Consonant and vowel transpositions effects during reading development: A study on Italian children and adults.

Giacomo Spinelli[1], Lucia Colombo[2], and Stephen J. Lupker[3]

[1] Università di Milano-Bicocca
[2] Università di Padova
[3] University of Western Ontario

Word count: 14853

Key words: Masked priming, transposition effect, consonant-vowel status, reading development

Address correspondence to:
Lucia Colombo
Dipartimento di Psicologia Generale
Via Venezia 8, 35141 Padova, Italy
Email: lucia.colombo@unipd.it

Abstract.

Recently, Colombo, Spinelli, and Lupker (2020), using a masked transposed letter (TL) priming paradigm, investigated whether consonant/vowel (CV) status is important early in orthographic processing. In four experiments with Italian and English adults they found equivalent TL priming effects for CC, CV and VC transpositions. Here, we investigated that question with younger readers (age 7 to 10) and adults, as well as whether masked TL priming effects might have a phonological basis. That is, because young children are likely to use phonological recoding in reading, the question was whether they would show TL priming that is affected by CV status. In Experiment 1, target words were preceded by primes in which two letters (either CV, VC or CC) were transposed versus substituted (SL). We found significant TL priming effects, with an increasing developmental trend, but, again, no letter type by priming interaction. In Experiment 2 the transpositions/substitutions involved only pairs of vowels with those vowels having either diphthong or hiatus status. The difference between these vowel clusters is only phonological, thus the question was would TL priming interact with this factor. TL priming was again found with an increasing trend with age, but there was no vowel cluster by priming interaction. There was, however, an overall vowel cluster effect (slower responding to words with hiatuses) which decreased with age. The results suggest that TL priming only taps the orthographic level, and that CV status only becomes important at a later phonological level.

The impact on the reading process of transposing two letters in a word has been investigated using a number of different experimental paradigms. The most popular is the masked priming lexical decision paradigm. In this paradigm, a forward mask (e.g., ######) is initially presented followed by a briefly presented prime which is then followed by a target word/nonword, to which a lexical decision is made. In investigations of transposed letter (TL) priming, the prime is typically either the target with two internal letters transposed (TL primes; e.g., chidlren) or the target with those same two letters substituted for (SL primes; e.g., chistren). The typical result is that lexical decisions are faster after a TL than an SL prime (a TL priming effect; Colombo, Spinelli, & Lupker, 2020; Forster, Davis, Schoknecht, & Carter, 1987; Lupker, Perea, & Davis, 2008; Perea & Lupker, 2003, 2004). This effect has been replicated in an increasing number of different languages including English, Spanish, French, Italian and Japanese.

Several explanations have been provided for the TL priming effect, all based, to some degree, on the idea that letter position and identity are coded somewhat independently during the orthographic coding process (Adelman, 2011; Davis, 2010; Gómez, Ratcliff, & Perea, 2008; Grainger & van Heuven, 2003; Norris & Kinoshita, 2008, 2012; Whitney, 2001). One thing that these models have in common is that the orthographic input coding scheme is assumed to be the same for consonants and vowels. In fact, word recognition models are, in general, silent about any potential consonant-vowel differences in the orthographic code.

Recently, Colombo et al. (2020) investigated a consonant-vowel contrast in a masked priming lexical decision task in both Italian and English, showing that the consonant-vowel status of the letters was not relevant to the size of their priming effects. They used TL and SL primes that involved only adjacent letters: either two consonants (CC; TL, elefatne, from ELEFANTE ‘elephant’; SL, elefable), a consonant and a vowel (CV; TL, eleafnte, SL, eleolnte) or a vowel and a consonant (VC; TL, elefnate; SL, elefmote). The three TL priming effects were, in all cases, significant and, more importantly for the present discussion, were equivalent in size in each of the four experiments, two carried out in Italian and two in English. The fourth experiment, in English, used a slight variation of the typical paradigm, called sandwich priming (Lupker & Davis, 2009), in order to increase the size of the priming effect and increase the probability of finding an interaction between the type of letters (i.e., consonants versus vowels) involved in the transposition and the size of the priming effect. Nonetheless, as just stated, the size of the TL priming effect was identical in the three conditions. That is, there was no evidence of a consonant-vowel difference in the size of TL priming effects.

In contrast to Colombo et al.'s (2020) results, the literature on masked TL priming does provide some suggestion that that there may be consonant-vowel differences in that transpositions involving two vowels (e.g., cinaso for the base word casino) can act somewhat differently than other types of transpositions (i.e., transpositions involving two consonants). Perea and Lupker (2004), for example, using Spanish word targets (e.g., casino) and transposing/substituting non-adjacent letters, either consonants (TL, caniso; SL, caviro) or vowels (TL, cisano; SL, cesuno), found significant TL priming from consonant transpositions, but not from vowel transpositions. In the following years, similar patterns have been found in other masked priming lexical decision experiments in Spanish (Acha & Perea, 2008; Comesaña et al., 2016; Perea & Acha, 2009; see also Carreiras, Vergara, & Perea, 2007, 2009; Johnson, 2007) and in English (Lupker et al., 2008). More recently, however, Yang and Lupker (2020) were unable to replicate the pattern in English (i.e., consonant and vowel transpositions produced equivalent TL priming effects), a result that led those researchers to conclude that differences between consonants and vowels typically emerge at later stages than the orthographic coding stage, the stage that is typically examined with masked priming. Nonetheless, at least in Spanish, a language that is similar to Italian in many respects, there seems to be evidence for the existence of a consonant-vowel difference in TL priming effects.

In thinking about how to interpret these results, what needs to be noted is that the crucial difference between consonants and vowels is phonological, not orthographic, in nature. Thus, the existence of empirical data from Spanish showing consonant-vowel differences in TL priming suggests that, although a consonant-vowel difference did not emerge in Italian adult readers in

Colombo et al.'s (2020) experiments, it may emerge in readers who rely on phonological information to a larger extent, specifically developing readers. In the present research we examined that possibility. That is, we compared the performance of three groups of children and one group of adults in a TL masked priming paradigm, manipulating the consonant-vowel status of the transposed/substituted letters.

When young children start reading, they typically rely on a procedure involving spelling-sound translation, in which each grapheme is converted to its corresponding phoneme before assembling the phonemes in order to identify the word. This process is much easier, and more efficient, for children learning to read transparent languages, like Italian, in which such a procedure always yields a correct pronunciation. From this slow and effortful process, the reading process eventually develops to become one in which phonological coding is a fast, automatic process. In addition to the phonological coding process, however, an orthographically-based reading process also develops in which readers are able to exploit units of increasing size in the orthographic code which allows them to access higher-level representations directly, that is, without involving a phonological activation process (Share, 1995; Ziegler & Goswami, 2005). Hence, it is, perhaps, not a great surprise that Colombo et al. (2020) observed no vowel-consonant differences in the adult readers in their experiments.

The dynamics of the phonological and orthographic processes in developing readers has been the motivating factor for a number of studies addressing TL effects and has given rise to two models. One is the multiple route model (Ziegler, Bertrand, Leté & Grainger, 2014; Grainger et al., 2012), in which a coarse-grained mechanism of orthographic coding is gradually acquired by children as they learn the connections between orthography and semantics. The relevant orthographic mechanism is very flexible concerning differences in letter position and, hence, is the source of TL effects.[1] Based on the idea that children at a very early stage of reading acquisition use a phonologically-based serial process of conversion from print to sound (which is ultimately replaced/supplemented by this alternative, coarse-grained orthographic coding mechanism), the prediction that TL priming effects should be small or nonexistent in young readers and constantly increase with reading experience follows directly from the model. Essentially, because the phonological procedure used by younger readers must necessarily specify the exact position of each letter in order to assemble the correct pronunciation of a word, TL effects should not be observed in those readers, with TL effects only emerging once those readers have developed experience using their coarse-grain coding scheme.

In an attempt to gain support for the multiple route model, Ziegler et al. (2014) compared TL and pseudohomophone priming effects as markers, respectively, of coarse-grained orthographic coding and of the involvement of phonological processing. They found that, in French, pseudohomophone priming effects remained constant through the five grades, supporting the idea that phonology is constantly involved in reading in younger readers, while TL priming effects tended to increase with age. It is worth noting that these results were obtained using standardized (z-transformed) RT data, i.e., data where absolute differences in RTs among groups differing in reading ability were removed (Faust, Balota, Spieler, & Ferraro, 1999). However, not all studies examining the development of TL priming effects applied a z-transformation to their latency data, a choice that, as will be discussed, may have played a role in the pattern of results obtained.

Ziegler et al.'s (2014) conclusions were further supported by Eddy, Grainger, Holcombe, and Gabrieli (2016) who showed through EEG measurements in children aged 8 to 10 that only TL priming effects, but not phonological priming effects, correlated with level of reading ability. These results suggest that TL priming effects reflect the increasing use of a coarse-grained orthographic coding mechanism as reading ability increases and do not emerge as a result of phonological coding mechanisms.

A different view – the lexical tuning hypothesis – has been proposed by Castles and collaborators (Castles, Davis, Cavalot, & Forster, 2007; Kezilas, McKague, Kohnen, Badcock, & Castles, 2017) who argued that sensitivity to letter position does not depend either on the dynamics between phonological and orthographic processes or the development of a course-coding scheme but on developmental changes in the flexibility of orthographic coding. According to this hypothesis, young readers use a very flexible mechanism of lexical tuning which associates the orthographic code with many orthographically similar words in the lexicon. Only with increased reading experience does this mechanism become more precise and, therefore, more intolerant to changes in letter position. The basic prediction from this view for young readers is just the opposite of the one from the multiple route model: TL priming effects should be larger for younger readers, because their orthographic coding is quite flexible, and should decrease with an increase in reading experience.

Castles and collaborators' lexical tuning hypothesis was supported by finding that prime-target pairs like lpay-PLAY (TL condition) produced priming compared to rlay-PLAY (a one-letter different condition) only in second-grade children, not in older children or adults in a masked priming lexical decision task (Castles et al., 2007), with older children producing equivalent priming effects in the two conditions (in comparison to an unrelated prime). Similarly, Acha and Perea (2008) reported (in Spanish) a larger TL-SL priming effect in third-grade children than in sixth-grade children and adults, although the effect was significant in all age groups. Note, however, that these results were obtained using untransformed ("raw") RT data, i.e., data where absolute differences in RTs among groups differing in reading ability were not removed.

Slightly in contrast, using z-transformed RTs, Kezilas et al. (2017) found (in English) a TL-SL priming effect in young readers with the effect being constant across ages (see also Paterson, Read, McGowan, & Jordan, 2015). Kezilas et al. also found, however, that compared to an identity condition (e.g., play-PLAY), TL primes produced an increasing cost across ages, with no cost observed for young readers. They interpreted this result as evidence that transpositions are well tolerated early in development, making TL primes as effective as identity primes, but with increasing reading experience, transpositions become less tolerated and produce larger processing costs. That is, in line with the lexical tuning hypothesis, these results suggest that considerable position flexibility is present in beginning readers, but may decline as reading experience increases.

Colombo, Sulpizio, and Peressotti (2017, 2019) tested these alternative views about the nature of TL priming in Italian, a language characterized by a stronger spelling-sound correspondence than exists in either English or French using a slightly different paradigm. Colombo et al. (2017) showed TL effects in the rejection time for nonwords created from real words (ALBERGO 'hotel'; TL nonword ablergo, SL nonword acmergo) in an Italian lexical decision task in children in second, third and fifth grade as well as adults. That is, lexical decision latencies to TL nonwords like ablergo were slower and more error prone than to SL nonwords. More relevantly, this difference increased with age (when analyzing z-transformed RTs), a pattern that is consistent with the idea that TL effects increase as the orthographic coding system develops, as predicted by the multiple route model.

Because the pattern observed by Colombo et al. (2017) in their unprimed lexical decision task may arguably have involved a phonological component, Colombo et al. (2019) conducted a follow-up using a masked priming manipulation in order to provide a stronger case for the idea that these effects are orthographically based. Colombo et al. (2019) used a masked priming lexical decision task with a sandwich priming procedure with Italian second, third and fifth graders and adults to further investigate the developmental trend of TL effects. They also manipulated the serial position of the transposition/substitution to see whether the priming effect would vary as a function of that factor. Similar to TL effects in their unprimed lexical decision task in which TL and SL nonwords served as targets (i.e., Colombo et al., 2017), Colombo et al. (2019) found a significant interaction between age group and priming effect in standardized RTs, with the latter being significant only in fifth graders and adults. Also important is the fact that Colombo et al. (2017), in their experiment involving responding to nonword targets, found a serial effect of the TL/SL position, with a larger effect when the transposition was towards the end of the letter string. In contrast, the masked priming experiment in Colombo et al. (2019) did not show any effect of position, suggesting that the phonological component was less, or not at all, involved. This contrast provides at least some evidence that the TL priming effect in Colombo et al. (2019), where no serial position effect was found, would be more relevant to understanding the development of the orthographic coding process.

Summarizing, there are two different views on how orthographic processes develop. According to the first view, represented by the multiple route model (Grainger et al., 2012; Ziegler et al., 2014), children initially use a phonological recoding procedure, but, while this procedure becomes increasingly automatic, a coarse (i.e., more imprecise) orthographic coding process responsible for TL effects starts developing. Because this process becomes increasingly used as reading skills develop, it follows that TL effects should increase with age. According to the second view, represented by the lexical tuning hypothesis (Castles et al., 2007; Kezilas et al., 2017), the orthographic coding process is more flexible initially, allowing large TL effects, but as it becomes more efficient with increasing reading ability, it also becomes less flexible. Thus, TL effects should decrease with age.

These contrasting views have both received empirical support, but, until now, no study has been able to definitely settle the issue. For example, in a recent study examining TL priming effects in developing readers, Hasenhäcker and Schroeder (2021) found a somewhat ambiguous pattern of results, partially consistent with both types of models, despite using a longitudinal design, a larger number of children and additional control conditions (a completely unrelated condition and an identity condition) in addition to the TL and SL conditions. Further, it is also worth noting that most of the data supporting the type of position taken by the multiple route model have come from analyses in which the latencies have been z-transformed, a procedure which tends to reduce effect sizes for beginning readers. In contrast, most of the data supporting the position taken by the lexical tuning hypothesis have come from analyses in which raw RTs were analyzed in spite of the fact that they are much longer for younger readers (Kezilas et al., 2017, appears to be one of the few exceptions). We will return to this issue in the Discussion of Experiment 1, however, we should note that, in the present paper, we have opted to report analyses using both approaches.

Returning to the issue of how and when consonant and vowel differences emerge during the reading process, as noted, we would expect that those differences are more likely to emerge when phonology is involved. Relevant to this point, Perea and Acha (2009) examined TL priming effects for vowels and consonants in adult participants using lexical decision and cross-case matching (also known as the same-different task). In the latter task, participants are presented with a reference stimulus followed by a target and are required to decide whether they are same or different (Norris & Kinoshita, 2008). The target was preceded by a TL or an SL prime. Because of its (presumably pre-lexical) nature, this task has been claimed to tap very early processes in word identification, before information about the consonant-vowel status of a word becomes available. Perea and Acha (2009) found a significant TL priming effect in both tasks, but they only found a consonant-vowel dissociation in lexical decision. In the cross-case matching task, the TL priming effect was equivalent for consonant and vowel transpositions. These results were explained by the authors by assuming that the cross-case matching task mainly reflects early orthographically-based processes which are insensitive to the nature of letters as vowels or consonants, while in lexical decision phonological effects may be to some extent involved.

The involvement of phonology in masked priming effects, and the consequent possible appearance of a CV dissociation was further examined in a developmental study by Comesaña, Soares, Perea, and Marcet (2016). They found that, in Spanish, a consonant-vowel difference in raw (untransformed) RTs effects for adults (who showed a TL priming effect for CC transpositions, but no such effect for VV transpositions, replicating Perea and Acha, 2009), but not for fourth graders (who showed equal-sized TL priming effects for CC and VV transpositions). Importantly, note that Comesaña et al. claimed, based on their results as well as the findings by Davis, Castles, and Iakovidis (1998) who found no phonological priming in children, that developing readers do not make use of phonology in reading. In any case, these results suggest that phonological coding not only may have some role in TL priming effects (because, as noted, the difference between consonant and vowels is phonological), but also that this role may become more relevant to shaping the pattern of TL priming effects with age. The more general point that the studies reviewed above suggest is that TL priming effects may depend on the interplay between phonological and orthographic processes rather than on one type of process in particular.

As noted above, Colombo et al. (2020) found no consonant-vowel differences in adults in masked priming experiments where consonants and vowels were transposed (i.e., CC versus CV versus VC transpositions). Given that little evidence of TL priming effects was found in second and third graders (Colombo et al., 2019) and that, at this grade level, Italian children are likely to be mostly using a phonological procedure, it is possible that the reason for the difficulty in observing TL priming in younger children is that any TL priming that emerges in younger children would be phonologically based and, hence, might be affected by consonant–vowel differences in the early stages of reading development, differences which would diminish in older children and, potentially, disappear in adults.

This hypothesis suggests a potential interaction between TL priming effects and consonant-vowel status in younger children. In younger children, some priming may be expected for CC transpositions because they involve letters belonging to the same phonological class (i.e., consonants). Thus, phonologically, the resulting nonword prime would resemble its base word quite well. This idea is consistent with Ziegler et al.'s (2014) data for TL priming on CC transpositions where, albeit small, a TL priming effect did seem to be present in younger children. In contrast, CV and VC transpositions may produce little or no priming because they involve the transposition of phonologically distinct letters (i.e., a consonant and a vowel). These transpositions, unlike CC transpositions, would result in a nonword prime that produces little priming because, phonologically, it bears little resemblance to its base word. This idea would also explain the fact that Colombo et al. (2019) found no overall TL priming in younger children in a study where CC, CV, and VC transpositions were used: With little or no priming from two-thirds of the transpositions (CV and VC) and some priming from the other third (CC transpositions), observing an overall TL priming would be difficult. In fact, a re-analysis of Colombo et al.'s (2019) results provides some support for this idea, because for third graders, CV and VC transpositions indeed produced negative and no priming, respectively, whereas CC transpositions produced regular priming. For second graders, however, the opposite was true, with CC transpositions producing no priming and CV and VC transpositions producing regular priming. However, that experiment was not designed to contrast different types of transpositions. A clearer picture should emerge from one that was specifically designed to do so.

The present Experiment 1 was such an experiment. In Experiment 1, as in Colombo et al.'s (2020) Experiment 2, consonants and vowels were manipulated with the transpositions/substitutions being CC (TL, ablergo; SL, acmergo), CV (TL, alebrgo; SL, alimrgo), or VC (TL, albrego; SL, albcigo). Specifically, we hypothesized that younger children would show smaller priming effects for CV and VC transpositions compared to CC transpositions, a difference that should diminish with age and disappear in adults.

Experiment 1.

Method.

Participants. One hundred and thirty-four elementary school children participated in this experiment. Seven participants were excluded from the analyses because they committed an error on more than 40% of the trials in the task (5) (following Ziegler et al., 2014), because they were diagnosed with a developmental disorder (1), or because they were recent immigrants to Italy (1). Out of the 127 remaining students, 45 were second graders (mean age 7.51, range 7-8, 28 females), 33 were third graders (mean age 8.45, range 8-9, 21 females), 18 were fourth graders (mean age 9.61, range 9-10, 8 females), and 31 were fifth graders (mean age 10.48, range 10-11, 14 females). For all the students, parents and the school director gave informed consent. All the students (even the dyslexic ones) had typical development (as reported by their teachers) and each performed a partial version of the standardized test of word reading (Sartori, Job, & Tressoldi, 1995) prior to performing the experiment. Children were tested during the second part of the academic year in a quiet room at their school. A further sample of 49 adults (university students - mean age 24.80, range 20-35, 35 females) was also tested to serve as a comparison with the children. All participants included in the analysis were native speakers of Italian and had normal or corrected-to-normal vision.

These sample sizes were determined so that we would have a power of at least .80 to detect, in at least one of our age groups, consonant-vowel differences of similar size as those reported by Comesana et al. (2016) for Spanish adults, i.e., η_p^2 = .46. An a-priori power analysis conducted with G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) revealed that 13 participants per group would be required to detect an effect of that size in one of the age groups. After fourth and fifth graders were aggregated (see below), Experiment 1 exceeded that target.

Materials. Ninety-six Italian words were selected from the version of the Lessico Elementare database (Marconi, Ott, Pesenti, Ratti, & Tavella, 1993) available in the Q2Stress database (Spinelli et al., 2017), to serve as target words. Lessico Elementare is based on both materials written by adults for children and materials written by children themselves, with the Q2Stress version of this database retaining only the former materials, along with some additional information about those words (see Spinelli et al., 2017). The mean length of the target words is 7.08 letters (range = 6-8), the mean neighborhood size is 3.22 (range 0-9), and the mean word frequency (per million) is 149.81 (range = 24.52-1500.75). Six different nonword primes were created for each target word: (1) CC transposition (puslante-PULSANTE), (2) CC substitution (purmante-PULSANTE), (3) CV transposition (pulasnte-PULSANTE), (4) CV substitution (pulucnte- PULSANTE), (5) VC transposition (plusante-PULSANTE), and (6) VC substitution (prosante-PULSANTE).

Transpositions never involved the first or the last letter of the targets words and CC, CV, and VC transpositions were matched on the position of the first transposed/substituted letter (*M* = 3.40, *M* = 3.33, and *M* = 3.25, respectively). Orthographic neighborhood size (Coltheart, Davelaar, Jonasson, & Besner, 1977) based on the CoLFIS database (Bertinetto et al., 2005), and bigram frequency based on the Lessico Elementare database (Marconi et al., 1993), were also matched across all six prime types.[2] Ninety-six orthographically legal nonwords were created to serve as nonword targets. They were paired with CC, CV, and VC transposition and substitution primes in the same proportions as in the parallel conditions for the word targets. The word targets were divided into six sets of 16 and each set was primed by primes from one of the six prime conditions for a given participant. Six lists were created to complete the counterbalancing, and participants were randomly (and equally) assigned to one of those lists. No counterbalancing was conducted for the nonword targets (i.e., each nonword was primed by the same prime across participants).

Procedure. Children were tested individually in a quiet room at their school. Adult participants were tested individually in a quiet laboratory at the University of Padova. E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA; see Schneider, Eschman, & Zuccolotto, 2002) was used for stimulus presentation and data collection. To amplify the priming effects, the sandwich priming procedure was used (Lupker & Davis, 2009) as adapted by Ziegler et al. (2014). Each trial started with a fixation point (*) that lasted 1000 ms and was followed by the target word/nonword in lowercase letters for 27 ms in the same location as the fixation point. The target was followed at the same location by a prime in lowercase letters for 67 ms. Finally, the target was again presented but in uppercase letters and remained on the screen until the participant responded or for a maximum of 4 s. All stimuli were presented in Courier New-18 pt. font centered on the screen. Participants responded by pressing either the ‘‘m” (for a word) or ‘‘z” (for a nonword) key on the keyboard. They were instructed to decide and respond as quickly and accurately as possible whether the letter string they saw on the screen was a real word or a nonword. Participants were not informed of the presence of the primes. Participants completed 24 practice trials followed by a randomized list of the experimental trials.

Results.

No items were excluded from the analyses because error rates were below 40% for all items (Ziegler et al., 2014). In this and the following experiment, nonword data were not analyzed because the prime conditions were not counterbalanced over the nonword targets. Incorrect responses were removed from the latency analyses. For both latency and error analyses, correct response times faster than 300 ms were removed as outliers for both children and adults (.6% of the observations). Correct response times more than 3 SDs (for children) or more than 2.5 SDs (for adults) from each participant's mean (for correct responses to the words) were also removed as outliers (2.2%; Colombo et al., 2017).

Latency analyses were conducted on both raw RTs and z-transformed RTs (z-RTs) obtained by subtracting, for each trial for each participant, the mean of the participant's latencies from the trial latency and dividing this difference by the standard deviation of the participant's latencies (e.g., Hasenäcker & Schroeder, 2021; Kezilas et al., 2017). The reason that this z-score calculation was based on individual trial latencies, as opposed to condition means (the most common procedure for z-score calculations; Faust et al., 1999) was to accommodate our use of mixed-effects modelling, an analytical procedure that is typically performed on individual trial latencies rather than condition means. Mixed-effects modelling permits use of both continuous and categorical variables (the fixed effects) while controlling for variance among the participants and the items being used (the random effects: Baayen, 2008; Baayen, Davidson, & Bates, 2008). The specific class of mixed-effects models that we used was a Generalized Linear Mixed Model (GLMM). Generalized linear models can be used to fit the dependent variable when the normal distribution assumed by linear models is not appropriate. Raw RTs, typically right-skewed, are fit reasonably well with the Gamma and Inverse Gaussian distributions (Lo & Andrews, 2015). Because the z-RTs obtained in the fashion that we described have a similar (right-skewed) distribution as raw RTs, either of those distributions would also be appropriate. Because, however, both of those distributions assume positive values, a constant value of 10 was added to all z-RTs to make all point values positive and allow model fitting. This addition only influences the intercept of the model (which changes from 0 to 10, i.e., the constant) and has essentially no bearing on the parameters and the statistical significance of the random and fixed effects.

For the purposes of the analyses, fourth and fifth graders were aggregated because there were comparatively few participants in those grades individually, especially in the fourth grade. For all dependent variables (i.e., raw RTs, z-RTs, and errors), a 3 (Letter Type: CC vs. CV vs. VC, within-subject and within-item) X 2 (Prime Type: TL vs. SL, within-subject and within-item) X 4 (Age Group: second graders vs. third graders vs. fourth-and-fifth graders vs. adults, between-subjects and within-item) design was used for the fixed effects. In addition, the grouping/counterbalancing factor was also included as a fixed effect in order to account for variance associated with the participant groups created for counterbalancing (Pollatsek & Well, 1995). Effects or interactions involving that factor are of no importance to the main questions and will not be reported. Subjects and items (i.e., the target stimuli) were treated as random effects for raw RTs and errors, whereas for z-RTs, items were treated as the only random effect. The reason is that, because the z-transformation essentially eliminates all existing variability between overall latencies in different participants, entering subjects as a random effect produces a singularity error (i.e., the variance associated with that random effect will be zero).

All analyses were conducted in R version 3.6.3 (R Core Team, 2020). Prior to running the models, R-default treatment contrasts were changed to sum-to-zero contrasts (i.e., contr.sum) to help interpret lower-order effects in the presence of higher-order interactions (Levy, 2014; Singmann & Kellen, 2019). The models were fit by maximum likelihood with the Laplace approximation technique. The lme4 package, version 1.1-23 (Bates et al., 2015), was used to run the GLMM. The function Anova in the car package, version 3.0-7 (Fox & Weisberg, 2019), was used to obtain probability values, specifying Type III Sums of Squares. The emmeans package, version 1.3.1 (Lenth, 2018), was used to conduct follow-up tests with the default adjustments for multiple comparisons. A Gamma distribution was used to fit raw RTs and z-RTs, with an identity link between fixed effects and the dependent variable (Lo & Andrews, 2015), whereas a binomial distribution with a logit link between the fixed effects and the dependent variable was used to fit the error data. Note that, in the current version of lme4, convergence failures for GLMMs, especially more complex models, are frequent, although many of those failures reflect false positives (Bolker, 2021). To limit the occurrence of convergence failures, for this and for the following set of analyses, we kept the random structure of the models as simple as possible by using only random intercepts for subjects and items. For the same reason, model estimation was performed using the BOBYQA optimizer and a maximum of 1 million iterations, settings which typically return equivalent estimates as the default settings but with fewer false-positive convergence warnings (Bolker, 2021).

In addition to GLMM analyses, traditional by-subject and by-item ANOVAs based on condition means were also conducted in R for raw RTs, z-RTs (obtained, in this case, using the standard procedure, that is, by subtracting the condition mean from the mean of the 6 conditions and dividing it by the standard deviation estimated based on the 6 conditions), and errors. For the sake of simplicity, we report the results of those analyses in the Supplementary Materials. The mean RTs, z-RTs, and error percentages based on by-subject means are presented in Table 1 (to make the results more interpretable, the z-RTs are presented before the constant was added). For this and for the following experiment, the raw data and the R files used for the analyses are publicly available at https://osf.io/fmau4/.

Table 1

Mean lexical decision RTs (in ms), z-transformed RTs, and percentage of errors (with the standard error of the mean in parentheses) for word targets in Experiment 1

		Prime Type								
		RTs			Z-RTs			Error %		
Age group	Letter type	TL	SL	TL priming	TL	SL	TL priming	TL	SL	TL priming
Second graders	CC	1545 (53)	1649 (59)	104	-.100 (.043)	.120 (.036)	.220	8.1 (1.5)	10.4 (1.6)	2.3
	CV	1561 (60)	1625 (62)	64	-.060 (.037)	.059 (.032)	.119	7.7 (1.3)	7.1 (1.3)	-.6
	VC	1594 (69)	1608 (63)	14	-.038 (.041)	.020 (.036)	.058	6.6 (1.2)	7.6 (1.1)	1.0
Third graders	CC	1225 (64)	1343 (65)	118	-.180 (.044)	.125 (.042)	.350	5.0 (1.1)	8.7 (1.5)	3.7
	CV	1249 (62)	1329 (65)	80	-.105 (.044)	.114 (.046)	.219	7.5 (1.0)	6.9 (1.2)	-.6
	VC	1262 (69)	1340 (66)	78	-.086 (.039)	.139 (.044)	.225	5.1 (1.1)	7.4 (1.2)	2.3
Fourth-and-fifth graders	CC	1029 (41)	1096 (43)	67	-.111 (.041)	.136 (.033)	.247	3.4 (.8)	5.2 (1.1)	1.8
	CV	1015 (42)	1102 (39)	87	-.160 (.031)	.198 (.037)	.358	2.9 (1.0)	4.3 (1.0)	1.4
	VC	1015 (36)	1074 (39)	59	-.127 (.037)	.073 (.035)	.200	4.2 (.8)	2.6 (.7)	-1.6
Adults	CC	559 (12)	615 (11)	56	-.188 (.041)	.211 (.033)	.399	2.0 (.6)	6.3 (.9)	4.3
	CV	554 (11)	611 (13)	57	-.222 (.039)	.216 (.044)	.438	2.7 (.6)	6.1 (1.0)	3.4
	VC	566 (13)	611 (14)	45	-.175 (.039)	.186 (.039)	.361	2.6 (.6)	5.6 (.9)	3.0

Note: TL = transposed letter prime; SL = substituted letter prime; CC = consonant-consonant transposition; CV = consonant-vowel transposition; VC = vowel-consonant transposition.

Raw RTs. The model failed to converge. As per the recommended troubleshooting procedure (see "convergence" help page in R), we restarted the model from the apparent optimum and tried all other available optimizers in addition to the BOBYQA optimizer. Although all optimizers failed to converge, all optimizers returned fairly similar results, suggesting that the convergence warnings were false positives. We report the results from the BOBYQA optimizer.

There was a main effect of Prime Type, $\chi^2 = 157.24$, $p < .001$, reflecting faster responses following TL than SL primes, the TL priming effect. There was also a main effect of Age Group, $\chi^2 = 1186.37$, $p < .001$. Follow-up tests indicated that each age group differed significantly from all other age groups, all *p*s < .001, with second graders being the slowest, followed by third graders, fourth-and-fifth graders, and adults. That is, there was a trend for latencies to decrease with age. Prime Type and Age Group interacted as well, $\chi^2 = 8.78$, $p = .032$. Follow-up tests comparing TL priming effects for each pair of age groups revealed that the source of this interaction was the fact that third graders produced an overall larger priming effect (90 ms) compared to adults (53 ms), $\beta = -40.06$, $SE = 14.8$, $z = -2.71$, $p = .007$. However, there was not a clear trend for priming effects to decrease with age, as the priming effect produced by the adults was comparable with that produced by second graders (60 ms), $\beta = -8.96$, $SE = 15.1$, $z = -.59$, $p = .554$ (the other comparisons were not significant either, all *p*s > .095). Note, further, that all age groups showed a significant effect of Prime Type when analyzed separately, all *p*s < .001. This pattern of results is represented in the left panel of Figure 1.

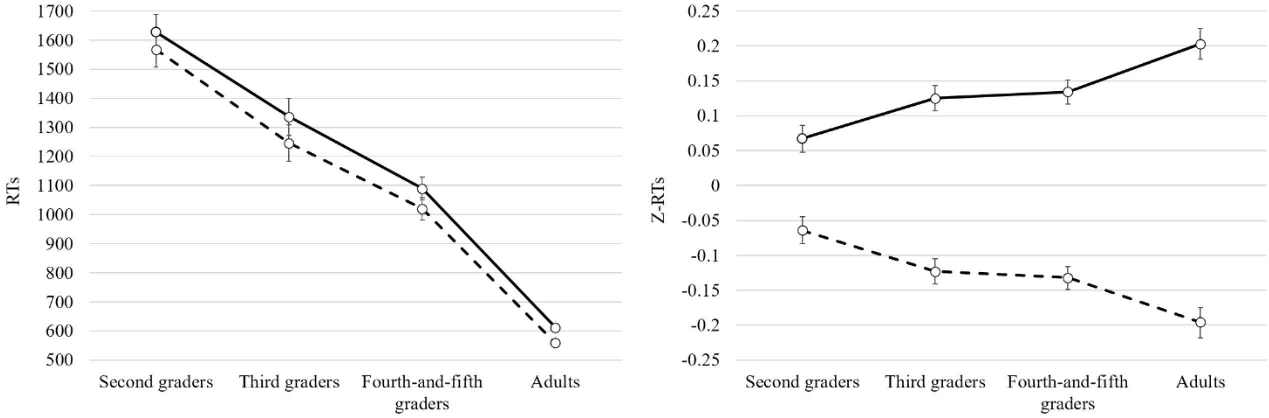
No other effect or interaction was significant (all *p*s > .3). In particular, replicating Colombo et al. (2020), there was no interaction between Prime Type and Letter Type, $\chi^2 = 2.40$, $p = .302$, indicating equivalent TL priming effects for CC, CV, and VC transpositions. Age did not modify this pattern, as there was no interaction between Prime Type, Letter Type, and Age Group, $\chi^2 = 4.80$, $p = .570$.

Z-RTs. There was a main effect of Prime Type, $\chi^2 = 308.26$, $p < .001$, indicating a regular TL priming effect (TL primes produce faster latencies than SL primes). Prime Type also interacted with Age Group, $\chi^2 = 42.85$, $p < .001$. Follow-up tests comparing TL priming effects for each pair of age groups revealed that, with one exception, each age group differed significantly from all other age groups, all *p*s < .014, with second graders producing the smallest effect, followed by third graders, fourth-and-fifth graders, and adults. The exception is the comparison between third graders and fourth-and-fifth graders, which was not significant, $\beta = .02$, $SE = .04$, $z = .34$, $p = .732$. Overall, unlike in raw RTs, there was a trend for priming effects to *increase* with age in z-RTs. Note that, albeit smaller, the effect of Prime Type still was significant for second graders when analyzed separately, $\beta = .14$, $SE = .03$, $z = 4.68$, $p < .001$. This pattern of results is represented in the right panel of Figure 1.

In this case as well, no other effect or interaction was significant (all *p*s > .1). In particular, replicating Colombo et al. (2020), there was no interaction between Prime Type and Letter Type, $\chi^2 = 3.93$, $p = .140$, suggesting that TL priming effects for CC, CV, and VC transpositions were equivalent. Age did not modify this pattern, as there was no interaction between Prime Type, Letter Type, and Age Group, $\chi^2 = 7.30$, $p = .294$.

Figure 1

TL priming effects as a function of age group in RTs (left panel) and z-RTs (right panel) in Experiment 1



Note: The solid line and the dashed line represent latencies for SL and TL primes, respectively. The error bars represent one standard error above and below the mean.

Error rates. The model produced unusually high standard errors, compromising the interpretability of tests for statistical significance. Inspection of the data revealed that the reason for this behavior was that the outcome variable completely separated certain combinations of the predictor variables. For example, no errors were committed by adults with CC transpositions in the TL condition in the fifth counterbalancing list. However, separation did not occur when the counterbalancing factor was removed. Therefore, we report the results from the model without that factor.

There was a main effect of Prime Type, $\chi^2 = 28.95$, $p < .001$, indicating that responses to TL primes were more accurate than responses to SL primes. There was also a main effect of Age Group, $\chi^2 = 29.43$, $p < .001$. Follow-up tests indicated that both second and third graders committed significantly more errors than fourth-and-fifth graders and adults, all *ps* < .02. Second and third graders did not differ from each other, nor did fourth-and-fifth graders and adults, both *p*s > .9. That is, following the third grade, error rates decreased to the level of adult performance.

There was also an Age Group X Prime Type interaction, $\chi^2 = 21.42$, $p < .001$. Follow-up tests comparing TL priming effects for each pair of age groups revealed that adults produced a significantly larger effect of Prime Type than all children groups, all *p*s < .002. No differences in the size of the Prime Type effect were found among the children groups, all *p*s > .3. Further, although, numerically, all age groups produced more errors with SL than TL primes, separate analyses for each age group revealed that adults and third graders were the only groups that produced a significant Prime Type effect, $\beta = -.99$, $SE = .16$, $z = -6.18$, $p < .001$ and $\beta = -.31$, $SE = .15$, $z = -2.15$, $p = .032$, respectively (other *p*s > .25).

There was also an interaction between Letter Type and Prime Type, $\chi^2 = 6.94$, $p = .031$. Follow-up tests comparing TL priming effects for each pair of letter types revealed that CC primes produced a larger effect of Prime Type than both CV primes, $\beta = -.36$, $SE = .16$, $z = -2.06$, $p = .040$, and VC primes, $\beta = -.44$, $SE = .18$, $z = -2.45$, $p = .014$, whereas the effect of Prime Type did not differ for CV vs. VC primes, $\beta = -.08$, $SE = .18$, $z = -.43$, $p = .665$. Further, separate analyses for

CC, CV, and VC primes revealed a significant effect of Prime Type for CC primes, β = -.66, *SE* = .13, *z* = -5.26, *p* < .001, and CV primes, β = -.30, *SE* = .12, *z* = -2.39, *p* = .017, and a borderline effect for VC primes, β = -.22, *SE* = .13, *z* = -1.71, *p* = .088. Note that this pattern was not modulated by Age Group as the three-way interaction between Prime Type, Letter Type, and Age Group was not significant, χ^2 = 8.99, *p* = .174. However, when adults were analyzed separately, there was no evidence that TL priming was affected by Letter Type, *F* = .54, *p* = .581. This null result replicates the general pattern reported by Colombo et al. (2020) for adults and suggests that the bulk of the somewhat small consonant-vowel differences reported here comes from the younger participants.

Discussion.

The results of Experiment 1 showed a clear effect of TL priming. This effect interacted with age group in all dependent variables, although the pattern of the interaction was not quite the same for all variables. In error rates and especially in z-RTs, there was tendency for the TL priming effect to increase from second grade to adulthood. In raw RTs, however, there was no evidence for the TL priming effect to increase with age. In fact, in raw RTs, the priming effect was *smaller* in adults compared to third graders, although there was no clear trend for a decrease with age either. These apparently contrasting results may depend on the large differences in average latencies among the groups. In particular, mean RTs for adults were approximately one-third the size of those of second graders (585 ms and 1597 ms, respectively). These large differences suggest caution when interpreting interactions with groups based on raw data (Faust et al., 1999).

The view that the use of raw data may be problematic, however, is not uncontroversial. In this regard, Gomez and Perea (2020) recently argued that raw RTs may be the most appropriate dependent variable to use in masked priming experiments, as they showed that, in children as well as in adults, identity priming appears to produce a shift in the response time distribution, as opposed to a change in its shape. This shift, which is typically around 50 ms in size in both children and adults, would index an advantage in the encoding process of the target given by the (typically, 50-ms) head start afforded by the related prime (see also Forster, 1998). Within such a model, transforming latencies by, e.g., scaling them by a measure of their variability, would be inappropriate, as doing so may cause spurious interactions to emerge (e.g., smaller priming effects for children, who have higher variability in their latencies compared to adults). It must be noted, however, that Gomez and Perea's (2020) claims are based on an analysis of two groups (second and fourth graders) with very similar mean latencies (1093 ms and 908 ms, respectively). Such small group differences typically would not require transformations in any case, as they would suggest similar processing abilities. The situation is different when comparing children and adults, however, as there is evidence that, in lexical decision, both evidence accumulation and decision criteria are qualitatively different in children and adults (Ratcliff, Love, Thompson & Opfer, 2012).

More generally, it has been argued (Faust et al., 1999) that it is important that absolute differences in RTs among groups differing in processing speed are removed in order to evaluate the impact of the experimental manipulation across the groups. For example, it is likely that encoding of the target takes more time for children than it does in adults (e.g., Spironelli & Angrilli, 2009). Thus, a 50-ms priming effect would not be as large an advantage for children as it is for adults, for which encoding of the target is a much faster process. That effect may in fact be inflated in raw latencies because children, especially younger children, are generally slow. Indeed, it has been argued that it is the failure to correct for differences in general performance across groups (i.e., using raw latencies) that could cause spurious interactions to emerge (in particular, larger effects for children than for adults, Zoccolotti et al., 2008). Based on this line of argumentation, z-transformations have recently become standard practice in developmental and aging research, including research addressing the development of TL priming effects (Colombo et al., 2019; Hasenäcker & Schroeder, 2021; Kezilas et al., 2017; Ziegler et al., 2014; but see Comesaña et al., 2016).

That is not to say that, if a transformation seems to be appropriate in order to account for latency differences across groups, this particular transformation (i.e.,, calculating z-scores) is necessarily the optimal transformation. Therefore, overall, it would seem appropriate that both raw and z-transformed latencies are considered when analyzing data in this context (Faust et al., 1999). Considering raw RTs, the data certainly did not show as compelling a picture as did Colombo et al., (2019) of larger TL priming effects with age and, therefore, a developing trend of a coarse-grained orthographic mechanism. However, as noted, such was not the case for z-RTs (in our opinion, for the reasons noted, a more appropriate dependent variable for latencies in the present situation) or the error rates, for which no transformations are typically performed.

More centrally, the data did not show any substantial effects of the type of letters involved in the transpositions, nor any interaction of this variable with age in the pattern of latencies.[3] Only in the accuracy data were some differences significant: In particular, there was robust priming for CC transpositions only. This letter type effect on TL priming did not change with age, however, it was not significant in adults when analyzed separately, replicating Colombo et al. (2020). This general pattern for a larger TL priming for CC transpositions compared to CV and VC transpositions, in any case, is the only hint that the inclusion of vowels in a transposition produces priming effects of different sizes.

Experiment 2.

Another way to investigate phonological influences in TL masked priming is to compare words that contain two adjacent vowels which have a different phonological nature. In Italian, adjacent vowel pairs form either a diphthong (biondo, ‘blonde’) or a hiatus (duello, ‘duel’) based on whether the two vowels are part of the same syllable (e.g., bion-do) or different syllables (e.g., du-el-lo), respectively (similar examples in English would be “naive”, where the “ai” is a hiatus, and “main”, where the “ai” is a diphthong).[4] Phonologically, these clusters can be classified as diphthongs or hiatuses depending on whether the two vowels do or do not belong to the same syllable. Specifically, if the vowel cluster includes either an “i” or a “u” which are pronounced as glides, then it is a diphthong, with both vowels belonging to the same syllable; otherwise it is a hiatus, with the two vowels belonging to different syllables. In Italian, diphthongs are more frequent than hiatuses (Marotta, 1987). For example, out of the 120,000 word forms in the Q2Stress database (Spinelli, Sulpizio, & Burani, 2017), 20,743 contain at least one diphthong whereas only 6,984 contain at least one hiatus. The expectation one could derive from this pattern is that diphthong words may be easier to process than hiatus words. In practice, however, diphthong and hiatus words may not be processed differently, at least in adults, although there is some evidence from Italian (Chetail, Scaltritti, & Content, 2014) that adults find a hiatus word more difficult than a control word when the task is to decide whether a word has two or more syllables. For example, when words with a hiatus, like “teatro” (theater), were presented, participants tended to underestimate the number of syllables, compared to words like “agosto” (August). According to Chetail et al., this result reflects readers’ sensitivity to the CV structure that constrains the parsing of a letter string into smaller units. In words like “teatro” there are two sequences of contiguous vowels, or vowel clusters (“ea” and “o”), that are mapped onto three syllables (te-a-tro), while in “agosto” the number of vowel clusters (“a”, “o”, and “o”) matches the number of syllables (a-go-sto). These considerations led Chetail et al. to claim that the CV structure is represented orthographically.

The contrast between hiatuses and diphthongs, however, is not a vowel cluster contrast because both hiatuses and diphthongs correspond to a single vowel cluster orthographically; where they differ is only at the phonological level. Thus, although hiatus words (duello) and diphthong words (biondo) have identical CV structures (and hence, the same number of vowel clusters) they may still differ phonologically. Because hiatuses and diphthongs do not differ orthographically in terms of their CV structure, if in skilled reading an orthographically based representation is used to drive processing in a lexical decision task, we would not expect any difference in priming effects between words with the two types of vowel clusters in adults. The early stages of reading development, on the other hand, are likely to reflect an involvement of phonology. Thus, whether or not priming will occur in young children and whether there would be a diphthong-hiatus difference in any age groups are empirical questions.

Summarizing, if Chetail et al. (2014) are correct in assuming that words are parsed orthographically based on vowel clusters, no difference is expected in masked TL priming effects between hiatus and diphthong words in adults. However, a difference may be possible in young children based on their strong reliance on phonology, even though, admittedly, prior studies (Colombo et al., 2019; Ziegler et al., 2014) showed that TL priming effects are small (if they appear at all) in younger children.

In both Colombo et al., (2020) and the present Experiment 1 there were no real differences in TL priming when both vowels and consonants were involved in the transposition/substitution in either (Italian) adults or children in the pattern of reaction times, although a small difference in priming effect sizes (1.3%) was found in the pattern of errors in Experiment 1. In Experiment 2, we attempted to examine the question of whether the (relatively small) priming effects we observed for the youngest age group might have been phonologically based. To do so, we used target words in which vowel clusters were present and we transposed the vowels within those clusters.

Perea and Acha (2009), consistent with our view, also suggested that a possible explanation of any dissociation between consonants and vowels in TL priming in the lexical decision task may be the involvement of phonology. Although they did not find a difference in the magnitude of the TL priming effect for vowels forming a diphthong vs. a hiatus in Spanish adult readers, such a difference may however emerge in children. Therefore, in this experiment, both children and adults were tested, because some effect might be apparent during reading acquisition and disappear in older readers (i.e., the younger readers may show a diphthong-hiatus difference whereas the adult readers may not). An identity prime (i.e., a prime that was the same as the target) was also included in order to examine to what extent TL primes produce less priming than that in the identity condition (following the logic of Kezilas et al., 2017).

Method.

Participants. Seventy-nine elementary school children participated in this experiment. Ten participants were excluded from the analyses because they committed an error on more than 40% of the trials in the task (2) and/or because they were recent immigrants to Italy (8). Out of the 69 remaining students, 24 were second graders (mean age 7.23, range 7-8, 10 females), 19 were third graders (mean age 8.16, range 7-9, 11 females), 15 were fourth graders (mean age 9.20, range 9-10, 9 females), and 13 were fifth graders (mean age 10.46, range 10-12, 2 females). For all the students, parents and the school director gave informed consent. All the students had typical development (as reported by their teachers) and each performed a partial version of the standardized test of word reading (Sartori et al., 1995) prior to performing the experiment. Children were tested during the second part of the academic year in a quiet room at their school. A further sample of 54 adults, university students (mean age 22.91, range 20-29, 28 females) was also tested to serve as a comparison with the children. Adult participants gave informed consent. All participants were native speakers of Italian and had normal or corrected-to-normal vision. Although we were unable to recruit as many participants for this experiment as we did for Experiment 1, the sample size for each age group (after fourth and fifth graders were aggregated) again exceeded the target sample size suggested by the power analysis discussed above for Experiment 1 (i.e., $n = 13$).

Materials. Eighty-four Italian words were selected from Lessico Elementare (Marconi et al., 1993, as in Spinelli et al., 2017) to serve as target words, of which 42 had at least one diphthong and 42 had at least one hiatus. The diphthong and hiatus words were matched on the most relevant variables (see Table 2).[5] For each target word, three prime types were created, an identity (ID) prime, a TL prime where the two vowels forming either a diphthong or a hiatus were transposed, and an SL prime where the same two vowels were substituted with two other vowels. As in Experiment 1, transpositions never involved the first or the last letter of the target words.

Eighty-four orthographically legal nonwords were created to serve as nonword targets.

These nonwords were created by modifying the target words while maintaining intact their CV structure and the vowel cluster used for the transposition/substitution (e.g., the word PAZIENZA became the nonword TOFIENGO). As a result, 42 of the nonwords (those derived from the diphthong words) had at least one vowel cluster which would be pronounced as a diphthong and 42 of the nonwords (those derived from the hiatus words) had at least one hiatus which would be pronounced as a hiatus. They were paired with ID, TL, and SL transposition and substitution primes created in the same way as was done in the parallel conditions for the word targets. Both diphthong and hiatus word targets were divided into three sets of 14 and each set was primed by primes from one of the three prime conditions for a given participant. Three lists were created to complete the counterbalancing for the word targets, and participants were randomly (and equally) assigned to one of those lists. There was only one list of stimuli involving the nonword targets.

Table 2

Mean stimulus characteristics (with range in parentheses) for diphthong and hiatus word targets and their transposition primes in Experiment 2

		Vowel cluster type		
Stimulus	Variable	Diphthong	Hiatus	Significance level
Baseword	Frequency (per million)	53.83 (2.45-225.60)	56.28 (0-480.63)	*p* = .885
	Length (in letters)	6.88 (5-9)	6.60 (4-10)	*p* = .306
	Orthographic neighborhood size	2.33 (0-5)	2.10 (0-5)	*p* = .435
	Bigram frequency	9.52 (8.67-10.09)	9.39 (8.15-10.21)	*p* = .101
TL prime	Position of transposition	2.88 (2-5)	2.62 (2-5)	*p* = .229
	Orthographic neighborhood size	.31 (0-4)	.43 (0-4)	*p* = .530
	Bigram frequency	9.07 (7.32-9.78)	9.26 (7.76-10.12)	*p* = .108
SL prime	Position of transposition	2.88 (2-5)	2.62 (2-5)	*p* = .229
	Orthographic neighborhood size	.26 (0-3)	.50 (0-3)	*p* = .152
	Bigram frequency	9.05 (7.34-10.03)	9.21 (8.18-10.02)	*p* = .143

Note: TL = transposed letter prime; SL = substituted letter prime. Frequency and bigram frequency were derived from Lessico Elementare (adjusted to 1 million; Marconi et al., 1993) whereas orthographic neighborhood size was derived from COLFIS (Bertinetto et al., 2005).

Results.

No items were excluded from the analyses because error rates were below 40% for all items (as noted above, the same criterion was applied to exclude participants and it resulted in the exclusion of 2 participants). Incorrect responses were removed from the latency analyses. For both latency and error analyses, correct response times faster than 300 ms were removed as outliers for both children and adults (.8% of the observations). Correct response times more than 3 SDs (for children) or more than 2.5 SDs (for adults) from each participant's mean (for correct responses to the words) were also removed as outliers (2.4%).

As in Experiment 1, fourth and fifth graders were aggregated because there were comparatively few participants in those grades individually, especially in the fifth grade in this case. For all analyses a 2 (Vowel Cluster Type: diphthong vs. hiatus, within-subject and between-item) X 3 (Prime Type: ID, TL vs. SL, within-subject and within-item) X 4 (Age Group: second graders vs. third graders vs. fourth-and-fifth graders vs. adults, between-subject and within-item) design was used for the fixed effects. The group/counterbalancing factor was also included as a fixed effect but effects and interactions involving that factor will not be reported. Subjects and items were considered as random effects for raw RTs and error rates, whereas items were the only random effect for z-RTs. Again, traditional by-subject and by-item ANOVAs based on condition means were also conducted in R for raw RTs, z-RTs (calculated on the basis of condition means), and errors, which are reported in the Supplementary Materials. The mean RTs, z-RTs, and error percentages based on by-subject means are presented in Table 3.

Table 3

Mean lexical decision RTs (in ms), z-transformed RTs, and percentage of errors (with the standard error of the mean in parentheses) for word targets in Experiment 2

		Prime type											
		RTs				Z-RTs				Error %			
Age group	Vowel cluster type	ID	TL	SL	TL priming	ID	TL	SL	TL priming	ID	TL	SL	TL priming
Second graders	Diphthong	1859 (98)	1854 (107)	1866 (75)	12	-.092 (.057)	-.124 (.049)	-.080 (.058)	.120	11.2 (2.5)	10.8 (1.7)	13.8 (2.0)	3.0
	Hiatus	1924 (102)	2036 (104)	1978 (105)	-58	.024 (.044)	.203 (.056)	.114 (.040)	-.256	19.8 (3.3)	14.7 (2.2)	19.5 (3.3)	4.8
Third graders	Diphthong	1492 (97)	1440 (102)	1559 (104)	119	-.099 (.057)	-.216 (.060)	-.007 (.056)	.719	8.8 (2.9)	4.1 (1.3)	6.4 (2.0)	2.3
	Hiatus	1588 (104)	1579 (94)	1630 (101)	51	.082 (.048)	.102 (.071)	.180 (.050)	.376	15.6 (4.3)	11.1 (2.3)	15.2 (3.6)	4.1
Fourth-and-fifth graders	Diphthong	1215 (63)	1219 (63)	1245 (56)	26	-.103 (.047)	-.098 (.041)	-.017 (.038)	.355	4.0 (1.3)	6.0 (1.3)	4.6 (1.3)	-1.4
	Hiatus	1267 (67)	1264 (58)	1306 (61)	42	.054 (.060)	.034 (.043)	.150 (.051)	.297	9.3 (1.6)	7.8 (1.9)	9.2 (2.1)	1.4
Adults	Diphthong	487 (12)	497 (13)	520 (12)	23	-.212 (.030)	-.096 (.031)	.177 (.037)	.778	3.3 (.8)	4.4 (.8)	6.0 (.9)	1.6
	Hiatus	498 (12)	499 (14)	530 (12)	31	-.069 (.032)	-.087 (.031)	.300 (.038)	1.149	4.1 (.9)	5.1 (.7)	7.7 (1.2)	2.6

Note: TL = transposed letter prime; SL = substituted letter prime; CC = consonant-consonant transposition; CV = consonant-vowel transposition; VC = vowel-consonant transposition.

Raw RTs. The model failed to converge. As per the recommended troubleshooting procedure (see "convergence" help page in R), we restarted the model from the apparent optimum and tried all other available optimizers in addition to the BOBYQA optimizer. Although all optimizers failed to converge, all optimizers except for the BOBYQA optimizer returned fairly similar results, suggesting that the convergence warnings were false positives. We report the results from the Nelder-Mead optimizer.

There was a main effect of Prime Type, $\chi^2 = 18.66$, $p < .001$, with follow-up tests indicating that responses following SL primes (1110 ms) were slower than both responses following TL primes (1081 ms; the TL priming effect), $\beta = 35.91$, $SE = 10.9$, $z = 3.29$, $p = .003$, and responses following ID primes (1074 ms), $\beta = -44.49$, $SE = 11.0$, $z = -4.09$, $p < .001$. However, responses were not significantly faster following ID primes than following TL primes, $\beta = -8.97$, $SE = 10.8$, $z = -.83$, $p = .682$. There was also a main effect of Vowel Cluster Type, $\chi^2 = 45.45$, $p < .001$, indicating overall faster responses to diphthong words than to hiatus words. There was a main effect of Age Group as well, $\chi^2 = 2103.47$, $p < .001$. Follow-up tests indicated that each age group differed significantly from all other age groups, all *p*s < .001, with second graders being the slowest, followed by third graders, fourth-and-fifth graders, and adults. That is, there was a trend for latencies to decrease as age increased.

Age Group and Prime Type did not interact in this experiment, $\chi^2 = 8.82$, $p = .183$. Focusing on the TL priming effect only (i.e., on the contrast between TL and SL primes), the pattern was somewhat similar to that reported for Experiment 1, with an overall larger TL priming effect for third graders (87 ms) compared to the other age groups. In this case, however, second graders showed a reverse TL priming effect (-25 ms) mainly driven by hiatus words. Overall, in any case, there was not a clear trend for priming effects to change consistently with age. This pattern of results is represented in the left panel of Figure 2.

Age Group also interacted with Vowel Cluster Type, χ^2 = 42.82, p < .001. Follow-up tests revealed that the source of this interaction was that increasing age was associated with a decrease in the size of the effect of Vowel Cluster Type (i.e., the slowdown to hiatus words compared to diphthong words). Second and third graders showed the largest effects (122 ms and 96 ms, respectively), followed by third graders (52 ms) and adults (8 ms). All comparisons between each pair of groups were significant except for the contrast between second and third graders, β = 2.62, *SE* = 32.7, z = .48, p = .936, and that between second graders and fourth-and-fifth graders, which was borderline, β = -53.60, *SE* = 28.3, z = -1.90, p = .058 (all other ps < .05). Note, further, that adults did not show a significant effect of Vowel Cluster Type when analyzed separately, β = -8.35, *SE* = 6.35, z = -1.32, p = .189, whereas all other age groups did, all ps < .001. This pattern of results is represented in the left panel of Figure 3.

Finally, Vowel Cluster Type did not interact significantly with Prime Type, χ^2 = 4.86, p = .088, although, overall, the TL priming effect was smaller for hiatus words, mainly because of the reverse TL priming effect produced by second graders for those words. However, there was not a three-way interaction between Vowel Cluster Type, Prime Type, and Age Group either, χ^2 = 8.22, p = .222.

Z-RTs. There was a main effect of Prime Type, χ^2 = 45.66, p < .001, with follow-up tests indicating that responses following SL primes were slower than responses following both TL and ID primes, β = .14, *SE* = .03, z = 5.55, p < .001 and β = -.16, *SE* = 14.8.03, z = -6.14, p < .001, respectively. Responses following ID primes, however, were not significantly faster than responses following TL primes, β = -.02, *SE* = .03, z = -.64, p = .798. There was also a main effect of Vowel Cluster Type, χ^2 = 4.77, p = .029, with faster responses to diphthong words than to hiatus words.

There was an Age Group X Prime Type interaction, χ^2 = 50.57, p < .001. To explore this interaction, follow-up tests were conducted contrasting TL and SL primes (the TL priming effect) for each pair of age groups (ID primes were not considered because the small advantage relative to

TL primes did not differ across groups, all *p*s > .1). Those tests revealed that, with two exceptions, each age group differed significantly from all other age groups, all *p*s < .05, with second graders producing the smallest TL priming effect, followed by third graders, fourth-and-fifth graders, and adults. The two exceptions related to the fourth-and-fifth graders, whose TL priming effect did not differ from either the second graders', $\beta = -.12$, $SE = .08$, $z = -1.52$, $p = .128$, or the third graders', $\beta = .05$, $SE = .08$, $z = .66$, $p = .510$ (in fact, numerically, fourth-and-fifth graders produced a smaller TL priming effect than did third graders, see the right panel of Figure 2).

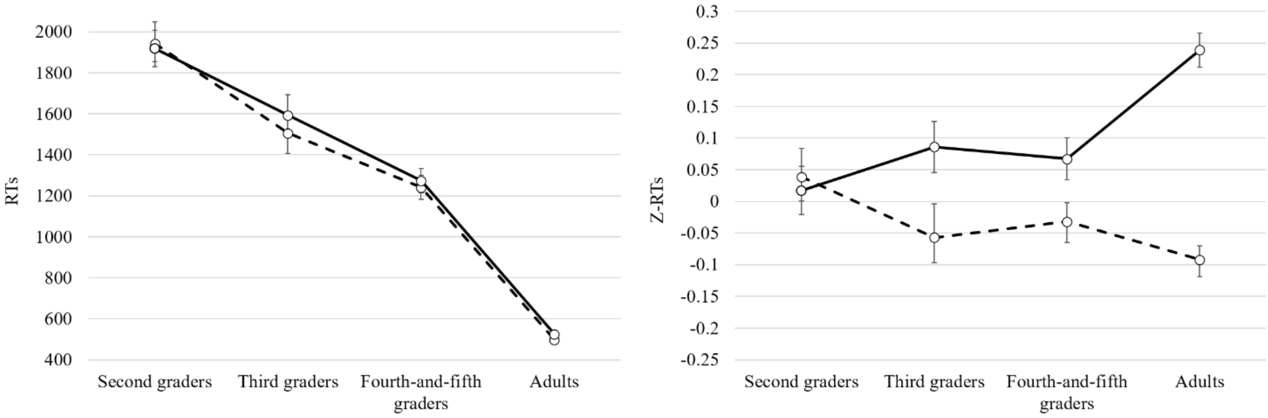
Note that, in this case, neither second graders nor fourth-and-fifth graders showed a significant effect of Prime Type when analyzed separately, $\beta = -.01$, $SE = .06$, $z = -.23$, $p = .972$ and $\beta = .10$, $SE = .05$, $z = 2.08$, $p = .093$, whereas a Prime Type effect emerged in third graders and adults, both *p*s < .05. Overall, however, similar to what was found in Experiment 1, the general tendency was for an increase in TL priming effects with age. This pattern of results is represented in the right panel in Figure 2.

Age Group was also found to interact with Vowel Cluster Type, $\chi^2 = 8.61$, $p = .035$. Follow-up tests revealed that the source of this interaction was the fact that the effect of Vowel Cluster Type (i.e., the slowdown to hiatus words compared to diphthong words) was smaller in adults compared to second graders, $\beta = -.12$, $SE = .06$, $z = -2.17$, $p = .030$, and third graders, $\beta = -.14$, $SE = .06$, $z = -2.47$, $p = .014$. No other pair of age groups differed significantly, all *p*s > .15. Further, separate analyses for each age group revealed that neither adults nor fourth-and-fifth graders showed a significant effect of Vowel Cluster Type, $\beta = -.10$, $SE = .09$, $z = -1.20$, $p = .231$ and $\beta = -.16$, $SE = .09$, $z = -1.77$, $p = .077$, respectively (for second and third graders, both *p*s < .05). Overall, similar to what was observed with raw RTs, the tendency was for a decrease in the effect of Vowel Cluster Type with age. This pattern of results is represented in the right panel of Figure 3.

Vowel Cluster Type, however, did not interact with Prime Type, $\chi^2 = .58$, $p = .749$, indicating equivalent priming effects for diphthong and hiatus words. Further, age did not modify this pattern, as there was no interaction between Vowel Cluster Type, Prime Type, and Age Group, $\chi^2 = 8.25$, $p = .220$.

Figure 2

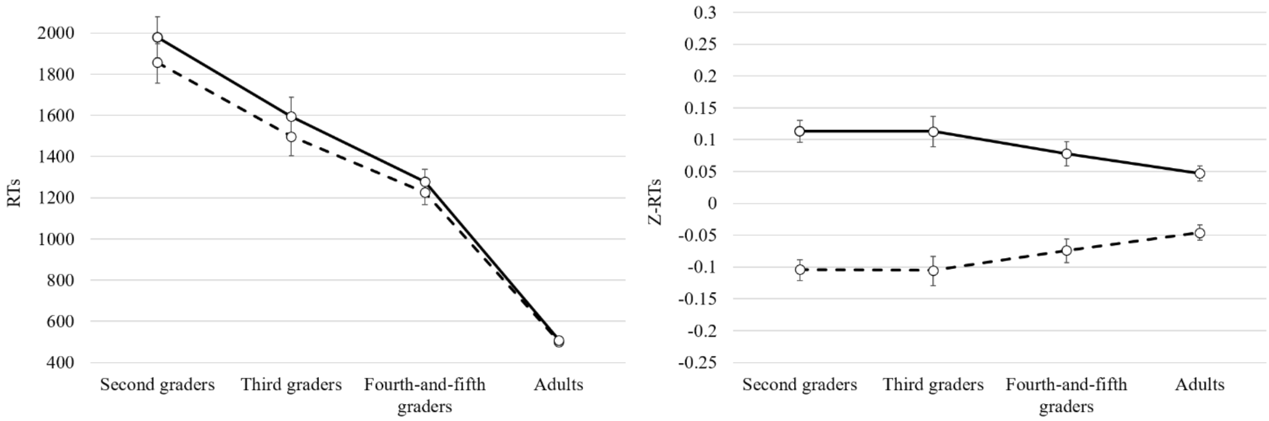
TL priming effects as a function of age group in RTs (left panel) and z-RTs (right panel) in Experiment 2



Note: The solid line and the dashed line represent z-transformed latencies for SL and TL primes, respectively. The error bars represent one standard error above and below the mean.

Figure 3

Vowel cluster type effects (hiatus vs. diphthong target words) as a function of age group in RTs (left panel) and z-RTs (right panel) in Experiment 2



Note: The solid line and the dashed line represent z-transformed latencies for hiatus and diphthong words, respectively. The error bars represent one standard error above and below the mean.

Error rates. There was a main effect of Prime Type, $\chi^2 = 8.97$, $p = .011$, with follow-up tests indicating that responses following TL primes (7.1% error rate) were more accurate than responses following SL primes (9.2%; the TL priming effect), $\beta = -.30$, $SE = .11$, $z = -2.81$, $p = .014$. There were no differences between ID primes (7.8%) and SL primes, $\beta = .23$, $SE = .11$, $z = 2.20$, $p = .072$, or between ID and TL primes, $\beta = -.06$, $SE = .11$, $z = -.57$, $p = .834$.

There was also a main effect of Vowel Cluster Type, $\chi^2 = 14.52$, $p < .001$, indicating higher error rates for hiatus than diphthong words. Age Group also produced a significant effect, $\chi^2 = 42.75$, $p < .001$. Follow-up tests indicated that second graders produced the most errors, followed by third graders, fourth-and-fifth graders, and adults. Fourth-and-fifth graders did not differ statistically from either second graders, $\beta = -.34$, $SE = .26$, $z = -1.34$, $p = .540$, or adults, $\beta = .25$, $SE = .21$, $z = 1.21$, $p = .619$, and adults differed from third graders only marginally, $\beta = -.60$, $SE = .23$, $z = -2.57$, $p = .051$. The other comparisons were significant (all *p*s < .05), however, supporting a general tendency for errors to decrease with age.

There was also an Age Group X Vowel Cluster Type interaction, $\chi^2 = 10.09$, $p = .018$. Follow-up tests comparing the effect of Vowel Cluster Type for each pair of age groups revealed that the source of this interaction was the fact that adults showed a marginally smaller effect of Vowel Cluster Type than second graders, $\beta = .35$, $SE = .21$, $z = 1.70$, $p = .088$, and a significantly smaller effect than third graders, $\beta = .76$, $SE = .26$, $z = 2.96$, $p = .003$, and fourth-and-fifth graders, $\beta = -.50$, $SE = .24$, $z = -2.11$, $p = .035$ (no other comparison was significant, all *p*s > .10). Indeed, when adults were analyzed separately, the effect of Vowel Cluster Type did not reach significance, $\beta = .20$, $SE = .20$, $z = 1.02$, $p = .306$, whereas it did for all other age groups (all *p*s < .005). .

There was also a marginal interaction between Age Group and Prime Type, $\chi^2 = 11.67$, $p = .070$. Follow-up analyses revealed that this marginal interaction mainly derived from priming effects differing between adults and third graders, with third graders showing higher error rates for

ID than TL primes and adults showing the reverse. Finally, Vowel Cluster did not interact with Prime Type, $\chi^2 = .67$, $p = .714$, nor was there a three-way interaction between Prime Type, Vowel Cluster Type, and Age Group, $\chi^2 = 3.14$, $p = .791$.

Discussion.

The data of Experiment 2 showed a significant TL priming effect regardless of whether the transposed letters were diphthongs or hiatuses. Further, diphthongs and hiatuses produced equivalent TL priming effects in all age groups. These data are somewhat similar to those of Perea and Acha (2009) who reported no differences in TL priming effects for diphthongs vs hiatuses in Spanish adult readers, although, in that case, neither diphthongs nor hiatuses produced a significant TL priming effect.

The TL priming effect also tended to increase with age in z-RTs, however, this pattern did not emerge in error rates nor in raw RTs. Therefore, as noted for Experiment 1, some caution should be exercised when interpreting this interaction. The main conclusion to draw here, however, is that TL priming is effective even with vowels transpositions, at least when they are adjacent. There was also a significant type of vowel cluster effect, which tended to decrease with age, being largest in the youngest children and smallest in adults. Words with hiatuses were more difficult to accept as words than words with diphthongs, especially for young children. As noted, the difference between diphthongs and hiatuses is phonological, more precisely at the level of syllable classification, with the vowels of the hiatus words belonging to different syllables, while the vowels of the diphthong belong to the same syllable. Thus, this result suggests that the phonological level was involved in processing the target. Importantly, this pattern emerged in all dependent variables, dispelling any potential suspicion that the data pattern observed may have been affected by the choice of the dependent variable.

One aspect to note is that the Vowel Cluster Type by Prime Type interaction was not significant, which, according to additive logic would imply that the two effects took place at different levels of processing. One clear possibility is that the TL priming effect reflects processing at the orthographic level, whereas the diphthong-hiatus difference may emerge during a later process, one involving the activation of phonology. Indeed, because the difference between diphthongs and hiatuses may only become clear once a phonological code is being produced, the relevance of this difference to the pattern of masked priming effects may be virtually zero.

Our manipulation of primes included an identity prime in Experiment 2, with no significant difference being found between TL and identity priming, indicating that there was virtually no cost for the TL manipulation for any of the participant groups. The fact that there was no cost for TL primes in our experiment with Italian readers, in our view, supports the idea that masked priming reflects an orthographic coding process in which letter identity is rapidly determined, but letter position is not. Kezilas et al. (2017), in contrast, found a significant cost for TL primes compared to identity primes, with an increasing developmental trend. In particular, English speaking early primary children did not show a significant difference between identity and TL primes, and the TL priming effect measured against a SL control was significant. In our study, in contrast, the youngest children did not show either TL priming or TL cost. Older children and adults showed significant TL priming only.

There may be several reasons for these discrepancies, some of which are methodological. The stimuli they used were shorter (on average, 5.5 letters) and higher in frequency. In addition, for transpositions they used consonants, and some of their trials involved non-adjacent ones (20%). Our selection of stimuli was constrained by the limited number of diphthongs and hiatuses, and only involved adjacent vowels. Moreover, the constant TL priming effect across ages they obtained might have been due to the limited size of TL priming effects across the age groups (around 10 ms, except for one group), which may have been due to their use of a standard masking priming paradigm rather than the sandwich priming paradigm. Note that also Hasenäcker and Schroeder

(2021) failed to find a clear increasing cost for the TL primes compared to the ID primes. In their longitudinal study, the only significant cost for TL primes was found for fourth graders.

General Discussion.

In two experiments we investigated whether the consonant/vowel status of letters involved in letter transpositions affects TL priming in a masked priming lexical decision task and whether it does so differently for children versus adults. Experiment 1, in which transposed and substituted letters were CC, CV and VC clusters, showed clear TL priming effects for all types of clusters. In z-RTs and error rates (but not in raw RTs), there was also a developmental trend for an increase in those effects across ages (i.e., little priming in the youngest children and larger effects in older children and adults). Consonant-vowel differences were not found in the pattern of latency effects. Only in errors was there a slight advantage for CC transpositions compared to CV/VC transpositions, an advantage that was mainly driven by the children participants. Thus, these results substantially confirm previous results obtained by Colombo et al., (2020), where the consonant/vowel nature of the letters involved in the transposition/substitution in the masked prime did not influence the TL priming effect in their lexical decision task.

In Experiment 2, where vowel clusters were used, there still were clear TL priming effects, with z-RTs (but not raw RTs or error rates) showing a developmental increase in those effects. In addition, there was also a robust effect of the type of vowel cluster, with diphthong words being easier than hiatus words, an effect that decreased with age. As noted, this vowel cluster difference should be considered as being phonological in nature, thereby suggesting that phonology was involved at some point in the processing sequence. The reason is that diphthong and hiatus words do not differ at the orthographic level, whereas at the phonological level, all else being equal, words including a hiatus are generally composed of an additional syllable than words including a diphthong. The key point to notice, however, is that TL priming and vowel cluster type did not interact. This result indicates that the two effects originated at different levels. And because the two developmental patterns would seem to go, if anything, in opposite directions (TL priming increasing with age, at least in z-RTs, see Figure 2, and vowel cluster type decreasing with age, see Figure 3), the overall pattern indicates that as age (and with it, presumable reading experience) increases, so does the involvement of orthography, while the opposite occurs for phonology. The juxtaposition of these two patterns further indicates that phonology likely does not play a role in TL priming, at least in more advanced readers.

Supporting these ideas are Carreiras et al.'s (2007) results from an experiment involving the recording of evoked potentials while participants made lexical decisions to nonwords derived from real words by transposing/substituting either consonants (relovution) or vowels (revulotion). The two transpositions conditions showed similar temporal patterns in the early stages (300-500 ms) and started differentiating only in later stages, after 600 ms. This result is consistent with the idea that the vowel cluster type effect appearing in our Experiment 2 was phonological, because it is at the phonological level that differences between vowel clusters appear and become relevant. It is indeed at this level that the attribution of the two vowels to the same or successive syllables is required. The results of Experiment 1, in contrast, suggest that the identity of letters is determined at the orthographic level, the same level at which TL priming effects occur.

As noted above, Comesaña et al. (2016) have, in fact, argued that any consonant-vowel dissociation in TL priming effects may depend on the degree of phonological involvement (see also Perea & Lupker, 2004). They found equal size priming effects for nonadjacent consonant (vecolidad-VELOCIDAD, 23 ms) and vowel (velicodad-VELOCIDAD, 21 ms) transpositions in children, while they found a smaller size effect for vowels than consonants in adults. As their fourth-grade children did not show phonological involvement in masked priming, while adults did, priming effects of the same size for consonants and vowels for children would imply that the consonant/vowel status was not processed at the orthographic level but only at the phonological level. Consequently, any consonant-vowel differences in adults are presumed to reflect a phonological involvement during orthographic processing. Comesaña et al. ultimately concluded that when phonology is not activated rapidly (e.g., as was presumed to be the case of their fourth-grade readers), consonant and vowel asymmetries in masked TL priming effects do not appear.

The implication is that, although phonological effects can arise in masked priming, this technique presumably does not typically capture effects arising later in processing. Indeed, in our results, the essential indicator of phonological involvement in masked TL priming, i.e., consonant-vowel differences, did not emerge even though the effect of the type of vowel cluster contained in the target words did emerge, with the size of that effect decreasing with age and becoming nonsignificant in adults. These results suggest that, in children, (later arriving) phonology was involved (at all age levels investigated here) in making lexical decisions, at least to some extent. Nonetheless, there was equivalent priming for both consonant and vowel transpositions, which for, vowels, was independent of type of vowel cluster. This pattern suggests that TL priming effects originate at the orthographic level, which would be the reason why they do not typically show up in younger children, for whom phonological recoding appears to be a later process.

Developmental trends and TL priming.

Our results indicate that, when latencies were standardized (i.e., transformed to z-scores), TL priming effects follow an increasing developmental trend in Italian, confirming previous findings (Colombo et al., 2019). Assuming that z-scores are the most appropriate dependent variable to consider in this context (see our discussion on this point in the Discussion section of Experiment 1), this pattern is consistent with the hypothesis that beginning readers use phonological recoding to translate letters into their corresponding sounds and then assemble them. This process, which can unfold differently depending, for example, on instruction and language, allows for the creation of an orthographic lexicon (e.g., see the self-teaching hypothesis by Share, 1995). With an increase in reading exposure, children will then automatize the phonological recoding process, causing it to become faster and more efficient and, at the same time, they will expand the number of orthographic representations in their lexicon. What TL masked priming effects show is that the orthographic representation for any word being read does not immediately represent all relevant information. Letter identity is rapidly processed, while letter position only later becomes available, which implies that the nature of the orthographic representation is initially imprecise, lacking detailed information such as the letters' position. This idea is consistent with our findings concerning the difference between identity and TL primes, which did not differ to any real degree in our data.

On the other hand, the present research leaves open the question of whether developmental changes in TL priming effects (if at all present) reflect a general improvement in reading skills or a change in the nature of the reading process (from serial to increasingly parallel). This issue is particularly important in thinking about and making hypotheses about developmental trends, as it might help explain the large variability found in children. Some previous studies argue against reading skills being the crucial factor, as better readers generally show smaller, not larger, TL effects (inconsistent with the fact that the skilled readers in our experiments, i.e., the adults, showed larger TL priming effects, at least in z-RTs). For example, Andrews and Cho (2012) reported smaller TL priming effects for adults scoring higher in measures of reading skill. Similarly, Perea et al. (2016) reported smaller TL effects for scrabble players (i.e., individuals with excellent letter position coding abilities) than for "control" individuals. Even within developing readers, superior reading skills seem associated with smaller rather than larger TL effects (Gomez, Marcet, & Perea, 2021; Pagán, Blythe, & Liversedge, 2021). Overall, these results are consistent with the idea that more skilled readers encode letter position better (Castles et al., 2007; Perfetti, 2017; see also Perfetti & Hart, 2002) and suggest that the skill-related differences in TL priming effects are of a different nature than what we reported here (i.e., larger TL priming effects in adults, presumably the more skilled readers, than in children): Our effects have to do with the reading process changing from serial to increasingly parallel through development.

Indeed, there is good reason to argue that the nature of the difference between skilled and less skilled readers, on one hand, and adults and children, on the other, are not the same. For example, as noted, Ratcliff et al. (2012) found that longer latencies in children are mostly dependent on longer encoding times and greater decision criteria thresholds, compared to adults. Such may not be necessarily the case for less skilled adult readers. The crucial aspects to note are that: i) developing trends are hardly linear and ii) the distinction between serial (sequential) and parallel processes may not be so clear cut. Thus, not all younger children are bound to use sequential processes, and not for all words; some younger children may have already developed sight-word reading (i.e., more parallel processing) and, in contrast, older children may sometimes use sequential processing (Altani, Protopapas and Georgiou, 2018). The characteristics of the orthographic system (e.g., more or less transparent) and of the specific words being used (e.g., length, frequency) also contribute to this discrepancy between age (or grade level) and type of reading processes involved (Altani et al., 2018; Di Filippo, et al., 2006). These differences may have played some role in the unstable developmental pattern of priming effects in the present experiments.

Adjacent and non-adjacent letters transpositions.

One final point to note concerns the issue of whether transposed letters are adjacent or not. Spanish studies have mostly made use of non-adjacent transpositions and the inconsistencies with our findings (that is, the fact that we found priming effects for VV transpositions, whereas several studies in Spanish did not) might have something to do with this difference. Perea, Duñabeitia and Carreiras (2008) measured the similarity distances and the estimates of potential priming among adjacent, 1- and 2- letter distance transpositions according to three models: Solar (Davis, 1999), Seriol (Whitney (2001), open bigram (Grainger & Van Heuven, 2003) and found a decreasing level of similarity from adjacent to 2-letters distance. They also tested the different levels of similarity in a masked priming lexical decision task, and found that the priming effect indeed was larger for adjacent than non–adjacent prime-target pairs. Thus, perhaps when comparing priming effects of different size (or testing whether priming effects are present) a better manipulation would be to use conditions which are known to provide the maximal probability of producing an effect. Given that VV transpositions have been shown to give small or null priming effects in some situations (Perea & Lupker, 2004; Lupker, Perea & Davis, 2008, but see Yang & Lupker, 2020), it would be important to test the VV transposition effect with a condition in which priming effects are maximized (i.e., with adjacent transpositions). It might be interesting, therefore, to see whether our results would be replicated if adjacent transpositions were used in Spanish.

In conclusion, we have found that adjacent transpositions of consonants and vowels (i.e., CC, CV, VC and VV transpositions) all provide reasonable-size priming effects, which implies that the simple involvement of a vowel in a transposition does not significantly reduce the priming effect. We have also shown that, when adjacent vowels are transposed, the TL priming effect does not interact with the type of vowel cluster (diphthong or hiatus) that those vowels form. Note also that although a main effect of vowel cluster was found, with diphthong targets being easier to process than hiatus targets, the target factor also did not alter the priming patterns (although the target type effect did decrease with age). Thus, the status of letters, as consonant or vowels, does not appear to be processed at the orthographic level, where TL priming occurs, but only matters at a level where phonology is activated.

Footnotes

[1] The point of this theoretical position is that a new coding process, one based on orthography rather than phonology, develops with reading experience. Because Ziegler et al.'s (2014) model is based on the idea of two processes/routes, it will be used as the prototype in our discussions. However, what should be noted is that there are a number of orthographic coding models that involve the concept of imprecision in the coding process on the orthographic route with the degree of imprecision determined by a parameter or set of parameters (e.g., Adelman, 2011; Davis 2010, Gómez, Ratcliff & Perea, 2008). Therefore, our use of this particular model should not be taken as an endorsement of it over the other extant models of orthographic coding.

[2] CoLFIS (Bertinetto et al., 2005) is a database of materials mainly written by and directed to adults. It was not possible to obtain orthographic neighborhood information for Lessico Elementare (Marconi et al., 1993). Note, further, that although our transpositions (and substitutions) never involved the first letter of the target word, they sometimes involved its first syllable, disrupting its integrity (e.g., for the target word albero ('tree'), the first syllable corresponds to its first two letters, "al"; the TL prime associated with that word, ablero, did not maintain those two letters). Preventing first-syllable disruptions or controlling them would have been challenging considering the other constraints that we had. Exploratory analyses, in any case, did not suggest that first-syllable disruptions mattered in either Experiment 1 or Experiment 2, as primes that disrupted the integrity of the target word's first syllable were associated, if anything, with *faster* latencies than primes that did not do so (rather than *slower* latencies, as the literature on syllable priming effects would suggest; see, e.g., Campos, Oliveira, & Soares, 2021).

[3] Note that although any transformation is going to have some impact on any interaction involving age group, this impact was likely not particularly strong for the interaction between prime type, letter type, and age group, the interaction that we were most interested in. The reason is that the relative magnitude of the priming effects across letter type conditions generally remained similar following the z-transformation. For example, the TL priming effects in Experiment 1 for second

graders were 104 ms, 64 ms, and 14 ms for CV, VC, and CC transpositions in raw RTs, which became .220, .119, and .058 in z-RTs. Although some discrepancies between raw RTs and z-RTs did occur (e.g., for third graders in Experiment 1, CV transpositions, compared to VC transpositions, produced slightly larger TL priming effects in raw RTs (80 ms vs. 78 ms) but slightly *smaller* TL priming effects in z-RTs (.219 vs. .225)), they were all minor and were not associated with statistically different patterns of results.

[4] Note that, in Italian, a vowel pair forms neither a diphthong nor a hiatus when one of the vowels is not pronounced because it has a diacritic function (e.g., giallo, ‘yellow’) or purely graphic function (e.g., cielo, ‘sky’; the i’s are not pronounced in either example). These cases were excluded from Experiment 2.

[5] Vowel cluster type (diphthong vs. hiatus) was manipulated between-items because there are very few words containing both a diphthong and a hiatus in Italian.

References.

Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: evidence with beginning, intermediate, and adult readers. *British journal of Psychology*, *99,* 245–264. https://doi.org/10.1348/000712607x224478

Adelman, J. S. (2011). Letters in time and retinotopic space. *Psychological Review*, *118,* 570–582. https://doi.org/10.1037/a0024811

Altani, A., Protopapas, A., & Georgiou, G. (2018). Using serial and discrete digit naming to unravel word reading processes. *Frontiers in Psychology*, *9,* 524. https://doi.org/10.3389/fpsyg.2018.00524

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 152-163. https://doi.org/10.1037/a0024953

Hasenäcker, J. & Schroeder, S. (2021). Transposed and substituted letter effects across reading development: A longitudinal study. *Journal of Experimental Psychology Learning Memory and Cognition.* 10.1037/xlm0001064.

Baayen, R. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511801686

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412. http://dx.doi.org/10.1016/j.jml.2007.12.005

Balota, D., Aschenbrenner, A., & Yap, M. (2013). Additive effects of word frequency and stimulus quality: The influence of trial history and data transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1563–1571. https://doi.org/10.1037/a0032186

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48. http://dx.doi.org/10.18637/jss.v067.i01

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton, A. M. (2005). CoLFIS (Corpus e Lessico di Frequenza dell'Italiano Scritto) [Corpus and Frequency Lexicon of Written Italian]. Institute of Cognitive Sciences and Technologies, CNR, http://www.istc.cnr.it/grouppage/colfisEng

Bolker, B. (2021). *GLMM FAQ*. Retrieved from https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html

Campos, A. D., Oliveira, H. M., & Soares, A. P. (2021). Syllable effects in beginning and intermediate European-Portuguese readers: Evidence from a sandwich masked go/no-go lexical decision task. *Journal of Child Language*, *48*, 699-716. https://doi.org/10.1017/S0305000920000537

Carreiras, M., Vergara, M., & Perea, M. (2007). ERP correlates of transposed-letter similarity effects: Are consonants processed differently from vowels? *Neuroscience Letters*, *419*, 219–224. https://doi.org/10.1016/j.neulet.2007.04.053

Carreiras, M., Vergara, M., & Perea, M. (2009). ERP correlates of transposed-letter priming effects: The role of vowels versus consonants. *Psychophysiology*, *46*, 34-42. https://doi.org/10.1111/j.1469-8986.2008.00725.x

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, *97*, 165–182. https://doi.org/10.1016/j.jecp.2007.01.006

Colombo, L., Spinelli, G., & Lupker, S. J. (2020). The impact of consonant–vowel transpositions on masked priming effects in Italian and English. *Quarterly Journal of Experimental Psychology*, *73*, 183–198. https://doi.org/10.1177/1747021819867638

Colombo, L., Sulpizio, S., & Peressotti, F. (2017). Serial mechanism in transposed letters effects: A developmental study. *Journal of Experimental Child Psychology*, *161*, 46–62. https://doi.org/10.1016/j.jecp.2017.04.002

Colombo, L., Sulpizio, S., & Peressotti, F. (2019). The developmental trend of transposed letter effects in masked priming. *Journal of Experimental Child Psychology, 186,* 117-130. https://doi.org/10.1016/j.jecp.2019.05.007

Chetail, F., Scaltritti, M., & Content, A. (2014). Effect of the Consonant–Vowel Structure of Written Words in Italian. *Quarterly Journal of Experimental Psychology*, *67,* 833–842. https://doi.org/10.1080/17470218.2014.898668

Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Lawrence Erlbaum.

Comesaña, M., Soares, A. P., Marcet, A., & Perea, M. (2016). On the nature of consonant/vowel differences in letter position coding: Evidence from developing and adult readers. *British Journal of Psychology*, *107, 651–674.* https://doi.org/10.1111/bjop.12179

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*, 713–758. https://doi.org/10.1037/a0019738

Davis, C., Castles, A., & Iakovidis, E. (1998). Masked homophone and pseudohomophone priming in children and adults. *Language and Cognitive Processes*, *13*, 625–651. https://doi.org/10.1080/016909698386401

Di Filippo, G., De Luca, M., Judica, A., Spinelli, D., and Zoccolotti, P. (2006). Lexicality and stimulus length effects in Italian dyslexics: role of the overadditivity effect. *Child Neuropsychology*, *12*, 141-149. https://doi.org/10.1080/09297040500346571

Eddy, M. D., Grainger, J., Holcomb, P. J., & Gabrieli, J. D. (2016). Orthographic and phonological processing in developing readers revealed by ERPs. *Psychophysiology*, *53*, 1776–1783. https://doi.org/10.1111/psyp.12763

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*, 777–799. https://doi.org/10.1037/0033-2909.125.6.777

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *The Quarterly Journal of Experimental Psychology*, *39A*, 211–251. https://doi.org/10.1080/14640748708401785

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression.* Thousand Oaks, CA: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Gomez, P., Marcet, A., & Perea, M. (2021). Are better young readers more likely to confuse their mother with their mohter?. *Quarterly Journal of Experimental Psychology*, *74*, 1542-1552. https://doi.org/10.1177/17470218211012960

Gomez, P., & Perea, M. (2020). Masked identity priming reflects an encoding advantage in developing readers. *Journal of Experimental Child Psychology*, *199*, 104911. https://doi.org/10.1016/j.jecp.2020.104911

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*, 577–600. https://doi.org/10.1037/a0012667

Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition*, *123*, 280–292. https://doi.org/10.1016/j.cognition.2012.01.003

Grainger, J., & van Heuven, W. J. B. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *The mental lexicon* (pp. 1–23). New York, NY: Nova Science.

Hasenäcker, J., & Schroeder, S. (2021). Transposed and substituted letter effects across reading development: A longitudinal study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0001064

Humphreys, G. W., Evett, L. J., & Quinlan, P. T. (1990). Orthographic processing in visual word identification. *Cognitive Psychology*, *22*, 517–560. https://doi.org/10.1016/0010-0285(90)90012-s

Johnson, R. L. (2007). The flexibility of letter coding: Nonadjacent letter transposition effects in the parafovea. In R. Van Gompel, M. Fisher, W. Murray & R. L. Hill (Eds.), Eye movements: A window on mind and brain (pp. 425–440). Oxford, England: Elsevier. https://doi.org/10.1016/b978-008044980-7/50021-5

Jordan, T. R., Thomas, S. M., Patching, G. R., & Scott-Brown, K. C. (2003). Assessing the importance of letter pairs in initial, exterior, and interior positions in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 883–893. https://doi.org/10.1037/0278-7393.29.5.883

Kezilas, Y., McKague, M., Kohnen, S., Badcock, N. A., & Castles, A. (2017). Disentangling the developmental trajectories of letter position and letter identity coding using masked priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 250–254. https://doi.org/10.1037/xlm0000293

Lenth, R. V. (2018). *emmeans: Estimated marginal means, aka least-squares means*. Retrieved from https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Levy, R. (2014). *Using R formulae to test for main effects in the presence of higher order interactions*. Retrieved from https://arxiv.org/abs/1405.2094

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. http://dx.doi.org/10.3389/fpsyg.2015.01171

Lupker, S. J., & Davis, C. J. (2009). Sandwich priming: A method for overcoming the limitations of masked priming by reducing lexical competitor effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 618–639. https://doi.org/10.1037/a0015278

Lupker, S. J., Perea, M., & Davis, C. J. (2008). Transposed letter priming effects: Consonants, vowels and letter frequency. *Language and Cognitive Processes*, *23*, 93–116. https://doi.org/10.1080/ 01690960701579715

Marconi, L., Ott, M., Pesenti, E., Ratti, D., & Tavella, M. (1994). *Lessico elementare: Dati statistici sull'italiano letto e scritto dai bambini delle elementari*. Bologna, Italy: Zanichelli.

Marotta, G. (1987). Dittongo e iato: una difficile discriminazione. *Annali della Scuola Normale*

Superiore di Pisa, *17*, 847-887.

Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General*, *137*, 433–455. https://doi.org/10.1037/a0012799

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review*, *119*, 517–545. https://doi.org/10.1037/a0028450

Pagán, A., Blythe, H. I., & Liversedge, S. P. (2021). The influence of children's reading ability on initial letter position encoding during a reading-like task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0000989

Paterson, K. B., Read, J., McGowan, V. A., & Jordan, T. R. (2015). Children and adults both see 'pirates' in 'parties': Letter-position effects for developing readers and skilled adult readers. *Developmental Science*, *18*, 335–343. https://doi.org/10.1111/desc.12222

Perea, M., & Acha, J. (2009). Does letter position coding depend on consonant/vowel status? Evidence with the masked priming technique. *Acta Psychologica*, 130, 127–137.

Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). Transposed-letter priming effects for close versus distant transpositions. *Experimental Psychology, 55*, 384–393. https://doi.org/10.1027/1618-3169.55.6.384

Perea, M., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: State of the art* (pp. 97–120). Hove, UK: Psychology Press.

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language, 51*, 231–246. https://doi.org/10.1016/j.jml.2004.05.005

Perea, M., Marcet, A., & Gómez, P. (2016). How do Scrabble players encode letter position during reading?. *Psicothema*, *28*, 7-12. https://doi.org/10.7334/psicothema2015.167

Perfetti, C. A. (2017). Lexical quality revisited. In E. Segers & P. van den Broek (Eds.), *Developmental perspectives in written language and literacy: In honor of Ludo Verhoeven* (pp. 51–68). Amsterdam: John Benjamins. https://doi.org/10.1075/z.206.04per

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–214). Amsterdam: John Benjamins. https://doi.org/10.1075/swll.11.14per

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 785–794. https://doi.org/10.1037/0278-7393.21.3.785

Ratcliff, R., Love, J., Thompson, C. A., & Opfer, J. E. (2012). Children are not like older adults: A diffusion model analysis of developmental changes in speeded responses. *Child Development*, *83*, 367-381. https://doi.org/10.1111/j.1467-8624.2011.01683.xSartori, G., Job, R., & Tressoldi, P. E. (1995). *The battery for evaluating dyslexia and dysorthography.* Firenze, Italy: Edizioni Os.

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

Scaltritti, M., & Balota, D. A. (2013). Are all letters really processed equally and in parallel? Further evidence of a robust first letter advantage. *Acta Psychologica*, 144, 397–410. https://doi.org/10.1016/j.actpsy.2013.07.018

Shillcock, R. C., & Monaghan, P. (2004). Reading, sublexical units, and scrambled words: Capturing the human data. In H. Bowman & C. Labiouse (Eds.), *Proceedings of the eighth neural computation and psychology workshop: Connectionist models of cognition and perception II. Progress in Neural Processing* (Vol. 15, pp. 221–230). Singapore: World Scientific. https://doi.org/10.1142/9789812702784_0023

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. Spieler & E. Schumacher (Eds.), *New methods in cognitive psychology* (pp. 4–31). New York, NY: Routledge.

Spinelli, G., Sulpizio, S., & Burani, C. (2017). Q2Stress: A database for multiple cues to stress assignment in Italian. *Behavior Research Methods, 49*, 2113–2126. https://doi.org/10.3758/s13428-016-0845-7

Spironelli, C., & Angrilli, A. (2009). Developmental aspects of automatic word processing: language lateralization of early ERP components in children, young adults and middle-aged subjects. *Biological Psychology*, *80,* 35-45. https://doi.org/10.1016/j.biopsycho.2008.01.012

Yang, H., & Lupker, S. J. (2020). A reexamination of consonant–vowel differences in masked transposed letter priming effects in the lexical decision task. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 74*, 92–110. https://doi.org/10.1037/cep0000195

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*, 221–243. https://doi.org/10.3758/bf03196158

Ziegler, J. C., Bertrand, D. Leté, B., & Grainger, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology, 50*, 1026–1036. https//doi.org/10.1037/a0035187

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: a cross-language investigation. *Psychological Science*, *21*, 551–559. https://doi.org/10.1177/0956797610363406

Ziegler, J.C. and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3-29. https://doi.org/10.1037/0033-2909.131.1.3

Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., & Martelli, M. (2008). Reading development in an orthographically regular language: effects of length, frequency, lexicality and global processing ability. *Reading and Writing, 22,* 1053-1079. https://doi.org/10.1007/s11145-008-9144-8

Appendix

Stimuli used in Experiment 1

	CC prime		CV prime		VC prime	
Target	TL	SL	TL	SL	TL	SL
ALBERO	ablero	amfero	alebro	alunro	albreo	alblao
ALFABETO	aflabeto	ardabeto	alafbeto	alelbeto	alfbaeto	alfrieto
ALLARME	allamre	allance	alalrme	aleprme	allrame	allmume
ALLEGRO	allergo	allelso	alelgro	alangro	allgero	allduro
ALTALENA	atlalena	adfalena	alatlena	aliglena	altlaena	altriena
ALTEZZA	atlezza	asdezza	aletzza	alanzza	altzeza	altniza
AMBIENTE	abmiente	asgiente	amibente	amecente	ambinete	ambidite
ANGELO	agnelo	aspelo	aneglo	anidlo	angleo	angrao
ANGOLO	agnolo	aldolo	anoglo	aneplo	angloo	angneo
ANTICO	atnico	avrico	anitco	anesco	antcio	antseo
APERTO	apetro	apengo	aeprto	aodrto	apreto	apdito
ARMADIO	amradio	atgadio	aramdio	arovdio	armdaio	armpeio
BAMBINA	babmina	bavcina	bamibna	bamosna	bmabina	brebina
BAMBOLA	babmola	bapnola	bamobla	bamidla	bmabola	blibola
BASTONE	batsone	bachone	basotne	basarne	bsatone	brotone
BEVANDA	bevadna	bevacta	beavnda	beosnda	bevnada	bevrida
BILANCIA	bilacnia	bilabpia	bialncia	biurncia	bilnacia	bilducia
CALCIO	caclio	carhio	calico	calaso	clacio	cnecio
CAMPIONE	capmione	cagbione	camipone	camamone	cmapione	cnepione
CANCELLO	cacnello	cafsello	caneclIo	canabllo	cnacello	cmecello
CANZONE	caznone	cablone	canozne	canebne	cnazone	clizone
CARTELLO	catrello	cadmello	caretllo	caragllo	cratello	cpetello
CARTONE	catrone	cahmone	carotne	carebne	cratone	cmetone
CASTELLO	catsello	camnello	casetllo	casocllo	csatello	cfetello
CESTINO	cetsino	cemlino	cesitno	cesusno	csetino	ctitino
COMPITO	copmito	cochito	comipto	comusto	cmopito	csapito
CONTENTO	cotnento	codsento	conetnto	conapnto	cnotento	cpitento
COPERTA	copetra	copecpa	coeprta	coatrta	copreta	copcata
DELFINO	deflino	despino	delifno	delegno	dlefino	dtifino
DESERTO	desetro	desepso	deesrto	deoprto	desreto	deslito
DIVERSO	divesro	divedlo	dievrso	diatrso	divreso	divdoso
DOMANDA	domadna	domatfa	doamnda	doilnda	domnada	dompida
ELEFANTE	elefatne	elefagre	eleafnte	eleilnte	elefnate	elefrite
ENERGIA	enegria	enevnia	eenrgia	eodrgia	enregia	engagia
ENORME	enomre	enonfe	eonrme	eabrme	enrome	entame
ENTRATA	etnrata	ezprata	entarta	entupta	entrtaa	entrsea
ESEMPIO	esepmio	esenhio	eesmpio	eitmpio	esmepio	esfipio
ESTATE	etsate	ernate	esatte	esopte	esttae	estmie
FARFALLA	fafralla	faslalla	faraflla	faroslla	frafalla	fcofalla
FINESTRA	finetsra	finercra	fienstra	fiagstra	finsetra	finpatra
FORESTA	foretsa	forebca	foersta	foalsta	forseta	forgita

FORTUNA	fotruna	fogmuna	forutna	forelna	frotuna	fsatuna
GIARDINO	giadrino	giacgino	giaridno	giarabno	giradino	ginedino
GIGANTE	gigatne	gigaple	giagnte	giolnte	gignate	gigrite
GIORNALE	gionrale	giompale	gioranle	giorotle	gironale	gifunale
GIORNATA	gionrata	giovdata	gioranta	giorimta	gironata	gicunata
GRANDE	gradne	graghe	garnde	golnde	grnade	grgode
INCONTRO	icnontro	ibsontro	inocntro	inidntro	incnotro	incfatro
INGRESSO	ignresso	ictresso	ingersso	ingacsso	ingrseso	ingrluso
INSALATA	isnalata	iclalata	inaslata	inumlata	inslaata	insniata
INTERO	itnero	ifmero	inetro	inibro	intreo	intsio
INVERNO	ivnerno	iflerno	inevrno	inacrno	invreno	invgano
LINGUA	lignua	lifrua	linuga	lineta	lnigua	lcogua
MARTELLO	matrello	mapsello	maretllo	maribllo	mratello	mbitello
MERCATO	mecrato	mesbato	meracto	meregto	mrecato	mdacato
MERENDA	meredna	merepla	meernda	mealnda	merneda	merpada
MOMENTO	mometno	momesvo	moemnto	moicnto	momneto	mombato
MORBIDO	mobrido	motsido	moribdo	morevdo	mrobido	mpebido
OMBRELLO	obmrello	osprello	omberllo	ombitllo	ombrlelo	ombrnilo
ORDINE	odrine	opline	oridne	oratne	ordnie	ordlue
OSPEDALE	opsedale	oltedale	osepdale	osamdale	ospdeale	ospnoale
OSPITE	opsite	oncite	osipte	osonte	osptie	osproe
PANCIA	pacnia	patsia	panica	panuha	pnacia	plicia
PARTENZA	patrenza	pablenza	paretnza	parosnza	pratenza	pcotenza
PARTITA	patrita	pasvita	paritta	parosta	pratita	pbitita
PENSIERO	pesniero	petviero	penisero	penebero	pnesiero	prasiero
PERFETTO	pefretto	pepmetto	pereftto	peruvtto	perfteto	perfnato
POLVERE	povlere	pomnere	polevre	polopre	plovere	ptivere
POSTINO	potsino	pognino	positno	posapno	psotino	ptitino
PRANZO	prazno	pravfo	parnzo	pognzo	prnazo	prvizo
PRESENTE	presetne	presedre	preesnte	preafnte	presnete	preshate
PRINCIPE	pricnipe	prifripe	pirncipe	popncipe	prnicipe	prgecipe
PROBLEMA	prolbema	pronsema	porblema	pasblema	prbolema	prlalema
RISPOSTA	ripsosta	rilmosta	risopsta	risatsta	rsiposta	rneposta
RITARDO	ritadro	ritapco	riatrdo	rielrdo	ritrado	ritsedo
RITORNO	ritonro	ritombo	riotrno	riaprno	ritrono	ritpino
SCARPE	scapre	scalde	sacrpe	subrpe	scrape	scnepe
SCELTA	scetla	scegra	seclta	saplta	scleta	scqota
SEGRETO	sergeto	senseto	segerto	segimto	sgereto	sbareto
SERPENTE	seprente	semsente	serepnte	serornte	srepente	sgipente
SGUARDO	sguadro	sguanpo	sugardo	sipardo	sgurado	sgusido
SILENZIO	sileznio	silecpio	sielnzio	siamnzio	silnezio	silbozio
SINISTRA	sinitsra	sinimbra	siinstra	sietstra	sinsitra	sinpetra
SORPRESA	soprresa	sotnresa	sorpersa	sorpalsa	sropresa	sgipresa
STANCO	stacno	stavlo	satnco	solnco	stnaco	stloco
STANZA	stazna	stafpa	satnza	solnza	stnaza	stliza
STRADA	srtada	splada	starda	stovda	strdaa	strpoa
STRANO	srtano	smnano	starno	studno	strnao	strseo
STREGA	srtega	snpega	sterga	stasga	strgea	strlia
TRISTE	tritse	trighe	tirste	tunste	trsite	trmote

ULTIMO	utlimo	unfimo	ulitmo	ulormo	ultmio	ultreo
USCITA	ucsita	uptita	usicta	usanta	usctia	uschea
VACANZA	vacazna	vacacpa	vaacnza	vaernza	vacnaza	vachiza
VERDURA	vedrura	vencura	verudra	verabra	vredura	vladura
VESTITO	vetsito	vecdito	vesitto	vesabto	vsetito	vtatito
VIOLENZA	violezna	violebca	vioelnza	vioirnza	violneza	violsiza

Stimuli used in Experiment 2

Diphthong words				Hiatus words			
Target	Identity prime	TL prime	SL prime	Target	Identity prime	TL prime	SL prime
ACQUARIO	acquario	acqaurio	acqierio	ALVEARE	alveare	alvaere	alvoire
ANZIANO	anziano	anzaino	anzeuno	BAULE	baule	buale	boile
BIONDO	biondo	boindo	baundo	BEATO	beato	baeto	buito
CAMION	camion	camoin	cameun	CAOS	caos	coas	ceis
CHIUSO	chiuso	chuiso	cheaso	CEREALE	cereale	ceraele	ceriole
CILIEGIA	ciliegia	cileigia	cilougia	CLIENTE	cliente	cleinte	claonte
CURIOSITÀ	curiosità	curoisità	curaesità	CREATURA	creatura	craetura	croitura
FIAMMA	fiamma	faimma	feomma	DIARIO	diario	dairio	deurio
FIANCO	fianco	fainco	fuenco	DUELLO	duello	deullo	daillo
FIORITO	fiorito	foirito	faerito	EGOISTA	egoista	egiosta	eguasta
GENIALE	geniale	genaile	genuole	EROINA	eroina	eriona	eruena
GUANCIA	guancia	gauncia	geoncia	FARAONE	faraone	faroane	faruine
GUANTO	guanto	gaunto	giento	GEOGRAFIA	geografia	goegrafia	guagrafia
GUARDIA	guardia	gaurdia	goerdia	IDEALE	ideale	idaele	idoule
GUERRA	guerra	geurra	goirra	INFLUENZA	influenza	infleunza	infloanza
ITALIANO	italiano	italaino	italueno	INVIARE	inviare	invaire	inveore
LAMPIONE	lampione	lampoine	lampuane	LEALE	leale	laele	loule
LIQUIDO	liquido	liqiudo	liqaedo	LEONE	leone	loene	luane
LUOGO	luogo	lougo	laigo	LEOPARDO	leopardo	loepardo	liupardo
MIELE	miele	meile	maole	MAESTÀ	maestà	meastà	moistà
MILIARDO	miliardo	milairdo	miluerdo	MAESTRA	maestra	meastra	miostra
NUOTO	nuoto	nouto	neato	NEONATO	neonato	noenato	naunato
ORIENTALE	orientale	oreintale	oruantale	OCEANO	oceano	ocaeno	ociuno
PAZIENZA	pazienza	pazeinza	pazounza	PAESAGGIO	paesaggio	peasaggio	pousaggio
PIANETA	pianeta	paineta	pouneta	PAESE	paese	pease	poise
PIANURA	pianura	painura	poenura	PAURA	paura	puara	peora
PIATTO	piatto	paitto	peutto	PERIODO	periodo	peroido	pereado
PIAZZA	piazza	paizza	puozza	POETA	poeta	peota	puita
PIEGARE	piegare	peigare	puogare	PREISTORIA	preistoria	priestoria	praustoria
PINGUINO	pinguino	pingiuno	pingeano	REALE	reale	raele	roile
PIOGGIA	pioggia	poiggia	pueggia	REAME	reame	raeme	riome
POMPIERE	pompiere	pompeire	pompaure	REAZIONE	reazione	raezione	rouzione
QUADERNO	quaderno	qauderno	qioderno	RIASSUNTO	riassunto	raissunto	roessunto

QUADRATO	quadrato	qaudrato	qiedrato	SCIARE	sciare	scaire	scuore
SENTIERO	sentiero	senteiro	sentauro	SPIONE	spione	spoine	speane
SGUARDO	sguardo	sgaurdo	sgeirdo	TEATRO	teatro	taetro	tuitro
SIEPE	siepe	seipe	soape	TEORIA	teoria	toeria	tuaria
SUONARE	suonare	sounare	sienare	TRIANGOLO	triangolo	traingolo	treungolo
TIEPIDO	tiepido	teipido	taupido	TRIONFO	trionfo	troinfo	traenfo
UGUALE	uguale	ugaule	ugiole	UBRIACO	ubriaco	ubraico	ubreoco
VIETATO	vietato	veitato	vautato	VEICOLO	veicolo	viecolo	voacolo
ZAINO	zaino	ziano	zueno	VIALE	viale	vaile	voele