and these hidden units and the phonemes they are connected to. Although, this can be interpreted as a myriad of micro-inferences occurring in parallel, none of these micro-inferences can be assimilated to a rule and none of the units can be construed as representing a particular rule or infra-lexical segment (coding rimes, graphemes, or specific letter groups). It is not obedience to spelling-sound rules but the degree of the readers’ familiarity or exposure to orthographic-phonological correspondences and the degree of systematicity between the pronunciation of a given word with regard to to the pronunciation of the other words of the language that accounts for reading performance. Connectionist networks reflect a knowledge of the specific pronunciation of irregular words such as *have*, as well as an influence of the pronunciation of visually similar words (lexical analogy), and a knowledge of regularities at various levels of generality (e.g., between body and rime, grapheme and phoneme, etc.). However, because of the format of the input and output representations and/or of the processes that govern the way regularities are captured, the emergent knowledge of print-to-sound relations will be predominantly one of body-rime relations instead of one of grapheme-phoneme relations. As Andrews & Scarratt (1998) showed, the PMSP model is predominantly sensitive to regularities affecting large-size units (body-rime). Note, however, that this does not apply to the ZHB model, which assumes direct relations between letters and phonemes and as a result sees the emergence of letter or letter-in-context regularities instead of body-rime regularities.

In addition, the difficulties of segmenting a letter string into a sequence of non-overlapping graphemes is a problem for any model that posits graphemes at the input of the system. This includes PMSP as well as DRC (or MROM-P). In the current versionof *PMSP*, the input units do not represent individual letters, but, in general, individual graphemes (where "grapheme" means "letter or letter group corresponding to a phoneme"). The tricky job of converting a string of letters into a string of graphemes is not performed by the model; the input submitted to the model is already coded as a set of graphemes. Representations at the output units are also localized: each unit at this level represents a particular phoneme, and any particular phoneme activates just one unit. Thus what the model has to learn is not a many-to-one conversion (the translation of a letter-string to a phoneme-string), but a one-to-one conversion (the translation of a grapheme-string to a phoneme-string).

Finally, connectionist models are also likely to encounter a significant decrease in performance when polysyllabic words are introduced since, due to stress and morphology, ambiguities are a lot more prominent in polsysyllabic words than in monosyllabic words and these models would therefore not benefit as much from the quasi-regular nature of the language (Plaut *et al.*, 1996; Zorzi *et al.*, 1998). The finding that grapheme-phoneme correspondences are a lot more systematic when information about stress and morphology can be provided by an external system might undermine the connectionist hypothesis of an unstructured knowledge of the print-to-sound associated. (cf van den bosch???).

In sum, at this stage, with the knowledge we possess, we cannot clearly declare the eventual superiority of one model over the others. Is it then that it indicates that not enough research effort has been put on the motivation of decisions in current models and that, basically, there is very little we can be sure about; or that it or that it suggests that 20 years later, we are still in the same impasse as the one denounced by Humphreys & Evett (1985): despite years of active research, it appears impossible to identify empirical evidence that could disentangle between models of reading? Should we use these findings to justify the very pessimistic view that a dramatic jump in our understanding of the cognitive processes underpinning reading will never take place?

We believe not. The problem is not the quality of the models available, the ability of the researchers in the field, or the richness of the data currently available. The problem is the approach we take to try to gain a better understanding of the reading processes. Because the field in the seventies, at a time where the recommended approach to science was one of constructing models, models have been designed using the data or intuitions available at the time. Then, the research effort has aimed at the evaluation of the validity of the models, to try to indentify the aspects of the models that conform or not with human performance. In very many studies, existing models have been used to guide experimental research.

The problem then, is that usual approach to model evaluation presents various disadvantages that only become to be addressed. A lot of efforts, in the past 30 years has been made to evaluate the validity of the solutions to the problem of reading proposed by box and arrow or computer models. The problem we have in the field is that what we have spent a lot more time to attempt to define and evaluate solutions (computer models taken or even the framework behind these models) than to try to understand the problem we are trying to find solve. How excellent, refined, clever, a solution can be, it will always be of poor value if it is not a solution to your problem.

Most models of readings assume some abstraction of the reading processes that start with letters and end with phonemes encoded in discrete slots; it is not clear at all that this corresponds to the input and output of a readers' cognitive system. On the contrary, xxx (seriol). Most researchers appear to accept the idea that the domain of one-syllable words is an appropriate abstraction of the material that readers are usually exposed to. This is not at all representative of the readers' experience with written text (this paper contains xx monosyllabic words and xx disyllabic words). In the Introduction section, we highlighted the way polysyllabic words were characterized by word properties very different than the ones of one syllable words. In the results section, we showed that they cannot be processed efficiently using the procedures used for one-syllable words.

[Personal note. Could do a parallel with Software dvt. Not interacting with clients -> solution to the wrong problem. Not scalable. Could also consider a parallel with less symbolic, more feature-based domain of vision. Vertical to horizontal. [Balota et al., 2003].

Rather than targeting a better understanding of the validity of the current solution, it would be more appropriate to target at better understanding of the problem that we are trying to solve. For this, it is crucial to develop richer statistics about words, as the ones of our an analysis of the print-to-sound relations in monosyllabic as well as polysyllabic words to gain some insight into the adequacy of these hypotheses. Then, these statistics should be used to isolate materials able invalidate assumptions made in these different models and design experiments which have the potential to disentangle the correct from the incorrect

Below, we list some questions which, if answered, would constraint models of reading.

There is evidence that graphemes are represented

The frequency effect reported in Lange & Content (1999) when print-to-sound consistency was tightly controlled for might indicate that graphemes are effectively coded after all. This possibility needs to be unambiguously dismissed with a demonstration that such effects cannot be simulated in a model such as *ZHB* model, which only assumes that letters are represented in context. Conversely, Zorzi et al.'s model could be uniquely supported by evidence which invalidates the assumption of the grapheme as a unit of representation in the human cognitive system.

But which graphemes?

If clear evidence can be provided that at least some multiletter units are activated in the course of print-to-sound conversion, then **strict guidelines must be introduced about what constitutes an acceptable grapheme and what does not**. We need a less ambiguous definition of what is included under the notion of grapheme. The marked grapheme option (e.g., $\mathbf{g}_{eiy}$ to /ʒ/) -- which undoubtedly contributes in an important way to the efficiency of the translation, as shown by the system consistency reduction-- does not fit the traditional definition of a grapheme, and neither do the letters that are associated to a sequence of phonemes rather than a single phoneme (e.g., **x** to /ks/ in *taxi* but also xxx, xxx, xxx in disyllabic words). Then, we need to find empirical evidence to determine which categories and possibly which specific graphemes are represented. In our work, different categories of graphemes are identified: silent letter, geminates, etc. The cohesive nature of each of these categories needs to be experimentally assessed. For example if the Rey *et al*. (2000) experiments are meant to prove the existence of graphemes, use that methodology.

What matters? Type or token counts?

Another issue is the exact nature of the system that uses such a statistical knowledge of grapheme-phoneme regularities. Computational models of word recognition such as Zorzi *et al*. (1998), for example, predict that nonword naming times will be predominantly influenced by print-to-sound consistency estimates that consider the number of words exemplifying each association (*type values*). But it predicts that words will be predominantly influenced by the cumulated [summed] lexical frequency (token values) of words exemplifying the associations (*token values*).

Contextual influences

In the same vein, if context is so central to the performance of the model when it is exposed to polysyllabic words, it is then necessary to specify what contexts are compatible within the DRC framework and which are not. Independently of assessing *DRC*, it is worthwhile to attempt to specify which pronunciations can be obtained on the basis of orthographic information alone (limiting the contexts under consideration to the predicted left or right phoneme) and which ones are the result of a transformation of the phonetic codes by a distinct process, external to grapheme-phoneme conversion but which eventually uses the result of the conversion as an input. This would potentially reduce the space of possible hypotheses as, in contrast with the DRC model (which assumes a highly structured and hierarchical system organized in successive levels of representation and which proceeds by applying knowledge of the most typical translation of the current grapheme, followed by the local and eventually lexical constraints which modify the pronunciation of the graphemes), all other models suppose that grapheme-to-phoneme transcoding proceeds from less structured knowledge which rests on the statistical regularities between written forms (orthography) and their realisation (phonology). They therefore predict that sensitivity to orthographic contextual regularities will be indistinguishable from sensitivity to contextual regularities, since both are associated with sub-patterns of regularities [Note however that this is not true if cleaning units interconned with the output level]. At least, there would be no reason to expect that the answer is modified at a very late stage in the conversion process [e.g., due to attractors] when the same answer is awaited by an automatic modification of the codes at the time of their articulatory realization.

Output representations

Another important unanswered question about stress and the way lexical information such as syntactic categories are sometimes associated with predictable changes in pronunciation is whether we keep a list of different ways of pronouncing graphemes (e.g., **ea** in *bead* and *head*) or whether we have rules to specify how alternative pronunciations can be produced: that is, do we keep a set of different ways of pronouncing a word like 'that' or 'there', or do we also have rules to specify how one form of the word may be changed into another? If we have rules to specify changes of form, then what are the units on which these rules apply, phonemes or features? As we mentioned, the centrality of the phoneme as the unit for representing speech as discrete segments is not that well established. Features would make the process of sound changes under lexical influence, stress value, or phonetic phenomena a lot more obvious. One feature or gesture is increased by lexical influence (stress, morphological boundary, syntactic class). Chomsky and Halle, on p. 65 proposed that we should assume "*specifical features marked as plus or minus; and that the phonological rules, as they apply to these representations, will gradually convert these specifications to integers*". Still, dual-route models need to explain how morpheme boundaries are integrated, that is, how a letter string that has been converted into the wrong number of phonemes due to missing a morpheme boundary can be correctly aligned onto a lexical representation (e.g., the second "n" in "greenness"). Do we have to suppose that the output of the two routes is an underlying phonology representation, where all positions match, or should we suppose that the output of the lexical and non-lexical routes are the effective realisation of the string (which then mismatch)? It would be interesting to investigate strings in which the alignment of the phonemes is not the same in the output of the lexical and non-lexical route. A model such as DRC, for instance, would predict an effect of the mismatch in representation: the earlier the boundary, the earlier the mismatch. [difference between irregular & complex words.]

Stress

We mentioned that due to the important drop in predictability of print-to-sound associations when no information about stress is provided, may give an advantage to an hierarchised system in which hypothesis on stress is made in parallel to print-to-sound conversion. It is of course impossible to make any firm claim in the absence of connectionist models adapted to the processing of polysyllabic words. Still, it is again possible to use grapheme-phoneme consistency estimates derived from different analyses to evaluate which one best fit with human performance. For instance, by comparing estimates which use information about stress with estimates without information about stress. If estimates reflecting a knowledge of the stress value correlate better with the naming times than the estimates blind to any information about stress, then, in agreement with Baker and Smith's (1976) data, it suggests that the information about stress is available during conversion rather than at later processing stages.

It would of course also be necessary to integrate findings about the way morphological structure influences both word naming (Henderson, 1985*c*) and stress assignment (Schiller et al., 2005)

Crosslinguistic comparison

Another area that needs to be investigated, in order to validate current theories of reading is cross-linguistic studies. [Frost et al. 2005. Neighborhood effects different in Hebrew – not published yet?]

The proposed empirical evaluations will hopefully benefit from access to exact statistics about language introduced in this study, for variable manipulation in experimental studies

As part of that attempt, we have set-up on-line resources which include a sample lexicon of over 5000 English with words segmented into graphemes and phonemes based on different coding options, the set of spelling rules obtained with these different segmentation options and a Sound Change Applier which automatically derives the pronunciation of words on the basis of a given set of rules. As already stated, these analyses are not at meant to provide a realistic description of what happens in the human mind. They however provide a detailed linguistic description of the way in which print-to-sound ambiguities are distributed in English. These resources may be used for reference (eg, to establish how English uses the digraph **ai**, or spells the diphthong /aɪ/); as a teaching resource (eg, to check that pupils or students can handle the most important English spelling patterns); for general acclimatization to the vagaries of English spelling; or simply for browsing in wonder at the infinite unpredictability with which today's premier international language is written.

ACKNOWLEDGEMENTS

This work in presents an extension of my PhD work, which had for topic "De l'orthographe a la pronunciation: Nature des processus et representations impliqués dans la conversion graphème-phonème" [From print to sound: Nature of the processes and representations involved in grapheme-phoneme conversion]. Initial development of these ideas benefited from discussions with Stephen Monsell and Karalyn Patterson on a M.Phil thesis, with Alain Content, Ronald Peereman, Marc Brysbaert on a PhD thesis. I am also grateful to Sally Andrews and Marcus Taft for their comments on an early draft of this paper. My thanks also go to Catherine Dickie for her careful proof-reading of the paper. Any remaining error is obviously mine.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Learning and thinking about print.* Cambridge, MA: MIT Press.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition, 10*, 565-575.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 234-254.

Andrews, S., & Scarratt, D. R. (1998). Rule and analogy mechanisms in reading nonwords: Hough dou peapel rede gnew wirds? *Journal of Experimental Psychology: Human Perception and Performance, 24*(4), 1052-1086.

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review, 105*(4), 678-723.

Aronoff, M., & Koch, E. (1996). Context-sensitive regularities in English vowel spelling. *Reading and Writing: An Interdisciplinary Journal, 8*, 251-256.

Attneave, F. (1959). *Applications of information theory to psychology: A summary of basic concepts, methods and results.* New York: Holt-Dryden.

Baayen, R. H., Piepenbrook, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM).* xxx: Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Baker, R. G., & Smith, P. T. (1976). A psycholinguistic study of English stress assignment rules. *Language and Speech, 19*, 9-27.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 386-393.

Barron, R. W. (1980). Visual and phonological strategies in reading and spelling. In U. Frith (Ed.), *Cognitive Processes in spelling.* (pp. 195-213). London, England: Academic Press.

Barron, R. W. (1981). Reading skill and reading strategies. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 299-327). Hillsdale, NJ: Erlbaum.

Bauer, D. W., & Stanovich, K. E. (1980). Lexical access and the spelling-to-sound regularity effect. *Memory and Cognition, 8*, 424-432.

Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review, 102*, 146-184.

Berndt, R. S., D'Autrechy, L. D., & Reggia, J. A. (1994). Functional pronunciation units in English words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 977-991.

Berndt, R. S., Reggia, J. A., & Mitchum, C. C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behavior Research Methods, Instruments, & Computers, 19*, 1-9.

Besner, D. (1987). On the relationship between orthographies and phonologies in visual word recognition. In A. Allport, D. MacKay, W. Pring & E. Scheerer (Eds.), *Language perception and production: Relationship between listening, speaking, reading, and writing* (pp. 211-226). London, England: Academic Press.

Besner, D. (1990). Does the reading system need a lexicon? In D. A. Balota, G. B. F. d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 73-99). Hillsdale, NJ: Erlbaum.

Bowey, J. A., & Hansen, J. (1994). The Development of Orthographic Rimes as Units of Word Recognition. *Journal of Experimental Child Psychology, 58*(3), 465-488.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read - A causal connection. *Nature, 301*, 419-421.

Browman, C. P., & Goldstein, L. (1990 *a*). Gestural specification using dynamically defined articulatory structures. *Journal of Phonetics, 18*, 299-320.

Browman, C. P., & Goldstein, L. (1990*b*). Tiers in articulatory phonology, with some implications for casual speech. In T. Kingston & M. E. Beckman (Eds.), Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech (pp. 341-376). Cambridge University Press.

Bullinaria, J. A. (1994). *Representation, learning, generalization and damage in neural network models of reading aloud.* Edinburgh: Neural Networks Research Group, Department of Psychology.

Burgess, S. R., & Lonigan. (1998). Bidirectional relations of phonological sensitivity and prereading abilities: Evidence from a preschool sample. *Journal of Experimental Child Psychology, 70*, 117-141.

Catach, N. (1984). *La phonétisation automatique du français.* Paris: Editions du CNRS.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* Harper & Row.

Clymer, T. (1963/1966). The utility of phonics generalizations in the primary grades. The Reading Teacher, 16/50, 252-258/182-185.

Collins, B. S., & Mees, I. (1984). *The sounds of English and Dutch*. Leiden: E.J. Brill / Leiden University Press.

Collyer, C. E. (1985). Comparing strong and weak models by fitting them to computer-generated data. Perception & Psychophysics, 38, 476-481.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). London: Academic Press.

Coltheart, M. (1979). When can children learn to read and when should they be taught? In T. G. Waller & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice.* (Vol. 1, pp. 1-30). London: Academic Press.

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In O. S. Mar & M. I. Posner (Eds.), *Attention and Performance XI* (pp. 3-37). Hillsdale, NJ: Erlbaum.

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In O. S. Mar & M. I. Posner (Eds.), *Attention and Performance XI* (pp. 3-37). Hillsdale, NJ: Erlbaum.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1197-1211.

Coltheart, M., Curtis, B., Atkins, P., & Haller, N. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*, 589-608.

Coltheart, M., Langdon, R., & Haller, M. (1996). Computational cognitive neuropsychology and acquired dyslexia. In B. Dodd, L. Worrall & R. Campbell (Eds.), *Evaluating Theories of Language: Evidence from Disordered Communication.* London, England: Whurr Publishers.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256.

Davis, C. (2000). *SOLAR: Self organising lexical acquisition and retrieval.* Unpublished PhD thesis, University of New South Wales, Australia.

Dell, G. S., Juliano, C., & Godvingdge, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17*, 149-195.

Dell, G.S., & Juliano, C. (1996). Phonological encoding. In T. Dijkstra & K. DeSmedt (Eds.), *Computational psycholinguistics: Symbolic and connectionist models of language processing*. London: Harvester-Wheatsheaf.

Drewnowski, A., & Healy, A. (1977). Detection errors on the and and: Evidence for readin units larger than the word. *Memory and Cognition, 5*, 636-647.

Fera, P., & Besner, D. (1992). The process of lexical decision: More words about a parallel distributed processing model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 749-764.

Forbes, C. A. (1964). Why Roman Johnny could read. In A. J. Mazurkiezwicz (Ed.), *New Perspectives in Reading Instruction.* New York, NJ: Pitman.

Forster, K. I. (1994). Computational modeling and elementary process analysis in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1292-1310.

Fowler, C. A., & Rosenblum, L. D. (1989). *The perception of phonetic gestures.* (No. SR 99/100).

Frauenfelder, U. H., Baayen, R. H., Hellwig, F. M., & Schreuder, R. (1993). Neighborhood density and frequency across languages and modalities. *Journal of Memory and Language, 32*(6), 781-804.

Frauenfelder, U., & Content, A. (1996). La perception de la parole et la reconnaissance des mots: apport d'une approche computationnelle. In *Actes des XXIèmes journées d'étude sur la parole* (pp. 1-9).

Gak, V. G. (1976). *L'orthographe du français: Essai de description théorique et pratique.* Paris: SELAF.

Gay, Thomas. (1981). Mechanisms in the Control of Speech Rate. Phonetica 38: 148-158.

Gimson, A. C. (1980, 3d Ed). *An introduction to the pronunciation of English*. London, England: Arnold.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 674-691.

Goswami, U., & Bryant, P. E. (1990). *Phonological skills and learning to read.* London, England: Erlbaum.

Grainger, J., & Jacobs, A. M. (1994). A dual read-out model of word context effects in letter perception: Further investigations of the word superiority effect. *Journal of Experimental Psychology: Human Perception and Performance (special Section: Modeling visual word recognition), 20*(6), 1158-1176.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*(3), 518-565.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences, 8*(2), 58-59.

Haas, W. (1970). *Phono-graphic translation*. Manchester: Manchester University Press.

Hanna , J. S., & Hanna, P. R. (1959). Spelling as a school subject: brief history. *National Elementary Principal, 38*, 8-23.

Hanna, J. S., & Hanna, P. R. (1965). The teaching of spelling. *National Elementary Principal, 45*, 19-28.

Hanna, P. R., Hanna, J. S., Hodges, R. E., & Ruderf, E. H. (1966). *Phoneme-grapheme correspondences as cues to spelling improvement.* Washington, DC: US Government Printing Office.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, Reading acquisition and Dyslexia: Insights from connectionist models. *Psychological Review, 106*(3), 491-528.

Harris, M., & Coltheart, M. (1986). *Language processing in children and adults.* London: Routledge & Kegan Paul.

Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic Press.

Henderson, L. (1985b). On the use of the term "grapheme". *Language and Cognitive Processes, 1*, 135-148.

Henderson, L. (1985c). Towards a psychology of morphemes. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1). Hove: Erlbaum.

Henderson, L., & Chard, J. (1980). The reader's implicit knowledge of orthographic structure. In U. Frith (Ed.), *Cognitive Processes in spelling.* (pp. 85-116). London, England: Academic Press.

Herdman, C. M., LeFevre, J.-A., & Greenham, S. L. (1994). Implicating the lexicon: Base-word frequency effects in pseudohomophone naming. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 575-590.

Hill, L. A., & Ure, J. M. (1962). *English sounds and spellings.* London, England: Oxford University Press.

Humphreys, G. W., Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *Behavioral and Brain Sciences, 8*, 689-740.

Jacobs, A., & Grainger, J. (1999) . Commentary on 'Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.'

Jacobs, A. M., Rey, A., Ziegler, J. C., & Grainger, J. (1998). MROM-P: An interactive activation, multiple read-out model of orthographic and phonological processes in visual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition*. Mahwah, NJ: Erlbaum.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36*(4), 505-529.

Jared, D.& Chateau.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29*, 687-715.

Kawamoto, A. H., Kello, C. T., Jones, R., & Bame, K. (1998). Initial phoneme versus whole-word criterion to initiate pronunciation: Evidence based on response latency and initial phoneme duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(4), 862-885.

Kay, J. (1987). Phonological codes in reading: Assignment of sub-word phonology. In D. A. Allport, D. Mackay, W. Prinz & E. Scheerer (Eds.), *Language perception and production: Shared mechanisms in listening, speaking, reading and writing* (pp. 181-196). London, England: Academic Press.

Kay, J., & Lesser, R. (1985). The nature of phonological processes in reading: Evidence from surface dyslexia. *Quarterly Journal of Experimental Psychology, 37A*(1-A), 39-81.

Kay, J., & Marcel, A. J. (1981). One process, not two, in reading aloud: Lexical analogies do the work of nonlexical rules. *Quarterly Journal of Experimental Psychology, 33A*, 397-413.

Kelly, M. H., Morris, J., & Verrekia, L. (1998). Orthographic cues to lexical stress: Effects on naming and lexical decision. *Memory and Cognition, 26*(4), 822-832.

Kessler, B. (1995). Regularity in the phonemic subsystem of English spelling: A study of the monosyllables. (Unpublished report).

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in english syllables. *Journal of Memory and Language, 37*, 295-311.

Klima, E. (1972). How alphabets might reflect language. In J. Kavanagh & I. Mattingly (Eds.), *Language by ear and by eye: the relationship between speech and reading* (pp. 57-80). Cambridge, MA: MIT Press.

Krakow, R. (1999). Physiological organization of syllables: a review. Journal of Phonetics, 27, 23-54

Kreidler, C. W. (1990). *The Pronunciation of English.* Malden, MA: Blackwell Publishing.

Kruisinga, E. (1957). *An introduction to the study of English sounds* (11th ed, 1st 1914 ed.). Groningen, Hollqnd: P. Noordhoof N.V.

Lange, M., & Content, A. (1999, June 20–26). *The grapho-phonological system of written French: Statistical analysis and empirical validation.* Paper presented at the 37th Annual Meeting of the Association for Computational Linguistics, University of Maryland, Wahsington DC (United States).

Lange, M., & Content, A. (2001). Activation of multiple phonemic associates of graphemes in visual word recognition. *Brain and Language, Submitted.*

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revisited. *Cognition, 21*, 1-36.

Liberman, A., Cooper, F., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431-461.

Lukatela, G., Eaton, T., Lee, C., & Turvey, M. T. (2001). Does visual word identification involve sub-phonemic level? *Cognition, 78*, B41-B52.

Marckwardt, A. H. (1957). *Introduction to the English Language.* (11th (1st, 1942). ed.). NY, USA: Oxford University Press.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101*(4), 653-675.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word-recognition. *Memory and Cognition, 2*(2), 309-321.

Mowrey, R. A., & MacKay, I. R. A. (1990). Phonological primitives: Electromyographic speech error evidence. Journal of the Acoustical Society of America, 88, 1299-1312.

Norris, D., & Brown, G. (1985). Race models and analogy theories: A dead heat? Reply to Seidenberg. *Cognition, 20*, 155-168.

Norris, D., & Brown, G. (1985). Race models and analogy theories: A dead heat? Reply to Seidenberg. *Cognition, 20*, 155-168.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53*, 13-24.

Paap, K. R., Noel, R. W., & Johansen, L. S. (1992). Dual-route models of print to sound: Red herrings and real horses. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 293-318). Amsterdam: Elsevier.

Parkin, A. J. (1982). Phonological recoding in lexical decision: Effect or spelling-to-sound regularity depend on how regularity is defined. *Memory and Cognition, 10*, 43-53.

Parkin, A. J. (1984). Redefining the regularity effect. *Memory and Cognition, 12*, 287-292.

Parkin, A. J., & Ellingham, R. (1983). Phonological recoding in lexical decision: The influence of pseudohomophone. *Language and Speech, 26*, 81-90.

Parkin, A. J., McMullen, M., & Graystone, D. (1986). Spelling-to-sound regularity affects pronunciation latency but not lexical decision. *Psychological Research, 48*, 87-92.

Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 335-359). Hillsdale, NJ: Erlbaum.

Peereman, R. (1991). Phonological assembly in reading: Lexical contribution leads to violation of graphophonological rules. *Memory and Cognition, 19*, 568-578.

Peereman, R., & Content, A. (1997). Quantitative analyses of orthography to phonology mapping in English and French. *Submitted to the computational Psycholinguistics workshop.*

Perfetti, C., Beck, I., Bell, L. & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal, Merrill-Palmer Quarterly, 33, 283-319.

Perry, C. (2000). German, English, the Length Effect, and Computational Models of Reading Aloud Speaker(s). Weekly seminars at the Macquarie Center for Cognitive Science, September 1, 2001.

Plaut, D. C. (1999). A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. *Cognitive Science, 23*(4), 543-568.

Plaut, D. C., & McClelland, J. L. (1993). *Generalization with componential attractors: Word and nonword reading in an attractor network.* Paper presented at the Proceedings of the 15th Annual Conference of the Cognitive Science Society, 824-829.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115.

Pring, L. (1981). Phonological codes and functional spelling units: Reality and implications. *Perception & Psychophysics, 30*, 573-578.

Rastle, K., & Coltheart, M. (1998). Whammies and double whammies: The effetct of length on nonword reading. *Psychonomic Bulletin & Review, 5*(2), 277-282.

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25*(2), 482-503.

Rastle, K., & Colthearth, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language, 42*(3), 342-364.

Reggia, J. A., Marsland, P. M., & Berndt, R. S. (1988). Competitive dynamics in a dual-route connecionist model of print to sound translation. *Complex Systems, 2*, 509-517.

Rey, A., Ziegler, J. C., & Jacobs, A. M. (2000). Graphemes are perceptual reading units. *Cognition, 75*(1), B1-B12.

Roach (2002). A little encyclopedia of phonetics. [on-line: http://www.personal.rdg.ac.uk/~llsroach/fue/encyc.pdf]

Roach, P. (1995, 2d Ed.). *English phonetics and phonology*. Cambridge, England: Cambridge University Press.

Roelofs, A. (1999). Phonological segments and features as planning units in speech production. *Language and Cognitive Processes, 14*(2), 173-200.

Rosson, M. B. (1985). The interaction of pronunciation rules and lexical representations in reading aloud. *Memory and Cognition, 13*, 90-99.

Schiller, N. O., Jansma, B. M., Peters, J., & Levelt, W. J. M. (in press). Monitoring metrical stress in polysyllabic words. Language and Cognitive Processes.

Schneider, T. D. (1995). *Information theory primer. [on-line:* ftp://ftp-ncifcrf.gov/pub/delila/primer.]

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19*, 1-30.

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

Seidenberg, M. S., & McClelland, J. L. (1990). More words but still no lexicon: Reply to Besner et al. (1990). *Psychological Review, 97*, 447-452.

Seidenberg, M. S., & Plaut, D. C. (1998). Evaluating word-reading models at the item level: Matching the grain of theory and data. *Psychological Science, 9*(3), 234-237.

Seidenberg, M. S., Plaut, D. C., Petersen, A. S., McClelland, J. L., & McRae, K. (1994). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1177-1196.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23*, 383-404.

Shallice, T., & McCarthy, R. (1985). Phonological reading: From patterns of impairment to possible procedure. In K. E. Patterson, J. C. Marshall & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 361-397). Hillsdale, NJ: Erlbaum.

Shallice, T., Warrington, E. K., & McCarthy, R. (1983). Reading without semantics. *Quarterly Journal of Experimental Psychology, 35A*, 111-138.

Shannon, C. E. (1948a). A mathematical theory of communication. Part I. *Bell System Technical Journal, 27*, 379-423.

Shannon, C. E. (1948b). A mathematical theory of communication. Part II. *Bell System Technical Journal, 27*, 623-656.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151-218.

Shillcock, R., & Monaghan, P. (1999). Bihemispheric representation, foveal splitting, and visual word recognition. *Behavioral & Brain Sciences, 22*(2).

Smith, P. T., Meredith, T., Pattison, H. M., & Sterling, C. (1983). The representation of internal word structure in English. In L. Henderson (Ed.), *Orthographies and Reading.* London, England: Erlbaum.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science, 8*(6), 411-416.

Stahl, 1992 – some references by looking in Google for ‘perfetti beck bell hughes 1987

Stanback, M. 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.

Stanhope, N., & Parkin, A. J. (1987). Further explorations of the consistency effect in word and nonword pronunciation. *Memory and Cognition, 15*, 169-179.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English Orthography. *Journal of Experimental Psychology: General, 124*(2), 107-136.

Treiman, R., Zukowski, A., & Richmondwelty, E. D. (1995b). What Happened to the "n" of Sink? - Children's Spellings of Final Consonant Clusters. *Cognition, 55*(1), 1-38.

Vallins, 1965

Venezky, 1999. Book, revised of 1970

Venezky, R. L. (1970). *The structure of English orthography.* The Hage, The Netherlands: Mouton.

Venezky, R. L. (1995). How English is read: Grapheme-phoneme regularity and orthographic structure in word recognition. In I. Taylor & D. R. Olson (Eds.), *Scripts and literacy: Reading and learning to read alphabets, syllabaries and characters. Neuropsychology and cognition* (Vol. 7, pp. 111-129). Dordrecht, Netherlands: Kluwer Academic Publishers.

Warren, P., & Marslen-Wilson, W.D. (1988) Cues to lexical choice: Discriminating place and choice. *Perception and Psychophysics*, 43, 21-30.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory and Cognition, 13*, 557-572.

Waters, G. S., Seidenberg, M. S., & Bruck, M. (1984). Children's and adults' use of spelling-sound information in three reading tasks. *Memory and Cognition, 12*, 293-305.

Weekes, B. (1997). Differential effects of letter number on word and nonword naming latency. *Quarterly Journal of Experimental Psychology, 50A*(2), 439-456.

Whalen, D. (1990). Coarticulation is largelly planned. *Journal of Phonetics, 18*, 3-35.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The seriol model and selective literature review. *Psychonomic Bulletin & Review, 8*(2), 221-243.

Wijk, A. (1966). *Rules of pronunciation for the English language.* Oxford, England: Oxford University Press.

Wylie & Durrell, 1970'

Zhang, J. (1995). *Automatic learning of english pronunciation rules.* Unpublished Master of Science, University of Regina, Regina, Saskatchewan.

Zhang, J. (1995). *Automatic learning of english pronunciation rules.* Unpublished Master of Science, University of Regina, Regina, Saskatchewan.

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for _OUGH and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments, & Computers, 29*(4), 600-618.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-route process model. *Journal of Experimental Psychology: Human Perception and Performance, 24*(4), 1131-1161.

[Levelt (in Gazzaniga, p.3), "Most neuroimaging work in word production involves monomorphemic words and hence reveals nothing about the production of complex morphology". [Indefrey, P., & **Levelt, W.J.M.** (2000). The neural correlates of language production. In M.Gazzaniga (Ed.), The new cognitive neurosciences (2nd ed.) (pp. 845-865). Cambridge, MA: MIT Press.]]