

InfoSyll: A Syllabary Providing Statistical Information on Phonological and Orthographic Syllables

Fabienne Chetail · Stéphanie Mathey

Published online: 27 December 2009
© Springer Science+Business Media, LLC 2009

Abstract There is now a growing body of evidence in various languages supporting the claim that syllables are functional units of visual word processing. In the perspective of modeling the processing of polysyllabic words and the activation of syllables, current studies investigate syllabic effects with subtle manipulations. We present here a syllabary of the French language aiming at answering new constraints when designing experiments on the syllable issue. The InfoSyll syllabary provides exhaustive characteristics and statistical information for each phonological syllable (e.g., /fi/) and for its corresponding orthographic syllables (e.g., *fi*, *phi*, *phy*, *fee*, *fix*, *fis*). Variables such as the type and token positional frequencies, the number and frequencies of the correspondences between orthographic and phonological syllables are provided. As discussed, such computations should allow precise controls, manipulations and quantitative descriptions of syllabic variables in the field of psycholinguistic research.

Keywords Written word processing · Phonological syllable · Orthographic syllable · Syllabic variables

Introduction

There have been a considerable number of studies on monosyllabic word processing in the psycholinguistic literature on reading. It now seems essential to focus on polysyllabic word

This research was supported by an MESR grant to the first author. We are very grateful to the reviewers for their constructive comments on an earlier version of the manuscript.

F. Chetail · S. Mathey
Laboratoire de Psychologie, Université Bordeaux 2, Bordeaux, France

F. Chetail (✉)
Laboratoire Cognition Langage et Développement, Université Libre de Bruxelles, Av. F. Roosevelt,
50-CP 191, 1050 Brussels, Belgium
e-mail: fchetail@ulb.ac.be

processing given that monosyllabic words account for a small part of the whole lexicon of a skilled reader (Brand et al. 2003). This has strong implications for reading models since they are currently dedicated to the processing of monosyllabic words (see Ans et al. 1998, for an exception). With regard to polysyllabic words, the syllable is a major infralexical unit that has received much attention in both the written and the spoken modalities. On the reading issue specifically, studies on the role of syllables have become more and more numerous and the syllabic effects investigated have become more and more subtle. This has engendered the need to use specific tools providing characteristics of syllables. In the first part of this paper, we present data from the literature on the role of syllables in visual word processing and on its relationships with orthographic information. Then we describe the interest of a new tool specifically focused on the syllables of the French language and dedicated to the construction of syllabic materials. In the second part, we present the InfoSyll syllabary, which is a database including numerous variables on the French syllables computed on the words of the lexical database Lexique (New et al. 2004). Finally, we present some quantitative descriptions of the syllabic characteristics of the French language computed from the syllabary.

The Syllabic Issue in Reading

Evidence for Syllabic Effects

Strong evidence in favour of the functional role of syllables in reading has been provided by the investigation of the syllable frequency effect (Carreiras et al. 1993). Words with a first syllable of high frequency were compared to words with a first syllable of low frequency in reading tasks. In the lexical decision task, Carreiras et al. (1993) observed that skilled readers in Spanish decided more slowly that words with a high-frequency first syllable were real words than words with a low-frequency first syllable. This inhibitory effect of syllable frequency has been widely replicated in follow-up studies in Spanish (e.g., Alvarez et al. 2000, 2001; Conrad et al. 2008; Perea and Carreiras 1998), French (e.g., Chetail and Mathey 2009b; Conrad et al. 2007; Mathey and Zagar 2002; Mathey et al. 2006) and German (e.g., Stenneken et al. 2007; Conrad and Jacobs 2004). In reading aloud also, syllable frequency effects have been reported in the naming task, the effects being either facilitative (e.g., Macizo and Van Petten 2007, in English; Perea and Carreiras 1998, in Spanish) or inhibitory (e.g., Carreiras et al. 1993, in Spanish; Conrad et al. 2006, in German).

Another key effect to assess the role of syllables in reading is the syllable congruency effect, initially investigated in the auditory modality by Mehler et al. (1981). They found that segments such as /pa/ or /pal/ were better detected in words when they exactly corresponded to the first syllable of the word (e.g., /pa/ in *pa.lace*, /pal/ in *pal.mier*) rather than when they corresponded to the first phonemes only (e.g., /pa/ in *pal.mier*, /pal/ in *pa.lace*).¹ In the visual adaptation of the design, a prime such as a segment (e.g., *pa&&&&*) or a pseudo-word containing the segment (e.g., *palute*) is briefly displayed before the target word (e.g., *PALMIER*) (see Ferrand et al. 1996). A facilitatory priming effect was obtained with prime duration up to 60 ms in the lexical decision task (Alvarez et al. 2004; Carreiras and Perea 2002, in Spanish; Chetail and Mathey 2009a, in French). The target words were recognized faster when they were preceded by a prime that shared the first syllable (e.g., *pa&&&& -PALACE*) rather than not (e.g., *pal&&&& -PALACE*). The data argued for the phonological nature of syllabic effects given that the syllable congruency effect was obtained even when

¹ Hereafter, the dots mark syllable boundaries, though the items presented did not contain the dots.

the pseudoword prime and the word target shared only the phonological feature of the first syllable (e.g., Alvarez et al. 2004, in Spanish : *bi.rel-VI.RUS*, *bi* and *vi* corresponding to the same oral syllable /bi/ while the spelling is different). Syllable congruency effects have also been found in the naming task, though the data are more contrasted. Some studies reported facilitatory syllable congruency effects (Chetail and Mathey 2009a; Ferrand et al. 1996 in French; Ferrand et al. 1997, in English) while others did not (e.g., Brand et al. 2003; in French; Schiller 1998, in Dutch; Schiller 2000, in English). In particular, segmental priming effects were found rather than syllable congruency effects. In that case, priming effects increased as the number of letters shared by the prime and the target increased, whatever the syllabic structure of the target word (e.g., priming effect larger in the condition *pi%%-PILOT* than in the condition *pi%%%-PILOT*) (e.g., Schiller 1998, 2000). Thus the reliability of the syllable congruency effects in reading aloud remains a debated question (see Schiller 2006, for a recent discussion).

Taken together, the findings concerning both syllable frequency and syllable congruency in lexical decision and naming tasks suggest that syllables are functional units of written word processing. To accommodate the lexical decision data, it has been proposed that not only a level of letter representations and a level of word representations are involved during the access to the mental lexicon (McClelland and Rumelhart 1981), but also that there is an intermediate level of syllable representations (Conrad et al. 2009; Mathey et al. 2006). In this framework, syllabic activation would ensue from two complementary processes. First, when a written word is displayed, the corresponding syllables are activated and activate in turn the corresponding lexical representation by means of facilitative connections. Second, the syllables also send activation to the other words sharing these syllables (i.e., the syllabic neighbours), and these words enter into competition with each other by means of the lexical inhibition process. Consequently, the processing of words with a high-frequency first syllable is slowed down compared to that of words with a low-frequency first syllable, given that in the former case, more words enter into competition with each other. In addition, it has been proposed to encode syllable frequency in the syllabic level by using a resting activation level based on frequency, which would also increase the activation speed of high-frequency syllables (Mathey et al. 2006). A similar interactive activation framework incorporating syllable units has been considered to account for syllabic effects in naming (e.g., Chetail and Mathey 2009a). In particular, inhibitory effects are assumed to ensue from lexical competition between syllabic neighbours while facilitatory effects are assumed to ensue from facilitation during the retrieval in the mental syllabary of the articulatory-phonetic syllable programs required for word pronunciation (see Levelt and Wheeldon 1994). In the latter case, the post-lexical process of syllable retrieval might shadow the lexical competition effects (Macizo and Van Petten 2007).

From a methodological point of view, the way syllabic variables are defined and manipulated is important because it underlies the empirical evidence. Usually, variables are selected similarly both in naming and lexical decision tasks given that the same (or the same kind of) materials is used in both tasks (e.g., Carreiras et al. 1993; Chetail and Mathey 2009a; Ferrand et al. 1996, 1997). From the first experiments on the syllable frequency effect, it has been clear that positional syllable frequency needs to be considered, mainly in bisyllabic words. Positional frequency refers to the number of occurrences of syllables at a given position in words. Taking into account positional frequency seems particularly relevant for the issue of syllabic activation, given that syllabic effects are position-sensitive, the first syllable acting more than the remaining ones (e.g., Alvarez et al. 2000; Tousman and Inhoff 1992). Hence, syllables at the first and second or even the third positions in words are usually either manipulated or controlled and positional syllable frequencies are used. Another criterion for

syllable frequency computations is related to the distinction between type count (number of words sharing a given syllable) and token count (sum of the printed frequency of all words sharing a given syllable). This distinction is important since type and token frequencies could tap into different processes and can lead to opposite effects, even though these two measures of syllable frequency are highly correlated (e.g., Alvarez et al. 2001). However, type and token syllable frequencies have been widely confounded in experiments on syllables to date. It has been shown only recently that manipulating the token count of syllable frequency leads to inhibitory syllable frequency effects, while manipulating the type count of syllable frequency yields facilitatory syllable frequency effects (Conrad et al. 2008). This suggests that these are important variables to be taken into account in current experiments.

In addition, investigating the syllable congruency effect has implied contrasting specific syllabic structures (commonly, consonant-vowel syllables such as /pa/ vs. consonant-vowel-consonant syllables such as /pal/) in bisyllabic words. Beyond the necessities of this specific experimental situation per se, the syllabic structure is a variable which has received little consideration to date in skilled reading. However, some data indicated that words beginning with a consonant-vowel-consonant first syllable were processed longer than words beginning with a consonant-vowel first syllable (Chetail and Mathey 2009a), and that syllable congruency effects were more reliable for words beginning with a consonant-vowel first syllable (Alvarez et al. 2004), thus suggesting the need to take this variable into account. In the same vein, little information is provided about the establishment of syllable boundaries when selecting words for experiments on syllables. However, to bypass the complex issue of setting boundary locations in polysyllabic utterances, words are usually selected so that there is no case of ambisyllabicity (i.e., when an internal consonant could belong either to the coda of first syllable or to the onset of the next syllable, such as in *panic* which can be segmented both into *pa.nic* and *pan.ic*). This precaution made it possible to manipulate syllables which are clearly extracted from speech utterances (e.g., Carreiras and Perea 2002; Chetail and Mathey 2009a; Ferrand et al. 1996). Finally, although polysyllabic words have usually been used in studies investigating syllabic effects, a few studies have also used monosyllable words to investigate the activation of the syllables in relation to orthographic information (Mathey et al. 2006, in the lexical decision task; see also Doignon-Camus et al. 2009; Seidenberg 1987; in the illusory conjunction task).

Syllabic Effects and Orthographic Information

Examining the role of syllables in reading differs substantially from examining their role in speech processing. Though syllables are phonological units, their orthographic features are also processed when displayed visually. This has led to examining the influence of orthographic information carried by words and syllables in syllabic effects and the relationships between syllabic and orthographic units.

This issue was first discussed by Taft (1979) in the English language. He argued that syllables are involved in visual word processing, but that the syllabification of words is based upon orthographic principles rather than pronunciation. This proposal was based on the strong ambisyllabicity of the English language, and on the cases of incompatibility between syllabification following phonological principles and the morphemic structure of words (e.g., the word *leaking* has a morphemic boundary after the letter *k* while the syllable boundary is after the letter *a*). Hence, Taft (1979) proposed that the basic orthographic syllabic structure (BOSS) is a functional unit of polysyllabic word processing, the BOSS referring to the first syllable of a word plus as many consonants following its first vowels while not disrupting the

morphemic structure (e.g., the BOSS of *leaking* is *leak*) (e.g., Rouibah and Taft 2001; Taft 1979, 2001; but see Lima and Pollatsek 1983, in lexical decision tasks). From the point of view of orthographic regularity, Seidenberg (1987) used the illusory conjunction task (i.e., to detect the colour of a letter in a briefly-presented word displayed in two colours) to assess to what extent the syllabic effects in English were subsumed by statistical orthographic regularities. More precisely, given that the bigram that straddles the syllabic boundary of words is usually less frequent than those that precede and follow this boundary, this would produce an orthographic cue to isolate perceptual units such as syllables. This particular pattern of bigram frequencies, referred to as a bigram trough, would thus be responsible for the emergence of syllabic effects. Additional data in French and Spanish actually suggest that the frequency of orthographic units at the syllable boundary can influence the strength of syllabic effects. Syllabic effects were obtained for words both with and without a bigram trough but the effects were stronger when the bigrams made salient the syllables of the words, that is, when there was a bigram trough at the syllable boundary rather than not (illusory conjunction task: Doignon and Zagar 2005; lexical decision task: Carreiras et al. 1993; but see Conrad et al. 2009). In line with these data, Mathey et al. (2006) showed that the frequency of the first bigram of French words influences the strength of the syllabic frequency effects in the lexical decision task, low-frequency bigrams yielding facilitatory syllable frequency effects while high-frequency bigrams yielded inhibitory syllable frequency effects.

Another approach to the relationships between syllabic effects and orthographic information lies in investigating the orthographic information contained by the syllables themselves rather than by the words. This leads to the important distinction between the phonological features and the orthographic features of the syllables, and thus to distinguishing between syllable units (phonological syllables) and the written transcription of these units (orthographic syllables). For example, the syllable /si/ can be spelled either *si* or *ci* in French (*sirène* vs. *citron*) and in Spanish (*ciclón* vs. *sitiar*). The question of the role of orthographic syllables in reading has emerged very recently in the literature on the syllable issue. Studies using the primed lexical decision task have shown that when the prime shared the first phonological syllable with the target word but not the orthographic transcription of this syllable, a syllabic priming effect was obtained (e.g., Alvarez et al. 2004). Applied to syllabic neighbourhood, the distinction between orthographic and phonological syllables leads to the distinction between phonological syllabic neighbourhood (i.e., whole words sharing a phonological syllable, like the syllable /fi/ in *ficelle*, *physique*, *philosophie*, in French) and orthographic syllabic neighbourhood (i.e., whole words sharing an orthographic syllable, like the syllable /fi/ spelled *fi* in *filer*, *ficelle*, *fillette*, *ficher*). When the effects of syllable frequency were separately examined according to the syllable features in French (Conrad et al. 2007), data yielded an inhibitory effect of phonological syllabic neighbourhood (orthographic syllabic neighbourhood held constant) and no effect of orthographic syllabic neighbourhood (phonological syllabic neighbourhood held constant). Furthermore, when the phonological syllabic neighbourhood and the orthographic syllabic neighbourhood were jointly manipulated, the inhibitory effect of syllable frequency was obtained only for low-frequency orthographic syllables in adults, while the reverse pattern was found in fifth graders in French (Chetail and Mathey 2009b). These data suggest that the orthographic transcriptions of syllables is an important variable and that orthographic information carried by syllables should be taken into account when studying syllable effects, at least in French.

The distinction between phonological and orthographic syllables is particularly relevant in languages with inconsistent phonology to spelling correspondences, such as French or English (see Ziegler et al. 1996). The question of consistency between phonology and orthography can be examined on sound-to-print relationships (feedback consistency: number of

ways to spell a pronunciation pattern) and on print-to-sound relationships (feedforward consistency: number of ways to pronounce a spelling pattern). In French specifically, these two measures of consistency are not symmetric, feedforward consistency being much higher than feedback consistency. This was found at different intra-syllabic levels. When considering the rime of monosyllabic words, 87.6% are feedforward consistent while only 20.9% of words are feedback consistent (Ziegler et al. 1996). Similarly on the onset, nucleus, and coda units, words are more feedforward consistent (respectively 77.30, 35.95, and 61.13%) than feedback consistent (respectively 52.02, 1.22, and 14.45%) (Peerean and Content 1999a). This means that in French, the pronunciation of intra-syllabic units in written words is somewhat ambiguous while such pronounced units lead to various spellings. As underlined by Peerean and Content (1999b), the consistency between prints and sounds can be analyzed for units at different levels of word structure. This seems particularly relevant when examining the role played by a given unit in a given language. In English for example, correspondences between phonology and orthography are more consistent at rime level than at a smaller level, supporting the idea that rimes are important units of word processing in this language (Treiman et al. 1995). In French, given that syllables are more relevant units than rimes, a natural question concerns the degree of consistency at the syllabic level compared to smaller levels. However, no answer could be provided until now given that correspondences between oral syllables and their spellings were not available.

Value of a French syllabary

As mentioned above, studies on the syllable issue have become more numerous during recent decades and the syllabic effects investigated have become increasingly precise. The syllable frequency effect was first examined in contrasting words with high-frequency syllables and words with low-frequency syllables (Carreiras et al. 1993). Since this seminal study, more complex issues have emerged, with subtle manipulations of syllabic variables such as the phonological feature of syllables (e.g., Alvarez et al. 2004), the way to compute syllable frequency (e.g., Conrad et al. 2007) or the frequency of syllable spelling (e.g., Chetail and Mathey 2009b). In addition, current models of visual word recognition incorporate syllable units (Conrad et al. 2009; Mathey et al. 2006) and even syllable frequency resting levels (Mathey et al. 2006), so it is important to provide adequate tools to test such model predictions in the French language. More generally, faced with the issue of the processing of polysyllabic words and the increasing complexity of experiments on the syllable issue, specific tools on syllable characteristics are therefore essential for designing further experiments. Specifically, a syllabary (i.e., a database based on syllables) would provide an exhaustive description of the characteristics of both the phonological and the orthographic syllables encountered in words. Hence, such a database would enable researchers to define precisely which syllables should be selected in order to meet an experimental objective, and then to select the words containing these syllables in current lexical databases. Syllable and lexical databases might therefore be used in a complementary way. In addition, such a syllabary makes it possible to select specific syllables in order to study their processing when they are presented in isolation (Stenneken et al. 2007) and also makes it possible to build pseudowords according to some syllabic criteria in order to investigate their processing (e.g., Carreiras and Perea 2004). A syllable database would also be useful for determining syllable characteristics of words not occurring in databases. Finally, beyond facilitating the construction and controls of materials for conducting psycholinguistic experiments, a syllabary can be used for the description of the syllabic characteristics of the language. As underlined by Peerean and Content (1999a), new relevant variables to be considered might therefore emerge from such descriptions.

Currently, the syllabic variables available in the French language concern the number of syllables of words (databases Brulex: [Content et al. 1990](#); Lexique : [New et al. 2004](#); Vocolex: [Dufour et al. 2002](#)), the syllabified forms of the words along with the abstract phonological syllabified form (e.g., the word /solej/ is syllabified in *so – lej*; [Dufour et al. 2002](#); [New et al. 2004](#)) and the frequencies of phonological syllables for each word ([New et al. 2004](#); [Peereman et al. 2007](#)). However, in these databases, syllabic measures are only a small part of the statistics on words. Specifically, current databases do not provide statistical information on the orthographic syllables of French (frequencies, structures, correspondence with the phonological syllables), though these data are essential to investigate the influence of the orthographic information in syllabic effects, and more generally, the activation of syllables in the visual modality (e.g., [Davis and Perea 2005](#), in Spanish).

To meet the need for detailed measures about syllables in French, we present here a new syllabary gathering precise information on the phonological syllables of French, along with detailed computations on their corresponding orthographic syllables. In light of previous studies and of current issues on syllables, data on the frequency of occurrences of the syllables at the first, second, third and last positions in words (i.e., positional frequencies) are presented. Given the differential role of type and token syllable frequencies, both types of computation are provided. The length and structure of the syllables are also indicated, which should help in examining the role of these factors in written word processing. In the syllabary, these syllabic variables (type and token positional frequencies, structure, length) are computed for both phonological and orthographic syllables. This means that for each phonological syllable, the number, list, frequencies and structure of each different spelling of the syllables are available. Conversely, the database allows to find quickly the different phonological syllables associated with an orthographic syllable. In addition, to go further into the issue of consistency between sounds and prints, measures of both feedback and feedforward consistency at a syllabic level are available. In that sense, the InfoSyll database constitutes an appropriate tool both to select syllabic materials for running experiments aiming at grasping precisely the process of activation of syllables in written word processing and to provide descriptions of syllabic regularities in the French language.

Statistical Information in the InfoSyll Syllabary

Lexical Corpus and Preliminary Computations

Before running the syllable computations for the InfoSyll syllabary, the phonological and orthographic syllabifications of a pool of words corresponding to an adult reader lexicon were required. For this purpose, we used the lexical database Lexique ([New et al. 2004](#)) which includes about 40,000 words along with their inflected forms extracted from a lexical corpus composed of texts written from 1950 to 2000 (total of 31 millions of words). For each word, information such as orthographic and phonological representations, grammatical category, lexical frequency, number of letters and phonemes, and number of orthographic and phonological neighbours is available. In addition, this lexical database includes the phonological syllabification of words, from which we computed the orthographic syllabic segmentation of the words.

The syllabified phonological words we used for InfoSyll consist of the lemmas of the Lexique database ([New et al. 2004](#)) from which we removed compound entries, abbreviations and onomatopoeias. This lexical corpus was composed of 8.87% monosyllabic words, 34.22% bisyllabic words, 35.86% trisyllabic words, 15.68% four-syllable words, and 5.42%

words with five or more syllables. The word phonological representations in Lexique, are syllabified according to an algorithm of syllabification (Pallier 1994), this algorithm being based on the rules of syllabification proposed by Pulgram (1970). In this algorithm, the final schwas in polysyllabic words were cancelled (e.g., /nɔ̃trə/ becomes /nɔ̃tr/). In regular sequences such as VCVCVC or CVCYV (e.g., /abisal/, /pɔ̃lɛ/ respectively), syllabification is performed so that open syllables are consistently created (V.CV.CVC: /a.bi.sal/, CV.CYV: /pɔ̃.lɛ/). Similarly, when two phonemes follow one each other such as in the CVVC sequence (e.g., /pɔ̃ɛt/), syllabification is performed between the two phonemes (/pɔ̃.ɛt/). Given that ambiguity in segmenting phonetic strings arises in presence of groups of medial consonants, rules of syllabification for such items are detailed here. The syllable segmentation of a phonetic string such as CVC₁C₂VC (e.g., /kapris/) is designed so that the median consonants C₁C₂ are part of the second syllable onset, insofar as C₁C₂, and C₂V are legal strings (maximum onset principle). In case of ambiguous segmentation of words with intervocalic consonant groups (ambisyllabicity), the segmentation generally takes place between two consonants (e.g., /dis-pās /, /gal-ri/, /kaf-te/). However, there are two exceptions for which the segmentation occurs before the two consonants. The first exception is when a liquid (*l*, *r*) comes after a plosive (e.g., /a-bri/, /kɔ̃-plɛ/), although this rule is not applied in case of an alveolar plosive (*t*, *d*) followed by a liquid (e.g., /at-las/). The second exception is when a labiodental fricative (*f*, *v*) is followed by a liquid (e.g., /kɔ̃-flɪ/, /a-vril/).

We computed the orthographic syllabification of words from the phonological syllabification field. Programs were written in dBase III code under the Visual FoxPro software to transcribe each phonological syllable into the corresponding orthographic syllable as a function of the word letters. For example, if the first phonological syllable was /pyr/, the first three letters were respectively *p*, *u*, *r*, and the fourth letter was different from *e*, then the first orthographic syllable computed was *pur* (e.g., *pur* in the word *purger* phonologically syllabified /pyr-ʒe/). If the fourth letter was *e* and the other conditions were fulfilled, then the first orthographic syllable computed was *pure* (e.g., *pure* in *purement* phonologically syllabified /pyr-mā/). The orthographic transcriptions of the initial, middle and final syllables were processed separately to avoid errors in the syllable correspondences. To check for transcription errors or omissions, the concatenation of all orthographic syllables of each word was compared to the field of the orthographic word form. If the two data compared were not similar, we considered that the lexical entry needed to be revised. These entries were then manually corrected to warrant correct syllable segmentations.

InfoSyll Variables

The syllable measures were computed on the basis of the phonological and orthographic syllabifications of the lexical corpus described above. The different phonological and orthographic syllables occurring in the French language were extracted from the words and organized, hence making up the InfoSyll syllabary. Type frequencies were computed by counting the number of words among the 39,700 words sharing a given syllable at a given position. Token frequencies were computed by cumulating the printed frequency of all words sharing a given syllable at a given position (in occurrences per million). The variable of word frequency used was *freqliv*, which represents the frequency of occurrences per million of words in the book corpus of Lexique (New et al. 2004). Frequencies for the first position were computed with all the words ($n = 39,700$). Frequency computations for the second and third positions were run respectively with words containing two syllables or more ($n = 39,179$) and with words containing three syllables or more ($n = 22,593$). Monosyllabic words were excluded from the last-position computations, given that they were

included for the first-position computations. All measures were computed using a program in dBase III code run under the Visual FoxPro software. The phonetic symbols in InfoSyll are similar to those used in Lexique (New et al. 2004) and are presented in Appendix 1.

The InfoSyll syllabary contains 9,729 entries which correspond to the different combinations between the orthographic and phonological features of the syllables in the French language. There are 52 fields for each entry. These fields are indexed in Appendix 2. Before the presentation of the syllabic variables per se, numbering information is provided across six fields to clarify the organization of the syllabary. The first three fields are used for finding orthographic syllables from phonological syllables, that is, from phonology to orthography (*PO_numb_ps*, *PO_numb_os*, *PO_rank_os* fields). The field *PO_numb_ps* indicates the numbering of the phonological syllables (from 1 to 4,591). The field *PO_numb_os* indicates the numbering of the orthographic syllables (from 1 to 9,729), each orthographic syllable being connected with one of 4,591 phonological syllables. Given that each phonological syllable is declined as many times as there are different orthographic syllables corresponding to this phonological syllable, the field *PO_rank_os* indicates the rank of the different orthographic syllables for each phonological syllable (from 1 to 26). This means for example that there are 26 ways to spell the phonological syllable /se/ (i.e., 26 corresponding orthographic syllables). Following a similar reasoning, the fourth follow-up fields are used to find phonological syllables from orthographic syllables, that is, from orthography to phonology (*OP_numb_os*, *OP_numb_ps*, *OP_rank_ps*). The field *OP_numb_os* indicates the numbering of the orthographic syllables (from 1 to 8,619). The field *OP_numb_ps* indicates the numbering of the phonological syllables (from 1 to 9,729), each phonological syllable being connected with one of 8,619 orthographic syllables. Given that each orthographical syllable is declined as many times as there are different phonological syllables corresponding to this orthographic syllable, the field *OP_rank_ps* indicates the rank of the different phonological syllables for each orthographic syllable (from 1 to 11).² These six fields make it possible to use the syllabary flexibly when searching either the orthographic characteristics of phonological syllables (sort on the *PO_numb_os* field) or the phonological characteristics of orthographic syllables (sort on the *OP_numb_ps* field).

The first class of variable described in InfoSyll concerns the characteristics of the orthographic syllables. The first indication for each entry is the orthographic form of the given syllable (*os* field), with its abstract phonological structure (*os_str*) coded in consonants and vowels (respectively C and V) and its number of letters (*os_nlet*). Then, the type and token frequency computations for each orthographic syllable are provided for the first (respectively *osp1_tyf* and *osp1_tof*), second (*osp2_tyf* and *osp2_tof*), third (*osp3_tyf* and *osp3_tof*) and last (*ospl_tyf* and *ospl_tof*) positions in words. The field *Oambi* indicates when a syllable ensues from a segmentation of an orthographically ambisyllabic word. These cases of ambisyllabicity, specific to written words, occur when words exhibit intern semi-vowels or mute letters. The word /crejɔ̃/ (*crayon*) for example is syllabified without ambiguity when it is pronounced (/cre.jɔ̃/), while in the written modality the syllabification can be either *cra.yon* or *cray.on*, the ambiguity coming from the fact that the letter y participates in the two

² The real number of distinct phonological and orthographic syllables is respectively 4,591 and 8,619 (see fields *OP_numb_os* and *PO_numb_ps*). However, a phonological syllable can be repeated several times in the field *sp* given that it is connected to different orthographic syllables. For example, the phonological syllable /dal/ is repeated on three rows, one involving the couple /dall-dal/, one involving /dall-dale/, and one involving /dall-dalle/. This explains why the numbering for phonological syllables (1–9,729, field *PO_numb_os*) exceeds the real number of phonological syllables (4,591). The same is true when considering the field of orthographic syllables; the numbering for orthographic syllables (1–9,729, field *OP_numb_ps*) exceeds the real number of orthographic number (8,619).

phonological syllables. Similarly, the word /koy/ (*cohue*) is syllabified phonologically /ko.y/ but its orthographic syllabification can be either *coh.ue* or *co.hue*, the ambiguity stemming from the fact that the letter *h* does not participate in any of the two phonological syllables. In the present database, syllabification of such words was performed so that the ambiguous letter systematically belongs to the first syllable. Given that the question of the segmentation of ambisyllabic written words is beyond the scope of the present article and requires future empirical work, we indicated which syllables can be involved in ambisyllabicity (2.11% of syllables). Finally, the variable *nps_tot* indicates the total number of phonological syllables of the orthographic syllable (range from 1 to 11), and the variables *nps_p1*, *nps_p2*, *nps_p3*, and *nps_pl* indicate respectively the number of phonological syllables in the first, second, third and last positions. An example of the characteristics of the different phonological syllables of the orthographic syllable *que* is presented in Table 1. There are 4 phonological syllables corresponding to the orthographic syllable *que* (/kə/, /ka/, /ke/ et /kɛ/) but only one of these phonological syllables appears in the first position of words (e.g., /kə/ in *querelle* for example). More precisely, there are seven words beginning by this syllable (*tyf*=7) and their summed lexical frequency is 2.1 (*tof*=2.1). Similar information is provided for the 8,619 orthographical syllables contained in the syllabary.

The second class of variable in the syllabary concerns the phonological characteristics of the syllables. For each orthographic syllable, the phonological form of the corresponding phonological syllable is presented (*ps* field), with its abstract phonological structure (*ps_str*) coded in consonants, vowels, and semi-vowels (respectively C, V, and Y), its phoneme number (*ps_npho*) and its type and token frequencies computed for the first (respectively *psp1_tyf* and *psp1_tof*), second (*psp2_tyf* and *psp2_tof*), third (*psp3_tyf* and *psp3_tof*) and last positions in words (*pspl_tyf* and *pspl_tof*). The field *mono* indicates that a monosyllabic word constitutes a phonological syllable on its own and that this phonological syllable does not occur in any other word (referred to as monosyllables). This was done to distinguish between low-frequency syllables due to low occurrence in polysyllabic words (e.g., /ny/ is a low-frequency syllable compared to /ā/, respectively 151.24 vs. 2,189.07 token occurrences in the first position) and low-frequency syllables due to a single occurrence in a monosyllabic word (e.g., /ʒãr/, with 155.20 token occurrences, is a low-frequency syllable which is entirely defined by the monosyllabic word *genre*). The variable *nos_tot* indicates the total number of orthographic syllables of the phonological syllable (range from 1 to 26), and the variables *nos_p1*, *nos_p2*, *nos_p3*, and *nos_pl* indicate respectively the number of orthographic syllables in the first, second, third and last positions. For example, there are 11 orthographic syllables corresponding to the phonological syllable /fi/ (*phi*, *fi*, *phie*, *fix*, *fit*, *ffi*, *phy*, *fis*, *fie*, *fhy*, and *fee*) but only four of these orthographic syllables appear in the first position of words (e.g., *phy* in *physique* for example, *fi* in *ficelle*, *phi* in *philatélie*, and *fee* in *feeling*). An example of the characteristic of the different orthographic syllables of the phonological syllable /fi/ is presented in Table 2. Similar information is provided for the 4,591 phonological syllables contained in the syllabary.

The third class of variables concerns the frequencies of correspondences between orthographic and phonological syllables. Type and token frequency computations for each syllable are provided for the first (respectively *cor_osp1_tyf* and *cor_osp1_tof*), second (*cor_osp2_tyf* and *cor_osp2_tof*), third (*cor_osp3_tyf* and *cor_osp3_tof*) and last (*cor_osp1_tyf* and *cor_osp1_tof*) positions in words. An example of a word (extracted from Lexique, New et al., 2004) that contains the orthographic syllable at the given position is also provided (e.g., fields *cor_osp1_ex*, *cor_osp2_ex*, *cor_osp3_ex*, and *cor_osp1_ex*). When searching for a given syllable, the fields *os* and *ps* can be used jointly, given that the entries in the database (*os* field) are organized so that the different orthographic syllables of a phonological syllable

Table 1 Summary of variables and computations provided by InfoSyll for the orthographic syllable *que*

	Nb. units	Abstract structure	Syllable position in words											
			First			Second			Third			Last		
			Tyf	Tof	Ex.	Tyf	Tof	Ex.	Tyf	Tof	Ex.	Tyf	Tof	Ex.
OS	3	CVV	9	3,325,94		22	207,60		10	20,63				
PS	/kɛ/	CV	46	185,23	coquetterie									
	/ke/	CV	104	229,72	quiquennial			99	281,02	déliquescent				
	/ka/	CV	341	571,81	fréquemment			162	263,44	conséquentment				
	/kə/	CV	7	2,1	querelle			46	185,23	mousquetaire	12	16,49	embarquement	

Notes: Nb. units, number of units (phonemes for phonological syllables, letters for orthographic syllables); Tyf, type frequency; Tof, token frequency; Ex, example of word containing the given syllable at the corresponding position; PS, phonological syllable; C, consonant; V, vowel; OS, orthographic syllable; an empty cell means that the syllable does not appear in words in a given position

Table 2 Summary of variables and computations provided by InfoSyll for the phonological syllable /fi/

	Nb. units	Abstract structure	Syllable position in words											
			First			Second			Third			Last		
			Tyf	Tof	Ex.	Tyf	Tof	Ex.	Tyf	Tof	Ex.	Tyf	Tof	Ex.
PS	/fi/	2 CV	167	580.42		163	513.7		132	120.88		69	130.74	
OS	fee	3 CVV	1	0.54	feeling									
	fhy	3 CCV				1	0.07	sulphydrique						
	fie	3 CVV							1	0.2	crucifiement			
	fis	3 CVC							1	1.01	salsifis	1	1.01	salsifis
	phy	3 CCV	19	43.8	physique	3	0.62	emphysème	14	7.25	géophysique			
	ffi	3 CCV				57	242.25	affiner	7	8.25	insuffisant	1	2.03	bouffi
	fit	3 CVC				3	23.58	confit	1	0.54	déconfit	4	24.12	confit
	fix	3 CVC				1	0	préfix	1	7.43	crucifix	2	7.43	crucifix
	phie	4 CCVV				2	0.41	graphie	2	4.46	atrophie	57	72.51	philosophie
	fi	2 CV	117	475.71	ficelle	78	240.6	définitif	98	90.58	indéfini	3	23.03	défi
	phi	3 CCV	30	60.37	phialéite	18	6.17	amphibie	7	1.16	hémophilie	1	0.61	amphi

Notes: Nb. units, number of units (phonemes for phonological syllables, letters for orthographic syllables); Tyf, type frequency; Tof, token frequency; Ex, example of word containing the given syllable at the corresponding position; PS, phonological syllable; C, consonant; V, vowel; OS, orthographic syllable; an empty cell means that the syllable does not appear in words in a given position

follow each other, phonological syllables being alphabetically ordered. Thus, it is easy to find all the orthographic syllables of a given phonological syllable and to select the more or less frequent syllable (and similarly when searching all the phonological syllables of a given orthographic syllable). Considering the syllable /dã/ for example, when the database is used in the PO form (sort on *PO_num_b_os*), the syllabary indicates that its token frequency in the first position is 8,449.9, and that five orthographic syllables are associated more or less frequently (*dam*: 1.69, *den*: 38.8, *dans*: 8299.1, *dan*: 96.16, and *dent*: 11.15). Conversely, considering the orthographic syllable *dan* when the database is used in the OP form (sort on *OP_num_b_ps*), the syllabary indicates that its token frequency in the first position is 96.23 and that two phonological syllables are associated (/dã/: 96.16 and /dan/: 0.07). Throughout the syllabary, an empty cell means that the syllable does not appear in words in a given position.

Computations on InfoSyll

Besides facilitating the construction of syllabic materials for psycholinguistic studies and allowing precise control for syllable variables even in experiments not dealing directly with the syllable issue, one of the advantages of the InfoSyll syllabary is that a large number of syllabic variables are gathered, so it is possible to establish some syllabic characteristics of the French language. In this perspective, we present general data on the syllabic redundancy in French as evidenced by InfoSyll.

An analysis on the whole syllable corpus provided by InfoSyll showed that there is a wide variety of syllabic structures in French. Concerning orthographic syllables, there are 148 different types of structures, with an average length of 4.32 letters (ranging from 1 to 10, SD=1.25), while there are 45 different types of phonological syllable structures, with an average length of 3.47 phonemes (ranging from 1 to 6, SD=0.77). When examining the ten most frequent syllabic structures, it can be observed that the CV and CVC syllables are by far the two most frequent syllabic structures, CV syllables being more frequent than CVC syllables, all positions in words considered (see Fig. 1). This result is obtained for both orthographic syllables (26.38% of CV syllables and 15.42% of CVC syllables) and phonological syllables (48.66% of CV syllables and 21.04% of CVC syllables). The fact that the CV syllable is the most frequent type of syllabic structure in written words is consistent with the findings of [Wioland \(1985\)](#) reported for a corpus of spoken words. According to [Meynadier \(2001\)](#), the predominance of CV syllables over the others could be due to the universal and

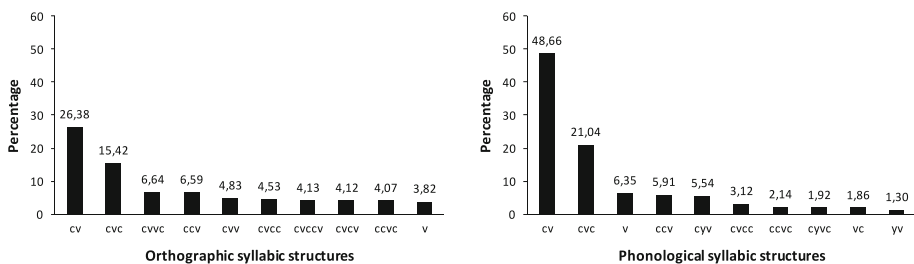


Fig. 1 Percentage of orthographic syllabic structures (*left*) and phonological syllabic structure (*right*) weighted as a function of their frequency of occurrences in written words, all positions confounded. Only the ten most frequent structures are presented (C consonant, V vowel, Y semi-vowel)

Table 3 Frequency and structure of the ten most frequent orthographic syllables (top) and the ten most frequent phonological syllables (bottom) according to their position in words (first, second, third, last, all positions considered)

Position		Second			Third			Last			All		
First	Fq	Syllable	struc.	Fq	Syllable	struc.	Fq	Syllable	struc.	Fq	Syllable	struc.	Fq
<i>Orthographic syllables</i>													
a	1840	ti	CV	731	ment	CVCC	848	ment	CVCC	2017	ment	CVCC	3065
dé	1452	ta	CV	381	ter	CVC	370	tion	CVVC	1310	a	V	2075
é	1093	to	CV	378	ti	CV	318	té	CV	961	dé	CV	1822
in	981	ter	CVC	317	li	CV	303	ter	CVC	637	té	CV	1633
re	699	ni	CV	309	té	CV	295	teur	CVVC	575	tion	CVVC	1586
i	593	ra	CV	306	tion	CVVC	247	ser	CVC	398	ter	CVC	1343
ca	591	di	CV	306	ta	CV	239	tique	CVCVV	357	é	V	1307
en	575	mi	CV	302	ra	CV	238	ler	CVC	264	ti	CV	1192
ré	564	ma	CV	294	ri	CV	204	rie	CVV	263	ma	CV	1009
con	562	co	CV	279	tique	CVCVV	199	ner	CVC	260	ca	CV	998
<i>Phonological syllables</i>													
/a/	1993	/tʰ/	CV	800	/mã/	CV	985	/mã/	CV	2147	/mã/	CV	3572
/de/	1491	/te/	CV	529	/te/	CV	723	/te/	CV	1640	/te/	CV	3043
/e/	1350	/to/	CV	494	/ne/	CV	444	/sjõ/	CVV	1423	/de/	CV	2259
/ê/	1229	/si/	CV	458	/ri/	CV	389	/ne/	CV	701	/a/	V	2187
/ã/	1148	/ri/	CV	432	/ti/	CV	365	/teer/	CVC	618	/sjõ/	CVV	1789
/i/	772	/ta/	CV	403	/li/	CV	358	/ze/	CV	511	/re/	CV	1603
/kõ/	747	/se/	CV	385	/se/	CV	326	/se/	CV	479	/ne/	CV	1552
/ka/	691	/li/	CV	385	/sjõ/	CVV	321	/le/	CV	470	/e/	V	1548
/re/	683	/le/	CV	367	/le/	CV	300	/re/	CV	463	/se/	CV	1448
/o/	640	/ni/	CV	360	/re/	CV	295	/je/	CV	426	/ti/	CV	1407

Notes: Struc., structure; Fq, type frequency; C, consonant; V, vowel; Y, semi-vowel

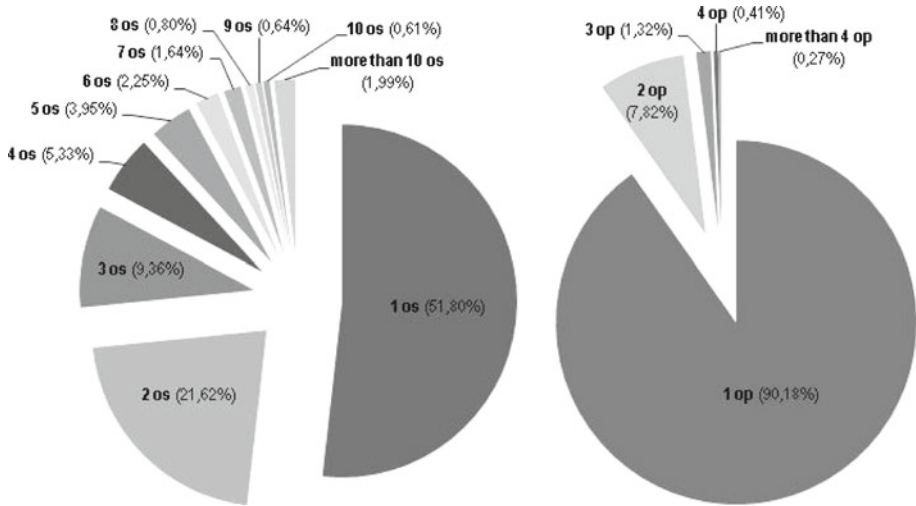


Fig. 2 Percentage of phonological syllables distributed according to the number of corresponding orthographic syllables (*OS*) (*left panel*) and percentage of orthographic syllables distributed according to the number of corresponding phonological syllables (*PS*) (*right panel*), all positions considered

canonical feature of the CV syllable (i.e., syllable encountered in all languages) and to its stable articulatory structure.

When looking at the ten most frequent syllables of the language (see Table 3), the data show that the syllable /mã/ spelled *ment* is by far the most frequent syllable in French all positions considered, which is due to its very high frequency and predominance in the third and last positions in words. The most frequent syllable in the first position in words is the syllable /a/ spelled *a*, and the most frequent syllable in second position is the syllable /ti/ spelled *ti*. In addition, the most frequent syllables are mostly CV syllables such as *dé*, *ca*, *re*, *ti*, *to*, *ta*, for orthographic syllables and /mã/, /kõ/, /de/, /si/, /ti/, /te/ for phonological syllables. This occurs for syllables in the second and third position in words in particular, and the ten most frequent syllables are almost exclusively CV syllables in the second position.

An advantage of InfoSyll consists in providing both the different spellings of each phonological syllable and the different pronunciations of each orthographic syllable. When considering these new variables, computations showed that about 60.48% of phonological syllables can be spelled in only one way (feedback consistency). The rate drops to 51.80% when monosyllables are excluded, the 48.20% remaining syllables being spelled more than one way, all positions considered. Conversely, 88.69% of orthographic syllables can be pronounced in only one way (monosyllables excluded), the 11.31% remaining syllables being pronounced more than one way (feedforward consistency) (see Fig. 2). In addition, data showed that the phonological syllables have on average 2.36 different orthographic transcriptions (ranging from 1 to 28, SD=2.48). Correlation tests indicated that this number of orthographic syllables is positively correlated to type frequency (all positions considered) of the phonological syllable ($r = .61, p < .001$). This means that the more frequent a syllable, the greater is its number of different spellings.

These data support the view that the French language is more feedforward-consistent than feedback-consistent also at a syllabic level. To compare the measures of consistency computed from InfoSyll on syllables with those previously reported on monosyllabic words

(Peereман and Content 1999a; Ziegler et al. 1996), we computed consistency measures at a word level for the first three syllabic positions in words of the lexical corpus used. Data showed that 69.96, 61.47, and 78.09% of the words had respectively their first, second or third orthographic syllable that can be pronounced only one way (feedback consistency), while only 8.72, 8.96, and 4.26% of the same words had respectively their first, second or third phonological syllable that can be written only one way (feedback consistency). Consistency measures are here lower than those previously reported at a rime or grapheme level. This can be explained by the fact that syllables can be composed of several inconsistent intra-syllabic units (e.g., /s/ and /ã/ in the syllable /sã/ are both inconsistent), which increases the spelling possibilities of a given syllable, thus decreasing the rate of feedback consistency. In previous works, evidence has been provided that feedback consistency can influence latencies of word processing in tasks such as spoken lexical decision, written lexical decision and naming, both in French and English (e.g., Stone et al. 1997; Ziegler et al. 2008). Given that, on the one hand, syllables are predominant units of word processing in French compared to infra-syllabic units, and that on the other hand, the present analysis shows that syllable transcriptions are relatively inconsistent in French, especially when the syllables are of high frequency, the number of orthographic syllables of a given phonological syllable could be a potentially important variable in visual word processing. The issue of syllabic consistency therefore requires further studies to assess whether and to what extent it influences visual word recognition.

Conclusion

The growing number of studies on the role of syllables in reading attests to the importance of this issue in the modeling of the processes underlying the visual recognition of polysyllabic words. Current questions on the syllable issue, such as the influence of the orthographic information contained by words and syllables in syllabic effects, require subtle manipulations of syllabic variables. We propose here a new tool, which is a syllabary of the French language gathering numerous measures on phonological and orthographic syllables encountered in texts, such as their positional type and token frequencies, structure, length, and correspondences between phonological and orthographic syllables. These various measures are essential for assessing the role of syllable units in visual word processing in line with current concerns on the syllable issue. The InfoSyll syllabary should help in addressing new questions about the role of syllables in activities such as reading and writing, since it makes it possible to select syllables to be manipulated according to various criteria. The database can also be used for issues not directly related to syllables, given that it allows syllable variables known to affect written word processing to be controlled precisely. In addition, the syllabic variables available can contribute to making quantitative descriptions of the syllabic properties of the French language. The InfoSyll syllabary should prove to be a useful tool for psycholinguistic research related to visual word processing.

Availability

The InfoSyll database can be freely downloaded from the following address: <http://lclld.ulb.ac.be/lequipe/fabienne-chetail> in three formats (text, Excel, dBase) so that it can be used with word processing, spreadsheet, or database software.

Appendix 1

see Table 4.

Table 4 Characters coding for the phonetic symbols in InfoSyll

Vowels and semi-vowels			Consonants			
Phonetic symbol	Example	InfoSyll code	Phonetic symbol	Category	Example	InfoSyll code
/a/ - /ɑ/	appel	a	/p/	bilabial plosive	plume	p
/ə/ (schwa)	fenêtre	*	/b/	bilabial plosive	balle	b
/œ/	beurre	9	/t/	alveolar plosive	terre	t
/ø/	deux	2	/d/	alveolar plosive	deux	d
/e/	café	e	/k/	volar plosive	casser	k
/ɛ/	neige	E	/g/	volar plosive	gare	g
/i/	pile	i	/f/	labiodental fricative	fer	f
/o/ - /ɔ/	chaud	O	/v/	labiodental fricative	vent	v
/y/	mur	y	/s/	alveolar fricative	sol	s
/u/	doux	u	/z/	alveolar fricative	oser	z
/ã/	lent	@	/ʃ/	post-alveolar fricative	chien	S
/õ/	bon	§	/ʒ/	post-alveolar fricative	page	Z
/ẽ/	cinq	5	/x/	velar fricative (jota)	jerez	x
/œ̃/	parfum	1	/m/	bilabial nasal	mer	m
/ɥ ^a /	huile	8	/n/	alveolar nasal	nuit	n
/w ^a /	bois	w	/ɲ/	palatal nasal	vigne	N
/j ^a /	paille	j	/ŋ/	velar nasal	camping	G
			/r/	liquid	verre	R
			/l/	liquid	lune	l

Note: ^a Semi-vowels

Appendix 2: Name and Description of InfoSyll Fields

Numbering

Use of the database from phonological to orthographic syllables:

PO_numb_ps: numbering of the phonological syllables (from 1 to 4,591)

PO_numb_os: numbering of the orthographic syllables (from 1 to 9,729)

PO_rank_os: rank of the different orthographic syllables for each phonological syllable (from 1 to 26)

Use of the database from orthographic to phonological syllables:

OP_numb_os: numbering of the orthographic syllables (from 1 to 8,619)

OP_numb_ps: numbering of the phonological syllables (from 1 to 9,729)

OP_rank_ps: rank of the different phonological syllables for each orthographic syllable (from 1 to 11)

Information on Orthographic Syllables

os: orthographic syllable forms corresponding to the phonological syllable

os_str: abstract orthographic syllable form of each orthographic syllable (C=consonant, V=vowel)

os_nlet: letter number of the orthographic syllable

osp1_tyf, *osp2_tyf*, *osp3_tyf*, and *ospl_tyf*: positional type frequency of the orthographic syllable, respectively at the first, second, third and last positions in words

osp1_tof, *osp2_tof*, *osp3_tof*, and *ospl_tof*: positional token frequency of the orthographic syllable, respectively at the first, second, third and last positions in words

Oambi: if equals 1, means that the orthographic syllable ensue from a segmentation of an orthographically ambisyllabic word

nps_tot: total number of phonological syllables corresponding to the orthographic syllable (from 1 to 11)

nps_p1, *nps_p2*, *nps_p3*, and *nps_pl*: number of phonological syllables respectively at the first, second, third and last positions in words

Information on Phonological Syllables

ps: phonological syllable form

ps_str: abstract phonological syllable form (C=consonant, V=vowel, Y=semi-vowel)

ps_npho: phoneme number of the phonological syllable

psp1_tyf, *psp2_tyf*, *psp3_tyf*, and *pspl_tyf*: positional type frequency of the phonological syllable, respectively at the first, second, third and last positions in words

psp1_tof, *psp2_tof*, *psp3_tof*, and *pspl_tof*: positional token frequency of the phonological syllable, respectively at the first, second, third and last positions in words

mono: if equals 1, means that a monosyllabic word constitutes the syllable on its own and that the syllable does not occur in any other word

nos_tot: total number of orthographic syllables corresponding to the phonological syllable (from 1 to 26)

nos_p1, *nos_p2*, *nos_p3*, and *nos_pl*: number of orthographic syllables respectively at the first, second, third and last positions in words

Information on Correspondences Between Phonological and Orthographic Syllables

cor_osp1_tyf, *cor_osp2_tyf*, *cor_osp3_tyf*, and *cor_ospl_tyf*: positional type frequency of correspondence between a phonological and an orthographic syllable, respectively at the first, second, third and last positions in words

cor_osp1_tof, *cor_osp2_tof*, *cor_osp3_tof*, and *cor_ospl_tof*: positional token frequency of correspondence between a phonological and an orthographic syllable, respectively at the first, second, third and last positions in words

cor_osp1_ex, *cor_osp2_ex*, *cor_osp3_ex*, and *cor_ospl_ex*: example of word containing the given syllable, respectively at the first, second, third and last positions in words

References

- Alvarez, C. J., Carreiras, M., & de Vega, M. (2000). Syllable-frequency effect in visual word recognition: Evidence of sequential-type processing. *Psicológica*, 21, 341–374.

- Alvarez, C. J., Carreiras, M., & Perea, M. (2004). Are syllables phonological units in visual word recognition?. *Language and Cognitive Processes*, *19*, 427–452.
- Alvarez, C. J., Carreiras, M., & Taft, M. (2001). Syllables and morphemes: Contrasting frequency effects in Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 545–555.
- Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, *105*, 678–723.
- Brand, M., Rey, A., & Peereman, R. (2003). Where is the syllable priming effect in visual word recognition?. *Journal of Memory and Language*, *48*, 435–443.
- Carreiras, M., Alvarez, C. J., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, *32*, 766–780.
- Carreiras, M., & Perea, M. (2002). Masked priming effects with syllabic neighbors in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1228–1242.
- Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain & Language*, *90*, 393–400.
- Chetail, F., & Mathey, S. (2009a). Syllabic priming in lexical decision and naming tasks: The syllable congruency effect re-examined in French. *Canadian Journal of Experimental Psychology*, *63*, 40–48.
- Chetail, F., & Mathey, S. (2009b). The syllable frequency effect in visual recognition of French words: A study in skilled and beginning readers. *Reading and Writing*, *22*, 955–973.
- Conrad, M., Carreiras, M., & Jacobs, A. M. (2008). Contrasting effects of token and type syllable frequency in lexical decision. *Language and Cognitive Processes*, *23*, 296–326.
- Conrad, M., Carreiras, M., Tamm, S., & Jacobs, A. M. (2009). Syllables and bigrams: Orthographic redundancy and syllabic units affect visual word recognition at different processing levels. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 461–479.
- Conrad, M., Grainger, J., & Jacobs, A. M. (2007). Phonology as the source of syllable frequency effects in visual word recognition: Evidence from French. *Memory & Cognition*, *35*, 974–983.
- Conrad, M., & Jacobs, A. M. (2004). Replicating syllable frequency effects in Spanish in German: One more challenge to computational models of visual word recognition. *Language and Cognitive Processes*, *19*, 369–390.
- Conrad, M., Stenneken, P., & Jacobs, A. M. (2006). Associated or dissociated effects of syllable frequency in lexical decision and naming. *Psychonomic Bulletin & Review*, *13*, 339–345.
- Content, A., Mousty, P., & Radeau, M. (1990). BRULEX: Une base de données lexicales informatisée pour le français écrit et parlé. *L'année Psychologique*, *90*, 551–566.
- Davis, C., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighbourhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*, *37*, 665–671.
- Doignon, N., & Zagar, D. (2005). Illusory conjunctions in French: The nature of sublexical units in visual word recognition. *Language and Cognitive Processes*, *20*, 443–464.
- Doignon-Camus, N., Zagar, D., & Mathey, S. (2009). Can we see syllables in monosyllabic words? A study with illusory conjunctions. *The European Journal of Cognitive Psychology*, *21*, 599–614.
- Dufour, S., Peereman, R., Pallier, C., & Radeau, M. (2002). VoCoLex: Une base de données lexicales sur les similarités phonologiques entre les mots français. *L'année Psychologique*, *102*, 725–746.
- Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: The role of syllabic units. *Journal of Memory and Language*, *35*, 708–723.
- Ferrand, L., Segui, J., & Humphreys, G. W. (1997). The syllable's role in word naming. *Memory & Cognition*, *25*, 458–470.
- Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary?. *Cognition*, *50*, 239–269.
- Lima, S. D., & Pollatsek, A. (1983). Lexical access via an orthographic code? The Basic Orthographic Syllabic Structure (BOSS) reconsidered. *Journal of Verbal Learning and Verbal Behavior*, *22*, 310–332.
- Macizo, P., & Van Petten, C. (2007). Syllable frequency in lexical decision and naming of English words. *Reading and Writing*, *20*, 295–331.
- Mathey, S., & Zagar, D. (2002). Lexical similarity in visual word recognition: The effect of syllabic neighbourhood in French. *Current Psychology Letters*, *8*, 107–121.
- Mathey, S., Zagar, D., Doignon, N., & Seigneuric, A. (2006). The nature of syllabic neighbourhood in French. *Acta Psychologica*, *123*, 372–393.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.
- Mehler, J., Dommergues, J., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, *20*, 298–305.

- Meynadier, Y. (2001). La syllabe phonétique et phonologique: Une introduction. *Travaux Interdisciplinaires du Laboratoire Parole Et Langage*, 20, 91–148.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36, 516–524.
- Pallier, C. (1994). *Rôle de la syllabe dans la perception de la parole: Etudes attentionnelles*. Ph.D dissertation. Ecole des Hautes Etudes en Sciences Sociales.
- Peereman, R., & Content, A. (1999a). LEXOP: A lexical database providing orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, 31, 376–379.
- Peereman R., & Content, A. (1999b). *LEXOP (ver. 2). User's manual*. (<http://leadserv.u-bourgogne.fr/bases/lexop/LexopMan.pdf>).
- Peereman, R., Lété, B., & Sprenger-Charolles, L. (2007). Manulex-infra: Distributional characteristics of grapheme-phoneme mappings, and infralexical and lexical units in child-directed written material. *Behavior Research Methods*, 39, 593–603.
- Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighbourhood frequency in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 134–144.
- Pulgram, E. (1970). *Syllable, word, nexus, cursus*. The Hague: Mouton.
- Rouibah, A., & Taft, M. (2001). The role of syllabic structure in French word recognition. *Memory & Cognition*, 29, 373–381.
- Schiller, N. O. (1998). The effect of visually masked syllable primes on the naming latencies of words and pictures. *Journal of Memory and Language*, 39, 484–507.
- Schiller, N. O. (2000). Single word production in English: The role of subsyllabic units during phonological encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 512–528.
- Schiller, N. O. (2006). Phonology in the production of words. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (pp. 545–553). Amsterdam: Elsevier.
- Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy?. In M. Coltheart (Ed.), *Attention and Performance, XII: The Psychology of Reading* (pp. 245–263). Hillsdale: Lawrence Erlbaum.
- Stenneken, P., Conrad, M., & Jacobs, A. M. (2007). Processing of syllables in production and recognition tasks. *Journal of Psycholinguistic Research*, 78, 36–65.
- Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Cognition*, 36, 337–359.
- Taft, M. (1979). Lexical access via an orthographic code: The basic orthographic syllabic structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*, 18, 21–39.
- Taft, M. (2001). Processing of orthographic structure by adults of different reading ability. *Language and Speech*, 44, 351–376.
- Tousman, S., & Inhoff, A. (1992). Phonology in multisyllabic word recognition. *Journal of Psycholinguistic Research*, 21, 525–544.
- 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, 107–136.
- Wioland, F. (1985). *Les structures syllabiques du français*. Genève: Editions Slatkine.
- Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers*, 28, 504–515.
- Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory recognition: Where do we stand after more than a decade?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 643–661.