Cross-linguistic investigations of oral and silent reading

Christophe Coupé, Yoon Mi Oh, François Pellegrino, Egidio Marsico

Laboratoire Dynamique du Langage UMR5596, Université de Lyon and CNRS, France

{Christophe.Coupe; Egidio.Marsico}@cnrs.fr
{yoon-mi.oh; francois.pellegrino}@univ-lyon2.fr

Abstract

Recent research on speech rate (Pellegrino et al., 2011) has shown that languages differ in terms of syllable rate, and that these differences are compensated by the average amount of information carried by syllables. The more syllables a language needs to express a given amount of information, the higher its syllable rate tends to be.

These results were obtained with subjects reading texts on a computer screen. The question arose whether silent reading rates would correlate with oral reading rates across languages. Although silent and oral reading fluency has been studied with respect to reading comprehension and how it develops in children, little literature focuses on comparing them in different languages.

We present here data for 8 languages (Cantonese, Finnish, French, Japanese, Korean, Mandarin, Serbian, and Thai). For each language, silent and oral syllables rates as well as reading durations were measured for several subjects and 15 different texts. Various comparisons were performed, and mixed-effects regression models were used to further evaluate the weight of the different variables (gender, language, speaker and text).

Most significantly, oral and silent reading rates appear to be well correlated, suggesting that language-specific syllabic complexity impacts silent reading in a similar way to oral reading.

Index Terms: Cantonese, Cross-language Study, Finnish, French, Japanese, Korean, Mandarin, oral reading, reading rates, Serbian, silent reading, Thai

10.21437/Interspeech.2014-127

1. Theoretical background

Silent and oral reading rates have often been utilized as indices during the last decades, from different perspectives ranging from the assessment of reading competence, the development of theoretical psycholinguistic models or cross-linguistic comparisons. A brief overview of these studies is provided in this section, as well as rationales for the present study.

1.1. Reading rate and the general study of reading

Oral reading rate has been investigated in relationship with the memory span of speakers and the number of constituents of words, and therefore related to both linguistic features and general cognitive processes [1]. Both oral and silent reading rates have also been studied with respect to reading development in children. While both oral and silent reading have been shown to become more fluent with age, and to gradually display more structures and stability, differences between them have also been highlighted, with silent reading showing greater flexibility [2]. Scholars focusing on reading instruction have been keen on developing efficient reading habits that maximize both fast reading rates and strong comprehension [3][4]. In this context, oral reading fluency has been established as a relevant indicator of overall reading competence (e.g. [5], [6]).

1.2. Reading rate and word predictability

Reading rates also provide an interesting window into cognitive language processing, beyond the assessment of general reading abilities. In self-paced reading experiments, each word's reading time depends on its expectation, which in turn varies with various syntactic, semantic, and pragmatic cues. The higher the expectation, the easier and faster the word is processed and integrated within the larger context of the sentence or text. Building on previous research, Smith and Levy explored the quantitative nature of the relationship between word expectation and reading time in English [7]. They estimated word predictability with n-grams derived from very large corpora and they conducted extensive experimentation to measure oral reading durations for different words in different contexts. They showed that, among the mathematical relations (linear, logarithmic, super logarithmic, etc.) predicted by theoretical models, the logarithmic relationship i) is the most likely and ii) is valid for a large span from very predictable to very unpredictable words. They thus provided strong arguments in favor of two models of reading (Norris' *optimal visual discrimination* [8] and their own *highly incremental processing* model).

1.3. Cross-linguistic studies of oral reading rate

Few studies have focused on cross-linguistic variations of reading rates, and as far as we know, none of them investigated silent reading rates with the exception of [9] in bilingual contexts. In a recent paper based on seven languages, Pellegrino et al. [10] investigated cross-linguistic variations in terms of oral syllabic rate, syllabic information density and information rate. Languages were shown to differ in terms of syllabic rate and syllabic information density, while being approximately similar in terms of information rate. The observed tradeoff between the first two variables suggests that different strategies are possible: some languages rely on more syllables produced at a faster rate (like Spanish and Japanese), while others favor slower syllabic rates and use fewer syllables (like Mandarin Chinese) in order to convey semantic information at around the same rate. Moreover, Syllabic complexity – defined by the average number of constituents (segments and possibly tones) of a syllable, weighted by the frequency of each syllable in a large corpus – was shown to correlate positively with information density and negatively with syllabic rate. A reasonable interpretation was therefore that more phonologically complex syllables pack more information, but take longer to produce.

1.4. Research goals

Investigating oral and silent rates cross-linguistically makes sense to better understand the processes underlying reading as a cognitive activity. For oral readings, as Naveh-Benjamin and

Ayres [1] and Pellegrino et al. [10] have suggested among others, syllabic complexity seems to have a direct impact on the production of speech; more complex and information-rich syllables take longer to articulate, and reduce the syllabic rate. As oral reading entails an overt articulation of constituents while silent reading doesn't, we could expect faster rates in the latter. However, both silent and oral readings involve a phonological cognitive processing sensitive to word structural complexity (e.g. [11] [12], [13] [14]). Consequently, silent reading rates should also reflect cross-linguistic differences, but with lower amplitude compared to oral reading. A first question is therefore how much silent rates vary from one language to another, and a second question whether results from [10] can be reproduced for oral rates, and expanded to silent rates.

While Smith and Levy [7] investigated silent reading duration with respect to word predictability, we thought relevant to analyze text reading duration on the basis of text length. Does the amount of time needed to read a text relate linearly or not to its length? For example, does reading speed decrease as text length increases because of higher cognitive load, or does it decrease as a result of more predictable words? This last option relies on the assumption that word-level effects stressed in [7] accumulate at the text level. If it holds, are such text-level effects present for silent reading, oral reading, or both?

The present study is a first approach to address the relationship between oral and silent reading rates from a cross-linguistic perspective. It is based on oral and silent reading rates computed over whole texts rather than on individual word reading times (as provided through self-paced reading or eye fixation durations) and it investigates the influence of several factors of variation (subject, language, etc.).

2. Data and Method

2.1. Written material

We adapted Pellegrino et al.'s approach to investigate a new set of languages. 15 English texts were selected among the 20 initial texts to minimize translation problems and the difficulty to deal with Western culture-specific situations. The initial texts in British English were taken from the Multext multilingual corpus [15] and were carefully translated into Cantonese, Finnish, French, Japanese, Korean, Mandarin, Serbian and Thai by native speakers. We could therefore study the reading of similar semantic content in different languages, and among different speakers of these languages.

2.2. Data collection

Data were collected for 8 languages (Cantonese Chinese YUE, Mandarin Chinese CMN, Finnish FIN, French FRA, Japanese JPN, Korean KOR, Serbian SRP, and Thai THA) by the first two authors in Lyon (France), Beijing (China), Hong Kong, Bangkok (Thailand), Seoul (South Korea) and Belgrade (Serbia). All subjects were native speakers of the target language, with additional efforts to focus on a specific variety of the language when possible – e.g. Mandarin spoken in Beijing, Serbian in Belgrade and Korean in Seoul. 5 men and 5 women were selected for this study for each language (totally 80 subjects for the 8 languages). No control was strictly enforced on age or on the social diversity of subjects, who were mainly students or members of academic or teaching institutions.

The RocMe! software [16] was used for presentation of experiment instructions and texts, as well as for recordings. During a first phase, subjects were asked to read the 15 texts silently one by one, from the beginning of the first sentence to the end of the last sentence. In a second phase, each text was then orally read twice by the subject, before being recorded. Reading a text three times – once silently, twice orally – before recording it orally allowed a subject to become familiar with its sentences, and to minimize reading errors. In case such an error occurred, the recording could be done again, which actually did not occur much. Texts were randomly ordered in each phase for each subject.

To measure reading duration, subjects were asked to press the space bar to display the text and start time measurement, and again to press the same key to stop the recording and move to the next text. Speech and reading durations were recorded for oral reading, while only duration was meaningful for silent reading.

2.3. Computing syllabic rate and information density

The *Praat* software [17] was used to discard pauses longer than 150ms in the oral recording.

We followed [10] to define syllabic rate and syllabic information density, adding a distinction between silent and oral (reading) rates (by reading rate, we therefore meant a syllabic reading rate). We left information rate aside for the sake of simplicity and brevity. We used Vietnamese as a reference language to normalize computations and avoid the problem of providing a quantitative measure of semantic content (see below). Using such an external anchor leads to erase relevantly the differences between the 15 texts in terms of amount of semantic information, under the assumption that twice the amount of information to convey requires twice as many syllables in Vietnamese, given its isolating nature. Data for Vietnamese came from [15].

For each text T_k in language L, composed of $\sigma_k(L)$ syllables, the average quantity of information per syllable is equal to the overall amount of semantic content S_k divided by the number of syllables:

$$I_L^k = \frac{S_k}{\sigma_k(L)} \tag{1}$$

S_k is independent of the language L, under the assumption that translating from English to other languages preserved the amount of semantic information. This amount is however difficult to estimate quantitatively, thus the use of Vietnamese ($L = VI$) as a reference to define information density:

$$ID_L^k = \frac{I_L^k}{I_{VI}^k} = \frac{S_k}{\sigma_k(L)} \times \frac{\sigma_k(VI)}{S_k} = \frac{\sigma_k(VI)}{\sigma_k(L)} \tag{2}$$

Denoting $SilD_L^{k,sp}$ and $OrD_L^{k,sp}$ the duration of respectively the silent and oral readings of text T_k by speaker sp in language L, we then defined silent and oral syllabic rates as follows:

$$SilSR_L^{k,sp} = \frac{\sigma_k(L)}{SilD_L^{k,sp}}, OrSR_L^{k,sp} = \frac{\sigma_k(L)}{OrD_L^{k,sp}} \tag{3}$$

2.4. Missing data and outliers

Statistical computations here and in the next sections were done with R (v. 3.0.3) and the statistical packages geoR [18], lme4 and lmerTest [19].

For the 1,200 (subject, text) pairs we analyzed, we found 3 missing values of silent durations, probably due to manipulation errors. Following a standard procedure, corresponding rates were replaced by the subject's average value, and durations derived from them.

Data were inspected for possible outliers departing from normal distributions. Average *SilSR*, *OrSR SilD* and *OrD* per subject were computed, after a boxcox transformation for *SilSR* [20] (the other variables were normally distributed). One Thai subject had *SilSR* and *SilD* values slightly below -2.5 times standard deviation. Besides, a few subjects had very high average silent reading rates (close to 20 syllables per second), which could however be explained by extensive reading habits. In absence of evident subject outlier, we kept the 80 subjects. We last investigated potential outliers at the global (subject, text) distribution, without averaging data. We applied boxcox transformations as above, detected outliers and deleted them. We removed a total of 39 (subject, text) pairs. We therefore conducted the analyses with 1,161 pairs.

3. Results

We applied the same procedure to address the questions raised in 1.4: we first looked at simple correlations using standard Pearson's R or Spearman's Rho coefficient, drew graphics easing the representation of data, then relied on various mixed-effects models to better take the various factors of our experiment into account.

3.1. Comparison of silent and oral syllabic rates

We first tried to address our hypothesis of an absence of correlation between silent and oral syllabic reading rates.

We computed coefficients of correlation between *SilSR* and *OrSR* for the 1,161 pairs (subject, text), and also between by-speaker and by-language average rates. Results are given in Table 1; Figure 1 illustrates the relationship at language level.

Data Set	Correlation coefficient	
All data (N = 1161)	Pearson's R:	.60***
Averaged by speaker (N = 80)	Pearson's R:	.67***
Averaged by language (N = 8)	Spearman's Rho:	.81**

Table 1. *Correlation between silent and oral SR.*

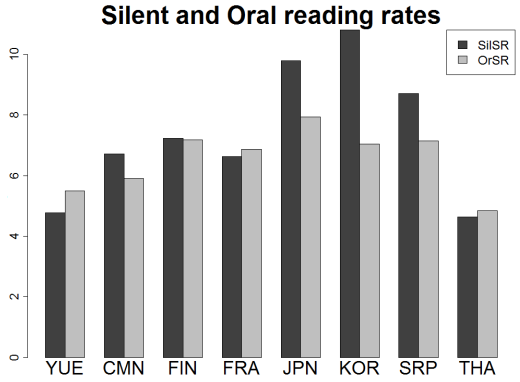

Figure 1: *Silent and oral reading rates (in #syl/s).*

The first result is that language-averaged *SilSRs* exhibit noticeable differences, even larger than *OrSRs*. Three languages show counterintuitive higher oral vs silent reading rates (Cantonese Chinese, French, and Thai). In Japanese, Serbian, and Korean, the average silent reading rates are remarkably higher than their oral counterparts, while in the other languages oral and silent reading rates are very similar.

Moreover, a significant correlation exists at all levels of aggregation, which suggests rejecting the hypothesis of disconnected oral and silent reading rates.

To further evaluate the influence of the factors involved, we first ran a mixed-effects model on all data with *SilSR* as dependent variable, *OrSR* as independent fixed descriptor, and Text and Subject Identity as independent random descriptors. This model aimed at checking the consistency of the relationship between *SilSR* and *OrSR* when the variability of texts and subjects was accounted for. A z-score transformation was applied to the numeric data in order to get effect estimates of comparable magnitudes.

We found a significant effect of *OrSR* ($t=6.93$, $p < .001$***), as well as significant effects of Text ($\chi^2(1) = 29.6$, $p < .001$***) and Subject Identity ($\chi^2(1) = 1497.5$, $p < .001$***). This result reinforces our previous statement obtained from simple correlations. We furthermore fitted a second model to *SilSR* with *OrSR*, Language and subject's Sex as fixed-effect predictors, and Subject Identity and Text as random-effect predictors. We once again found a significant effect of *OrSR* ($t=5.38$, $p < .001$***), of Text ($\chi^2(1) = 30.2$, $p < .001$***) and Subject Identity ($\chi^2(1) = 1221.8$, $p < .001$***). We found an additional effect of Language ($F=7.38$, $p < .001$***), but no effect of Sex.

It therefore appears that silent and oral reading rates are strongly (and positively) related. Detailed analysis of the estimates for the various languages in the last model suggests that most of the effect of Language comes from Serbian, Japanese and Korean, as visible in Figure 1.

3.2. Balance between Information Density and Syllabic Rate

As previously stated, Pellegrino et al. [10] showed a negative correlation between information density and syllable rate for a different set of languages. We assessed this result with our data, and found a strong negative correlation between information density and both silent and oral reading rates at the language level (in both cases (N=8), Spearman's Rho = -.81, p=.021**). Figure 2 displays the distribution of information density and oral syllabic reading rate.

As previously, we further assessed the result with random-effects models. We first fitted a model to *SilSR* with *ID*, Language and Sex as fixed effect predictors, and Subject Identity and Text as random-effect predictors. We got a significant effect of *ID* ($t=-5.58$, $p < .001$***), of Text ($\chi^2(1) = 39.1$, $p < .001$***) and Subject Identity ($\chi^2(1) = 1335.3$, $p < .001$***). We found an additional effect of Language ($F=8.52$, $p < .001$***), but no effect of Sex. Replacing *SilSR* with *OrSR* as the dependent variable, we found significant effects for *ID* ($t=-7.29$, $p< .001$***), Text ($\chi^2(1) = 200$, $p < .001$***) Subject Identity ($\chi^2(1) = 1200$, $p < .001$***), Language ($F=28.25$, $p < .001$***) and Sex ($F=5.78$, p=.019**). Using Maximum Likelihood (ML) to compare our models, we found *OrSR* to be much better predicted by its predictors than *SilSR* (AIC=557 versus AIC=1262). Text and Language appear to be the two

predictors explaining most of this difference. Significant differences additionally exist between languages.

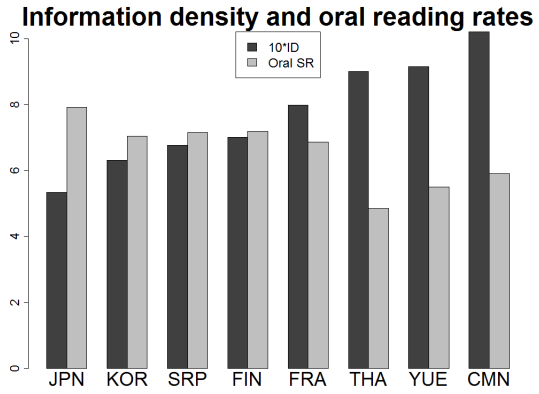

Figure 2: *Information density (unitless) and oral syllabic rate (#syl/s).*

We can therefore conclude that both silent and oral syllabic reading rates significantly negatively correlate with information density: the more information a syllable contains in average, the slower the reading rates are (at language level). Interestingly enough, we observe an effect of Sex for oral rates, with females speaking slower than males, but no effect for silent reading.

3.3. Relation between Duration and Text length

While there is only a weak correlation between *SilD* and the number of syllables (σ) (Pearson's R = .11, p < .001***), a stronger relationship exists between *OrD* and σ (Pearson's R = .71, p < .001***).

In the line of the results reported in [6], we compared 3 different mixed-effects models with *OrD* as dependent variable, Sex and either i) σ, ii) log(σ) or iii) exp(σ) as fixed predictor, and Language, Text, and Subject Identity as random predictors. We compared these 3 models in terms of AIC to identify the best linear and non-linear fit between σ and *OrD*. Results are given in Table 2.

Model	AIC (ML)
OrD ~ σ	435.5
OrD ~ log(σ)	335.9
OrD ~ exp(σ)	1515.2

Table 2. *Estimation of the best relationship between oral duration and number of syllables.*

In all models, we found significant effects for Text, Subject Identity, Language and Sex (with females having longer durations than men). Employing the logarithm of the number of syllables significantly improves the model compared with the other two (p < .001***).

Replacing *OrD* with *SilD* as the dependent variable, we found that a logarithmic relation results in the best fit as well. Two differences are however worth mentioning: i) AIC are significantly lower for *SilD* than for *OrD*, and ii) contrary to previous models, no effect of Sex is ever observed.

4. Discussion

Whereas previous studies on differences between oral and silent reading have focused only on one or two languages (in bilingual context), this paper proposes a first evaluation on a larger language set and yields promising results.

The main result is that oral and silent reading rates are significantly correlated thus confirming that cross-linguistic differences in word structure complexity influences the phonological processing in both reading modes. Moreover, both silent and oral reading rates significantly negatively correlate with information density, extending results from [10] in two directions: new languages and silent reading modality.

Second, a non-linear (logarithmic) relationship is found between text lengths and reading durations, suggesting that predictability increases with longer texts. This account is compatible with the idea that word-level effects shown in [7] accumulate at text level. Moreover, this non-linear trend is observed for both oral and silent reading, although weaker for silent reading.

An additional result is that Sex appeared to be a significant predictor in all models with oral rate or duration as dependent variable, but not in models addressing silent reading. This could suggest different processes taking place, or more likely a sociolinguistic effect when it comes to orality (e.g. [14]).

This study finally suggests that the script type is not the main factor when it comes to reading speed. Ideographic Cantonese and Mandarin indeed showed rates similar to languages using alphabets like Finnish, French or Thai. One non-alphabetic script – Japanese – behaved similarly to alphabet-based Korean and Serbian.

This line of research paves the way for future research in several directions. First, we did not check participants for text comprehension during silent reading. A more precise control (with self-paced reading for instance) may result in a reduction of inter-individual variation. Second, for 5 out of 8 languages, silent and oral rates are very similar but for Japanese, Korean and Serbian, differences are much larger. This may be due to either linguistic parameters (these languages exhibit the lowest Information density in the sample) or differences in reading abilities among the groups of participants. This last hypothesis is based on two Korean and one Serbian subjects who exhibit very high silent reading rates, suggesting the possible use of speed reading. Finally, a better experimental structure could be designed to record silent and oral rates in a more symmetrical fashion, rather than recording one after the other.

5. Acknowledgements

The authors would like to warmly thank Dominic Massaro for his suggestion to work on silent reading rate and Dr. Ching-Pong Au, Pr. Feng Wang, Ms. Ubonwan Chamnansan, Ms. Mladenka Kovačević, Mrs. Miyuki Ishibashi, Mrs. Keiko Guillaume, Dr. Hannu Laaksonen and the *Centre for Research in Speech and Language Processing* at Chulalongkorn University for help with translation and recording. They are grateful to the LABEX ASLAN (ANR-10-LABX-0081) of Université de Lyon for its financial support within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) of the French government operated by the National Research Agency (ANR).

6. References

[1] Naveh-Benjamin M. and Ayres T. J. “Digit Span, Reading Rate, and Linguistic Reactivity”, Q. J. Exp. Psychol., 38A: 739–751, 1986.

[2] O’Brien B. A., Wallot S., Haussmann A. and Kloos H. “Using Complexity Metrics to Assess Silent Reading Fluency: A Cross-Sectional Study Comparing Oral and Silent Reading”, Sci. Stud. Read., 1–20, 2013.

[3] Hiebert E. H. and Samuels S. J. “Comprehension-Based Silent Reading Rates: What Do We Know? What Do We Need to Know?”, Lit. Res. Instr., 51: 110–124, 2012.

[4] Massaro, D. “Acquiring Literacy Naturally”, American Scientist, 100(4): 324, 2012.

[5] Fuchs L. S., Fuchs D., Hosp M. K. and Jenkins J. R. “Scientific Studies of Reading Oral Reading Fluency as an Indicator of Reading Competence : A Theoretical, Empirical, and Historical Analysis”, Sci. Stud. Read., 5(3): 239–256, 2009.

[6] Denton C. A., Barth A. E., Fletcher J. M., Wexler J., Vaughn S., Romain M., and Francis D. J. “The Relations Among Oral and Silent Reading Fluency and Comprehension in Middle School: Implications for Identification and Instruction of Students With Reading Difficulties”, Sci. Stud. Read., 15(2): 109–135, 2011.

[7] Smith N. J. and Levy R. “The effect of word predictability on reading time is logarithmic”, Cognition, 128(3): 302–19, 2013.

[8] Norris, D. “The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process”, Psychological Review, 113(2): 327–357, 2006.

[9] Wright, J. A. “The Impact of Oral Fluency and Silent Fluency on the Comprehension of Fourth Graders”, Diss. Louisiana State University, 2011.

[10] Pellegrino F., Coupé C., and Marsico E. “Across-language perspective on speech information rate”, Language 87: 539–558. 2011.

[11] Chetail, F. “Effect of number of syllables in visual word recognition: new insights from the lexical decision task”, Language, Cognition and Neuroscience. DOI: 10.1080/23273798.2013.876504. 2014.

[12] Ferrand, L. “Reading aloud polysyllabic words and nonwords: The syllabic-length effect re-examined”, Psychonomic Bulletin & Review, 7: 142-148, 2000.

[13] New, B., Ferrand, L., Pallier, C. and Brysbaert, M. “Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project”, Psychonomic Bulletin & Review, 13: 45-52. 2006.

[14] Spoehr, K. T. “Phonological encoding in visual word recognition”, Journal of Verbal Learning and Verbal Behavior, 17: 127-142. 1978.

[15] Campione E. and Veronis J. “A multilingual prosodic database”, Proc. of the 5th International Conference on Spoken Language Processing (ICSLP'98), Sydney, Australia, 3163-3166. 1998.

[16] Ferragne, E., Flavier, S. and Fressard, C. “ROCme! Software for the recording and management of speech corpora”, Proc. of Interspeech 2013, Lyon, France, 25-29 august. 2013.

[17] Boersma, P. “Praat, a system for doing phonetics by computer”, Glot International 5(9/10): 341-345. 2011.

[18] Ribeiro Jr., P. J. and Diggle P.J. “geoR: A package for geostatistical analysis”, R-NEWS 1(2), ISSN 1609-3631, 2001.

[19] Douglas B., Maechler M., Bolker B. and Walker S. “lme4: Linear mixed-effects models using Eigen and S4”, R package version 1.1-5. http://CRAN.R-project.org/package=lme4. 2014.

[20] Box, G. E. P. and Cox D. R. “An analysis of transformations (with discussion)”, Journal of the Royal Statistical Society B 26: 211–252, 1964

[21] Jacewicz, Ewa, Fox, R.A., O'Neill C. and Salmons, J. “Articulation rate across dialect, age, and gender”, Language Variation and Change. 21(2): 233-256, 2009.