When orthography is not enough: The effect of lexical stress in lexical decision

Lucia Colombo · Simone Sulpizio

Abstract Three lexical decision experiments were carried out in Italian, in order to verify if stress dominance (the most frequent stress type) and consistency (the proportion and number of existent words sharing orthographic ending and stress pattern) had an effect on polysyllabic word recognition. Two factors were manipulated: whether the target word carried stress on the penultimate (dominant; "graNIta," "seNIle"—*slush*, *senile*) or on the antepenultimate (non-dominant) syllable ("MISsile," "BIbita"—*missile*, *drink*), and whether the stress neighborhood was consistent (graNIta, MISsile) or inconsistent (seNIle, BIbita) with the word's stress pattern. In Experiment 1, words were mixed with nonwords sharing the word endings, which made words and nonwords more similar to each other. In Experiment 2, words and nonwords were presented in lists blocked for stress pattern. In Experiment 3, we used a new set of nonwords, which included endings with (stress) ambiguous neighborhoods and/or with a low number of neighbors, and which were overall less similar to words. In all three experiments, there was an advantage for words with penultimate (dominant) stress and no main effect of stress neighborhood. However, the dominant stress advantage decreased in Experiments 2 and 3. Finally, in Experiment 4, the same materials used in Experiment 1 were also used in a reading-aloud task, showing a significant consistency effect but no dominant stress advantage. The influence of stress information in Italian word recognition is discussed.

Keywords Lexical stress · Lexical decision · Stress dominance · Stress consistency

L. Colombo (✉)
Department of General Psychology, University of Padua, Via
Venezia 8, 35131 Padova, Italy
e-mail: lucia.colombo@unipd.it

S. Sulpizio
Department of Psychology and Cognitive Science (DiPSCo),
University of Trento, Trento, Italy

S. Sulpizio
Fondazione Marica De Vincenzi ONLUS, Trento, Italy

There has been increasing interest in the role that lexical stress plays in the pronunciation of a word from its orthography. The issue is particularly relevant for free-stress languages such as English or Italian, in which readers must assign stress before articulation may start (e.g., Colombo, 1992; Colombo, Deguchi, Boureux, 2014; Colombo & Zevin, 2009; Perry, Ziegler, & Zorzi, 2010; Sulpizio, Arduino, Paizi, & Burani, 2013; Sulpizio, Spinelli, & Burani, 2015). Important factors related to lexical stress that have been investigated, in tasks such as reading aloud and lexical decision, are: Dominance, the most frequent stress pattern in a language (Colombo & Zevin, 2009), also known as regularity or typicality (Arciuli & Cupples, 2006; Jouravlev & Lupker, 2014; Kelly & Bock, 1988; Rastle & Coltheart, 2000; see below); orthographic correlates of stress (e.g., affixes; Rastle & Coltheart, 2000); stress neighborhood consistency (i.e., the more or less constant correspondence between spelling patterns and stress; Arciuli & Cupples, 2006; Burani & Arduino, 2004; Burani, Paizi, & Sulpizio, 2014; Colombo, 1992; Jouravlev & Lupker, 2014; Pagliuca & Monaghan, 2010; Paizi, Zoccolotti, & Burani, 2011; Seva, Monaghan & Arciuli, 2009; Sulpizio & Colombo, 2013); and grammatical class (typical stress differs for nouns and verbs in English; Arciuli & Cupples, 2006; for different grammatical classes in Russian; Jouravlev & Lupker; 2014; for a review of all the above issues, see Sulpizio, Burani, & Colombo, 2015). Among these factors, the most relevant for the present study is dominance.

The first studies focusing on the effect of the most frequent stress pattern in reading aloud showed that the emergence of this effect is influenced by word frequency. Both Monsell, Doyle, and Haggard (1989) and Brown, Lupker, and Colombo (1994) found evidence for a stress effect in naming English words and an interaction with frequency. The same pattern was also established in naming Italian words by Colombo (1992), who found that low-frequency words with the most frequent stress pattern (on the penultimate syllable of polysyllabic words, as in "bamBIno," *child*) showed an advantage compared to low-frequency words with antepenultimate stress (e.g., "TAvolo," *table*), which is much less frequent (with a proportion of about 0.2 to 0.8). However, Rastle and Coltheart (2000) found a stress effect in English only when the word contained affixes, which cued the correct stress pattern. Yet again, Burani and Arduino (2004) found no evidence for an advantage of the dominant stress in Italian, arguing that participants use only stress neighborhood consistency in reading aloud, that is, the association between certain orthographic patterns and stress (see below for a more detailed description). Finally, Colombo and Zevin (2009) also found no clear evidence for an advantage of words with dominant stress over those with non-dominant stress in a priming paradigm with a naming task.

While the majority of studies have been based on tasks involving production (reading aloud and naming), the investigation of lexical stress has been further extended to the lexical decision process. Specifically, Arciuli and Cupples (2006) investigated the effect of frequency of stress (which they called "typicality") in English with naming and lexical decision, manipulating both typicality of stress and grammatical class (nouns vs. verbs). In fact, they did not explore the most frequent stress pattern in the language overall, but the most frequent/typical pattern within each grammatical class. They found an advantage for typically stressed words (i.e., second-syllable stress for verbs, first–syllable stress for nouns) in both tasks, but only in the pattern of errors. They argued that because the same pattern was obtained in both naming and lexical decision, and because lexical decision can be carried out based on orthography alone, these results suggest that stress may be cued directly by orthography. Furthermore, through an analysis of the distribution of word endings and their correlation with stress pattern, they found that word endings were able to predict both grammatical class and lexical stress. Mundy and Carroll (2013), on the other hand, only found evidence for an effect of the consistency between orthographic endings and stress in a lexical decision task. Kelly, Morris, and Verrekia (1998) investigated the role of particular spellings in word endings, which, in their hypothesis, might signal a specific stress pattern (e.g., "-et" associated to first-syllable stress, as in *comet*, and "-ette" associated to second-syllable stress, as in *roulette*). They found supporting evidence, in both latencies and error rates, in naming and lexical decision.

Further evidence for an effect of stress in lexical decision comes from a study on Russian (Jouravlev & Lupker, 2014). In Russian disyllabic words there is no overall dominant pattern, but initial and final stress are present in different proportions in each grammatical category. Nouns and verbs have a similar proportion of initial and final stress, while adjectives are more frequently stressed on the initial syllable. Using disyllabic stimuli, Jouravlev and Lupker found a significant interaction of stress type with grammatical class: No stress type effect was found in either naming or lexical decision for nouns and verbs, while a significant advantage for initial compared to final stress adjectives was found in both naming and lexical decision.

The aim of the present study was to investigate the role of lexical stress in word recognition and whether this role depends on stress neighborhood consistency, that is, the correspondence, or lack of correspondence, between orthography and stress. These issues were investigated in Italian, a transparent language in spelling-sound correspondence but not in lexical stress. Considering stress types in Italian, penultimate syllable stress is the most frequent pattern ("bamBIno," *child*), since around 80 % of Italian polysyllables bear such stress (Thornton, Iacobini, & Burani, 1997). One of the first proposals about Italian stress representation and processing in reading (Colombo, 1992) claimed that stress can be represented both at the lexical level, as part of the phonological representation of lexical entries, and at the sublexical level, with spelling patterns of word endings directly cueing stress. Moreover, in agreement with models of production (Levelt, Roelofs, & Meyer, 1999) stress can be abstractly represented within a metrical frame specifying the stressed syllable position. This is an abstract representation because it is independent of segmental information and can be derived by statistical distributional information about the most frequent stress type, or, in some frameworks, by rule (Levelt et al., 1999).

This view is partially consistent with that proposed in a computational model, the CDP++ model for polysyllabic word reading in Italian (Perry, Ziegler, & Zorzi, 2014). In the computational perspective provided by Perry et al. (2014) prosodic information is abstractly coded in a stress buffer, connected with both the lexical and the sublexical pathways. While the lexical route would activate the correct stress pattern within the stress output nodes, the sublexical route likely reflects consistency effects. Using this framework, which includes an abstract representation of stress (the "stress nodes"), Perry et al. (2014) simulated the results of several studies on stress carried out in Italian with the reading-aloud task (e.g., Burani & Arduino, 2004; Colombo, 1992; Colombo & Zevin, 2009). In particular, they simulated the dominant stress advantage, the stress type by word frequency interaction and the stress type by stress neighborhood consistency interaction found in Colombo (1992). But they were also able to find the same pattern as Burani and Arduino

(2004), with a significant stress-neighborhood consistency effect independent of main stress. Being able to simulate most, even contrasting results, they concluded that the divergences in the different studies were due to the particular nature of the experimental materials. They also claimed that stress and consistency effects likely reflect an interaction between lexical and sublexical pathways. In fact, simulations carried out by alternatively removing lexical and sublexical phonology clearly showed that both are required to simulate the main effects in Italian. Note, however, that Perry and colleagues did not simulate lexical decision, and a prediction based on this model would be highly speculative, particularly because a clear pattern of behavioral data in lexical decision is missing.

While at least one study has found an effect of the most frequent stress in English lexical decision (Arciuli & Cupples, 2006), two studies in Italian showed conflicting results. One study (Colombo, 1992) found an effect of stress dominance on low-frequency words in the pattern of errors; the other (Burani & Arduino, 2004) did not find any effect at all. There might be different reasons for this inconsistency in results, probably grounded on the nature of the experimental stimuli (see discussion in Burani & Arduino, 2004). However, clearly, it is important to further investigate this issue and verify if a robust pattern can be established in lexical decision.

Some studies show that access to the phonological information of the stimulus may be fast and automatic (Ashby, Sanders, & Kingston, 2009; Wheat, Cornelissen, Frost, & Hansen, 2010; see Frost, 1998, for a review). Moreover, phonological variables can sometimes affect decision (Halderman, Ashby, & Perfetti, 2012), as phonological information may help word recognition by enhancing orthography-to-phonology mapping. As noted, Italian is very regular in spelling–sound correspondences, with the only exception of stress pattern, and phonological assembly may be very fast. Thus, in recognizing a word, readers are very likely to use both phonological recoding procedures and whole-word lexical processing. Supporting evidence comes from a study by Burani and Cafiero (1991) who manipulated subsyllabic and consonant-vowel structure of Italian words in lexical decision. They found that lexical decision on words with consonant–vowel structure controlled were faster and less error prone when they contained simpler and more frequent subsyllabic units, as, for example, single-consonant onsets and codas, which are more frequent in Italian compared to complex (CC) onsets and codas, suggesting a role for these phonological units in lexical access.

Concurrent evidence from other types of effects supports the idea that both lexical and sublexical phonology are equally used in Italian. Peressotti and Colombo (2012) found a pseudohomophone advantage in reading aloud compared to control nonwords. Since pseudohomophone effects are considered markers of phonological involvement (Jacobs & Grainger, 1994) they interpreted this result as evidence of an interaction in the phonemic buffer between output lexical phonology activated directly from orthography and sublexical processing. Given the evidence for Italian suggesting that phonology is automatically activated, an effect of the most frequent stress type might be expected in lexical decision. Such an effect might be driven from lexical and/or sublexical phonology, consistently with what is predicted in Perry et al. (2014). In particular, it might reflect faster access to lexical representations in the phonological lexicon, faster activation of sublexical phonology, or both.

Naming a word can be helped by its ending, as shown by the literature on neighborhood consistency effects in English. Note, however, that in such literature, the term ending was defined in different ways. First, it was used to refer to consistency between orthography and segmental phonology (in particular, rhymes) mostly in the English language and for monosyllabic words (Jared, McRae, & Seidenberg, 1990; Seidenberg & McClelland, 1989; Treiman, Mullennix, Bijelac-Babic, & Richmond-Welty, 1995) on one hand, and, on the other hand, to consistency between orthography and stress, in Italian and Russian (Burani & Arduino, 2004; Colombo, 1992; Jouravlev & Lupker, 2014). Second, it has been defined at different levels of granularity, with the extreme case of the study by Monaghan, Arciuli, and Seva (in press) who investigated the relation to stress of initial and ending units of different sizes, from one to five letters.

A particular relevance in the investigation of consistency effects has been given to the rhyme. Many studies indeed show that rhyme is an important unit in reading (e.g., Ziegler & Goswami, 2005). The nature of endings assumed to be relevant for Italian reading was defined by Colombo (1992) as the sequence formed by the nucleus of the penultimate syllable and the last syllable: the rhyme. For examples, in the word "la-VO-ro" the rhyme is "-oro," formed by the nucleus of the penultimate syllable plus the last syllable. Italian words with many stress-consistent neighbors (words sharing orthographic ending and stress pattern, also called stress *friends*) are named faster than words with many stress inconsistent neighbors (words sharing orthographic ending but with a different stress pattern, also called stress *enemies*; Arciuli & Cupples, 2006; Burani & Arduino, 2004; Colombo, 1992). For example, TRAgica (tragic) and forMIca (ant) share the orthographic ending -ica, which is mostly included in words with initial stress: TRAgica is consistent with the stress neighborhood of -ica, whereas forMIca is stress inconsistent. Note that in some studies (Colombo, 1992; Colombo & Zevin, 2009) stress neighborhood consistency effects interacted with stress dominance and were only apparent for non-dominant stress words. This result was confirmed by simulations in the connectionist computational model of Italian reading by Pagliuca and Monaghan (2010), which also showed this interaction.

The effect of word endings might extend to lexical decision, with faster latencies and more accurate responses for

stress-consistent than stress-inconsistent words. Word endings might exert their effect at the orthographic level. Support for the idea that endings may provide orthographic cues to stress comes from the connectionist computational model implemented by Arciuli, Monaghan, and Seva (2010), which did not include a phonological component and showed sensitivity to the statistical properties of the orthography. The network was trained to learn from the orthography and was required to assign initial or second-syllable stress to the disyllabic words. By identifying orthographic regularities in word endings, the network was able to assign stress to disyllabic stimuli. The authors concluded that stress information is orthographically represented.
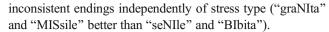
In the present study, we manipulated lexical stress and stress neighborhood consistency in lexical decision and reading aloud. We ran three lexical decision experiments in which word stimuli were maintained constant, while contextual variations were introduced. Specifically, in Experiment 1, words were presented mixed together with nonwords, most of which had a final sequence strongly associated either with dominant or non-dominant stress. In Experiment 2, we presented the same stimuli as in Experiment 1, but in two separate blocks, where stress was held constant. In Experiment 3, we again adopted the mixed stress presentation of Experiment 1, but with a new set of nonwords, which were built with weak final sequences in terms of orthography-to-stress association. The idea underlying the introduction of contextual variations in Experiments 2 and 3 was to decrease the tendency to rely on lexical consultation because of the pure lists (only one stress type; Experiment 2), and because of the greater word/ nonword dissimilarity on the basis of endings (Experiment 3).

In all the experiments, two variables were manipulated: type of stress (dominant vs. non-dominant) and stress neighborhood consistency (consistent and inconsistent; see Table 1 for examples).

If word recognition benefits from phonological activation, and if the frequency of the stress type is important in discriminating words and nonwords, we might expect an advantage for dominant stress words ("graNIta," *slush*, and "seNIle," *senile*, compared to "BIbita," *drink*, and "MISsile," *missile*), independently of the type of ending. If, instead, the consistency of endings is important and affects word recognition at an orthographic level as much as it affects word naming, we might expect an advantage for consistent compared to

inconsistent endings independently of stress type ("graNIta" and "MISsile" better than "seNIle" and "BIbita").

Moreover, we might also expect an effect of the endings on nonwords. In particular, given that nonwords with ambiguous endings or from small neighborhoods are less word-like, it might be easier to classify them as nonwords, compared to nonwords with larger and more consistent neighborhoods.

Former studies using the lexical decision task found stress dominance effects mostly in the pattern of errors (Arciuli & Cupples, 2006; Colombo, 1992), but the error rate is usually very low in lexical decision. Thus, in the present study, we carried out speeded lexical decision, with a 600 ms deadline for a response. This procedure has been found to increase the number of errors without changing the nature of the effects or processes involved (Colombo & Tabossi, 1992; Parkin, 1982).

Finally, Experiment 4 was carried out with a reading-aloud task, as a comparison to lexical decision, and to replicate former effects found in the literature.

Experiment 1

Method

Participants Thirty-four students (12 males; mean age = 23.03, SD = 1.09) from the University of Padua took part in the experiment. They were all native Italian speakers with normal or corrected-to-normal vision.

Materials Four sets each with 30 three-syllabic low-frequency words were selected from the CoLFIS database (Bertinetto et al., 2005) and were used as stimuli. The four sets were obtained by combining two experimental factors: stress dominance (words bearing dominant vs. non-dominant stress) and stress neighborhood consistency (stress-consistent neighborhood vs. stress-inconsistent neighborhood). Stimuli were matched on frequency, length in letters, orthographic neighborhood size, and summed frequency of orthographic neighbors (see Table 2). Familiarity ratings, on a 5-point scale, were also collected by a group of 15 participants who did not participate in any of the lexical decision experiments. We did not match stimuli on bigram frequency because this choice would have strongly limited the number of items in each condition; however, because bigram frequency has been found to affect lexical decision (see, e.g., Conrad, Carreiras, Tamm, & Jacobs, 2009), we statistically controlled for such measure in all our analyses.[1]

Table 1 Examples of word stimuli for each experimental condition

Stress neighborhood	Stress type	
	Dominant stress	Non-dominant stress
Consistent	graNIta (*slush*)	MISsile (*missile*)
Inconsistent	seNIle (*senile*)	BIbita (*drink*)

---

[1] Another variable that affects lexical decision in Spanish (Carreiras, Alvarez, & Devega, 1993) and German (Conrad, Stenneken, & Jacobs, 2006) is the frequency of the first syllable. However, the effect has not been found in Italian, where syllable frequency only affects the production stage (Sulpizio & Job, 2010).

Table 2 Summary statistics: Means (and standard deviations) for the words used in the experiments; words with dominant stress and consistent ("GraNIta," *slush*), and inconsistent ("seNIle," *senile*) stress neighborhood; words with non-dominant stress and consistent ("MISsile," *missile*) and inconsistent ("BIbita," *drink*) stress neighborhood. Examples of target words are in parentheses

Item Variables	Dominant stress		Non-dominant stress	
	Consistent stress neighborhood (graNIta)	Inconsistent stress neighborhood (seNIle)	Consistent stress neighborhood (MISsile)	Inconsistent stress neighborhood (BIbita)
Frequency	9.43 (11.31)	12.2 (21.88)	8.13 (12.54)	7.4 (5.72)
Length in letters	6.4 (0.67)	6.43 (0.67)	6.6 (0.62)	6.6 (0.62)
Number of orthographic neighbors (*N*)	3.1 (1.58)	2.33 (1.58)	2.9 (1.97)	2.8 (1.58)
Mean neighbors' frequency	5.6 (5.44)	8.36 (13.76)	5.73 (6.68)	4.63 (5.97)
Mean familiarity	2.41 (0.61)	2.14 (0.6)	2.26 (0.47)	2.16 (0.6)
S-neighborhood size (total number)*	289 (78.09)	299 (48.99)	452 (39.6)	593 (134.2)
Proportion of stress consistent words+	0,68 (.04)	0,15 (.03)	0,86 (.02)	0,16 (.01)

* Total number of words in each stress neighborhood defined by endings (e.g., neighbors in -Ita)

+ Proportion of words with the same ending as a target word and congruent stress (e.g., "graNIta," "Gita," "adiBIta," "conDIta," but not "Bibita")

The size of each word stress neighborhood and the proportion of stress consistent words were not matched, but they favored words with non-dominant stress (against our hypothesis of an advantage for dominant stress), as non-dominant neighborhoods have fewer word-ending types, but each word ending is included in a large number of words and of consistent neighbors (see Table 1).

A set of 120 filler nonwords was included: Forty-five nonwords ended with a final sequence mainly associated with dominant stress (e.g., "balona"), 45 nonwords with a final sequence mainly associated with non-dominant stress (e.g., "necile"), and 30 nonwords with a final sequence neither biased toward dominant nor toward non-dominant stress (e.g., "gorafo"). The nonwords were classified according to the likely stress pattern used in former studies, where their stress pattern was recorded (Colombo et al., 2014).

The experiment had a 2 (stress dominance: dominant vs. a non-dominant stress) × 2 (consistent vs. inconsistent neighborhood) design, and both factors were within participants. Stimuli were presented in one block, in a random order.

Procedure The experiment was run using the E-Prime Software (Psychology Software Tools, Pittsburgh, PA; www.pstnet.com). Stimuli were presented on a computer screen, at a distance of about 50 cm from the participant. Each trial started with a fixation cross, presented for 400 ms in the center of the screen. Then, a stimulus word/nonword appeared in the same position and was presented until the participant's response or for a maximum of 600 ms. The interstimulus interval was

1,500 ms. Participants were tested individually. They were asked to indicate, as quickly and accurately as possible, whether each letter string was a real word or not: The response was given by pressing either key 1 or key 5 of a five-keys response-box. Key selection was counterbalanced across participants. A set of 60 practice trials preceded the experiment.

Results

Because the task required a binary decision, only data from participants with mean accuracy above 60 % were kept. Thus, three participants were excluded from the analyses because of their low level of accuracy (48 %, 52 %, and 54 % of accuracy, respectively). Nonwords were analyzed separately from words.

Reaction times and errors for the word and nonword data were both analysed using mixed-effects models (Baayen, Davidson, & Bates, 2008). The models were fitted using the *lmer* function (*languageR* package) in R software (version 2.11); *p* values were calculated using the MCMC procedure, sampling 10,000 times (Baayen et al., 2008). Participants and items were treated as random factors. Results are reported in Table 3.

Words

Reaction times Results are reported in Table 3. Only correct responses were analyzed. RTs were log transformed to reduce

Table 3 Mean RTs for correct responses and percentage of errors by condition (with standard deviations), in Experiment 1 (mixed block)

Stress type	Stress neighborhood consistency			
	Consistent		Inconsistent	
	Mean RTs	%E	Mean RTs	%E
Dominant stress	462 (13)	12.16 (9.25)	454 (14)	7.17 (9.6)
Non-dominant stress	478 (18)	34.34 (12.14)	474 (20)	27.85 (10.83)
Stress effect	16	22.2	20	20.7

skewness of data. A mixed-effects model was performed with RTs as dependent variable and stress type (dominant vs. non-dominant) and stress neighborhood consistency (consistent vs. inconsistent neighbors) as fixed factors. Bigram frequency was also entered as predictor. The same analysis was also adopted in Experiments 2 and 3.
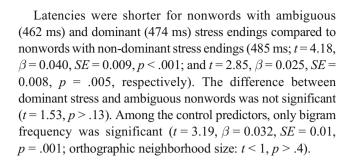
The model showed a significant effect of stress type ($t = 2.92$, $\beta = 0.038$, $SE. = 0.013$, pMCMC = .004): Participants were faster to identify the target stimulus as a word when it had dominant stress than when it had non-dominant stress. No further effect reached significance (stress neighborhood consistency: $t = -1.8$, $p > .05$; stress type by stress neighborhood consistency interaction: $t < 1$; bigram frequency: $t < 1$).

Response accuracy A mixed-effects model was performed with response accuracy as the dependent variable and stress type (dominant vs. non-dominant) and stress neighborhood consistency (stress consistent vs. stress inconsistent) as fixed factors. Bigram frequency was also entered as a fixed factor. The same analysis was also adopted in Experiments 2 and 3.

Overall error rate was 18.46 % of all data points. A main effect of stress type was found ($z = -4.46$, $\beta = -1.987$, $SE = 0.445$, $p < .001$). Participants were less accurate in categorizing stimuli as words when they had non-dominant stress than when they had dominant stress. No further effect reached significance (stress neighborhood consistency: $z < 1$; stress type × stress neighborhood consistency interaction: $z < 1$; bigram frequency: $z < 1$).

Nonwords

Reaction times Only correct responses were analyzed. RTs were log transformed to reduce skewness of data. A mixed-effects model was performed with RTs as dependent variable and type of ending (dominant vs. non-dominant vs. neutral) as fixed factor. Orthographic neighborhood size and bigram frequency were also entered as predictors to control for their effect on response latencies.

Latencies were shorter for nonwords with ambiguous (462 ms) and dominant (474 ms) stress endings compared to nonwords with non-dominant stress endings (485 ms; $t = 4.18$, $\beta = 0.040$, $SE = 0.009$, $p < .001$; and $t = 2.85$, $\beta = 0.025$, $SE = 0.008$, $p = .005$, respectively). The difference between dominant stress and ambiguous nonwords was not significant ($t = 1.53$, $p > .13$). Among the control predictors, only bigram frequency was significant ($t = 3.19$, $\beta = 0.032$, $SE = 0.01$, $p = .001$; orthographic neighborhood size: $t < 1$, $p > .4$).

Response accuracy Overall error rate was 16.27 %. Nonwords with ambiguous (91 %) and dominant (86 %) stress endings were more accurate than nonwords with non-dominant stress endings (76 %; $z = -4.951$, $\beta = -1.262$, $SE = 0.255$, $p < .001$; and $z = -3.898$, $\beta = -0.859$, $SE = 0.220$, $p < .001$, respectively). The difference between dominant stress and ambiguous nonwords was not significant ($z = 1.49$, $p > .13$). Among the control predictors, both orthographic neighborhood size and bigram frequency were significant ($z = -3.02$, $\beta = -0.227$, $SE = 0.075$, $p = .002$; and $z = -2.04$, $\beta = -0.529$, $SE = 0.259$, $p = .04$, respectively).

Discussion

The data of Experiment 1 show that target words with dominant stress were recognized faster and more accurately than non-dominant stress targets, and there was no effect of stress neighborhood consistency. Possibly, participants activated lexical phonology because of the different stress patterns in the stimuli, and dominant stress words, as more typical or familiar stimuli, were facilitated. Moreover, nonwords most likely associated with non-dominant stress were slower and more error prone than both words with dominant and with ambiguous stress. These effects were significant, although orthographic variables were included in the model to control for confounding factors. Thus, the disadvantage of nonwords with non-dominant stress suggests the involvement of phonology in nonword lexical decision.

In Experiment 2, we investigated if different processing might occur when words were presented in lists blocked by stress. The idea was as follows: Processing in Experiment 1 may have compelled lexical phonological contribution because of the simultaneous presence of stimuli with different stress, and of the resulting tendency to activate the phonological lexicon in order to discriminate words from nonwords. When all stimuli in a block have the same stress, the tendency to activate lexical phonology might be less strong, because, in principle, stress might be assigned following the list suggestion: If all the stimuli in a list have dominant stress, participants might be inclined to be consistent with the list and apply dominant stress. Thus, the stress dominance effect should diminish. If, in contrast, the tendency to activate lexical

phonology is not subject to strategic adjustments, the dominant stress advantage should still occur.

We used the same stimuli as in Experiment 1. In each block words either had dominant or non-dominant stress. Nonwords in each block were selected on the basis of the probability of being named with dominant or non-dominant stress, and were included so as to be congruent with the word list. Thus, within each block words and nonwords with the same stress pattern were presented.

Experiment 2

Method

Participants Twenty-nine participants (7 males; mean age = 23.06, SD = 2.61) from the University of Padua took part in the experiment. They were all native Italian speakers with normal or corrected-to-normal vision. None had participated in the previous experiments.

Materials and design Words and nonwords were the same as in Experiment 1. Stimuli were presented in two blocks. Each block was composed of half words and half nonwords. All stimuli within a block had the same stress.

Procedure The same procedure as in Experiment 1 was adopted, except that the whole set of stimuli was divided into two blocks. Each participant was presented with two blocks, one with dominant stress words and nonwords and one with non-dominant stress words and nonwords. Stimuli were randomized within each block, and block order was counterbalanced across participants.

Results

Five participants were excluded from the analyses because of a very low level of accuracy (40 %, 50 %, 51 %, 53 %, and 45 % of accuracy, respectively). Words and nonwords were separately analyzed.

Reaction times and errors were both analyzed using mixed-effects models (Baayen et al., 2008). Participants and items were treated as random factors. Results are reported in Table 4.

Words

Reaction times The mixed-effects model on log RTs showed that the main effect of stress type was significant ($t = 4.12$, $\beta = 0.04$, $SE = 0.009$, pMCMC $<. 001$). Participants were slower when categorizing word stimuli with non-dominant than with dominant stress. No further effect reached significance (stress

Table 4 Mean RTs for correct responses and percentage of errors by condition (with standard deviations) in Experiment 2 (pure blocks)

Stress type	Stress neighborhood consistency			
	Consistent		Inconsistent	
	Mean RTs	%E	Mean RTs	%E
Dominant stress	445	21.13	447	18.87
	(27)	(13.72)	(27)	(16.75)
Non-dominant stress	465	26.87	459	25.29
	(27)	(13.41)	(31)	(12.48)
Stress effect	20	5.74	12	6.42

neighborhood consistency: $t < 1$; stress type × stress neighborhood consistency: $t = -1.3$, $p > .1 < 1$; bigram frequency: $t < 1$).

Response accuracy Overall error rate was 22.89 %. No effect reached significance (stress type: $z < 1$; stress neighborhood consistency: $z < 1$; stress type × stress neighborhood consistency: $z < 1$; bigram frequency: $z = 1.5$, $p > .1$).

Nonwords

Reaction times Latencies were shorter for nonwords with ambiguous (455 ms) and dominant stress endings (458 ms) compared to nonwords with non-dominant stress endings (471 ms; $t = 3.16$, $\beta = 0.028$, $SE = 0.008$, $p = .002$; and $t = 3.37$, $\beta = 0.027$, $SE = 0.008$, $p < .001$, respectively). The difference between dominant stress and ambiguous nonwords was not significant ($t < 1$). Of the other predictors, only orthographic neighborhood size was significant ($t = 2.20$, $\beta = 0.006$, $SE = 0.028$, $p = .02$).

Response accuracy Overall error rate was 22.7 %. Nonwords with ambiguous (83 %) and dominant stress endings (79 %) were more accurate that nonwords with non-dominant stress endings (71 %; $z = -3.506$, $\beta = -0.638$, $SE = 0.182$, $p < .001$; and $z = -3.563$, $\beta = -0.574$, $SE = 0.161$, $p < .001$, respectively). The difference between dominant stress and ambiguous nonwords was not significant ($z = < 1$). Of the other predictors, only orthographic neighborhood size was significant ($z = -4.52$, $\beta = -0.245$, $SE = 0.054$, $p < .001$).

Discussion

The results of Experiment 2 replicated those of Experiment 1. There was again an advantage for dominant stress words compared to non-dominant stress words. Although overall the pattern was the same as in Experiment 1, the size of the effect showed a tendency to decrease, at least in the analysis of response accuracy. Words with dominant stress showed an increase in error rate compared to Experiment 1, although

there was a slight reduction in latencies. Apparently, then, participants were able to take some advantage of the constant stress within a list, at least with non-dominant words. As to nonwords, the same pattern was found as in Experiment 1, with nonwords having dominant or ambiguous endings easier and more accurate than nonwords with non-dominant endings. The nonwords types also differed in bigram frequency and proportion of words with dominant stress sharing the same endings. However, the bigram frequency of nonwords with ambiguous (11.3) and non-dominant (11.4) endings did not differ, while there was a significant difference in latencies and accuracy between the two nonwords types. Thus, although we cannot unambiguously determine the cause of the difference, possibly it was due to the nature of endings in the two nonword types.

In Experiment 1 and 2, endings of words and nonwords partially overlapped (39 %). Moreover, endings of all stimuli in the first two experiments belonged to large-sized neighborhoods, and were strongly biased toward one or the other stress pattern. Thus, endings could not help participants in discriminating words from nonwords.

In Experiment 3, we examined the possibility that the results of the two experiments were mainly determined by the difficulty to discriminate words from similar nonwords because of the overlapping endings, and by the consequent tendency to activate the lexical phonological representation of words. It is well known that lexical decision is affected by strategic manipulations depending, for example, on the type of nonword included (James, 1975; Shulman & Davison, 1977; Stone & Van Orden, 1993; Yap, Balota, Cortese, & Watson, 2006). Thus, in Experiment 3 we presented nonwords that were more dissimilar to words, compared to Experiments 1 and 2. We created a new set of nonwords, which did not share endings with words. These nonword endings belonged to small or to stress-ambiguous neighborhoods that would not provide robust cues to stress (Colombo et al., 2014; Sulpizio et al., 2013). These nonwords were also lower in bigram frequency, orthographic neighborhood size and orthographic neighborhood frequency (see Table 5), which increased their dissimilarity to words. If lexical decision can be affected by contextual effects of word endings, smaller or no effects of word stress would be expected in Experiment 3.

Table 5 Comparison of the mean values of the variables for the nonwords in Experiments 1–2, and 3

Variables	Experiments 1–2	Experiment 3	t test
Bigram frequency	11.53	11.26	7.03, p < .001
N size	1.13	0.78	7.8, p < .001
N frequency	18.02	0.16	1.9, p = .05

Experiment 3

Method

Participants Twenty-four participants (4 males; mean age = 21.41, SD = 0.82) from the University of Padua took part in the experiment. They were all Italian native speakers with normal or corrected-to-normal vision. None had participated in the previous experiments.

Materials and design The same words as in Experiment 1 were used. A new set of 120 filler nonwords was included by using mainly final sequences belonging to small or ambiguous neighborhoods, neither biased toward dominant nor toward non-dominant stress (e.g., -odo), as verified in former studies (Colombo et al., 2014). The new set of nonwords differed from that used in Experiments 1 and 2 on the following dimensions: bigram frequency, orthographic neighborhood size, and orthographic neighborhood frequency (see Table 5). The same design as in Experiment 1 was adopted.

Procedure The same procedure as in Experiment 1 was adopted.

Results Two participants were excluded from the analyses because of a very low level of accuracy (46 % and 53 %, respectively). Nonwords were only used as fillers and were not analyzed (mean RTs = 477 ms; mean error rate = 31.1 %). For words, reaction times and errors – 22.89 % of all data points – were both analysed using mixed-effects models (Baayen et al., 2008). Participants and items were treated as random factors. Results are reported in Table 6.

Reaction times The model on log RTs showed a main effect of stress type ($t = 3.27$, $\beta = 0.047$, $SE = 0.014$, pMCMC < .001): Participants were slower with non-dominant than with dominant stress words. No further effect reached significance (stress neighborhood consistency: $t < 1$; stress

Table 6 Mean RTs for correct responses and percentage of errors by condition (with standard deviations) in Experiment 3

Stress type	Stress neighborhood consistency			
	Consistent		Inconsistent	
	Mean RTs	%E	Mean RTs	%E
Dominant stress	447 (23)	26.49 (13.84)	452 (24)	30.31 (15.94)
Non-dominant stress	469 (31)	39.97 (17.73)	461 (23)	34.17 (14.15)
Stress effect	22	13.48	9	3.86

type × stress neighborhood consistency: $t < 1$; bigram frequency: $t = 1.73$, $p > .05$).

Response accuracy The mixed-effects model on response accuracy showed a main effect of stress type ($z = -3.1$, $\beta = -0.680$, $SE = 0.219$, $p < .001$), with participants being less accurate when categorizing words with non-dominant than with dominant stress. No further effect reached significance (stress neighborhood consistency: $z < 1$; stress type × stress neighborhood consistency: $z = 1.4$, $p > .1$; bigram frequency: $z < 1$).

Discussion

The results of Experiment 3 show that changing the nonword context did not greatly affect the pattern of data. Words with dominant stress were again recognized faster and more accurately than words with non-dominant stress. Although the general pattern remained the same in the three experiments, the size of the effects was reduced in Experiment 3, compared to Experiment 1, in both latencies and error rate. This reduction in effect size was supported in the joint analysis of the two experiments on both errors and latencies. In the RTs' analysis, stress ($t = 3.25$, $\beta = 0.038$, $SE = 0.011$, $p = .001$); experiment ($t = -3.36$, $\beta = -0.043$, $SE = 0.012$, $p = .001$) and consistency × experiment were significant ($t = -3.20$, $\beta = 0.031$, $SE = 0.009$, $p = .001$). The experiment and stress factors indicated that latencies were faster in Experiment 3 than in Experiment 1, and for dominant than for non-dominant stress words. The interaction showed that the effect of consistency was different in the two experiments, with slower latencies for consistent over inconsistent words in Experiment 1 ($t = -2.071$, $\beta = -0.0204$, $SE = 0.009$, $p = .04$) but no effect in Experiment 3 ($t < 1$).

In the analysis of errors, there were more errors in Experiment 3 than in Experiment 1 ($z = -5.03$, $\beta = -1.175$, $SE = 0.233$, $p < .001$), and the stress effect was significant ($z = -5.70$, $\beta = -1.628$, $SE = 0.285$, $p < .001$), but smaller in size in Experiment 3 than in Experiment 1 (stress × experiment; $z = 4.45$, $\beta = 0.837$, $SE = 0.187$, $p < .001$). Consistency × experiment ($z = -3.49$, $\beta = -0.759$, $SE = 0.217$, $p < .001$) and the three-way interaction stress × consistency × experiment were also significant ($z = 2.54$, $\beta = 0.704$, $SE = 0.276$, $p = .01$). The three-way interaction showed that the dominant stress advantage was reliable in Experiment 3 for consistent words (dominant stress advantage: 13.48 %; $z = -3.345$, $p < .001$), while substantially decreasing for inconsistent words (dominant stress advantage: 3.86 %; $z = -1.06$, $p > .2$). This reduction however was not apparent in Experiment 1 (dominant stress advantage for consistent words: 22.2 %; $z = -4.178$, $p < .001$; dominant stress advantage for inconsistent words: 20.7 %; $z = -7.33$, $p < .001$).

The comparison between experiments also showed a trade-off in Experiment 3, with a decrease in latencies, but an increase in error percentage, compared to Experiment 1. This trade-off suggests that the change in the nonword context affected processing, with participants tending to give a fast response, that often was mistaken, and with a reduction of the dominant stress advantage in Experiment 3. Exactly which characteristics of nonwords produced this reduction is not clear, given that the three nonword types were significantly different for bigram frequency, orthographic neighborhood size, and proportion of words with dominant stress sharing their endings. These aspects were not controlled.

The lack of a clear consistency effect in the present study, where lexical decision was used, stands in strong contrast with the results obtained in former studies with a reading-aloud task. In our view, this depends on processing differences due to the task, but it might also be that our results were strongly affected by the stimuli we used. To rule out such a possibility, we ran a control reading-aloud experiment, in which the same stimuli as Experiment 1 were used. We tested whether the same words would produce the typical stress neighborhood effect often reported in the literature on reading aloud.

Experiment 4

Method

Participants Twenty-eight participants (15 males; mean age = 23.28, $SD = 3.12$) from the University of Trento took part in the experiment. They were all native Italian speakers with normal or corrected-to-normal vision. None had participated in the previous experiments.

Materials and design The same as in Experiment 1.

Procedure Participants were tested individually. They were instructed to read the targets as quickly and accurately as possible.

Stimuli were displayed in black uppercase letters, centered on the computer screen. Before the presentation of each stimulus, a fixation cross was displayed for 500 ms. Each stimulus disappeared at pronunciation or after 1,500 ms. There was an interstimulus interval of 1,500 ms. The experiment was preceded by a practice session with stimuli not included in the experimental trials. The experimenter noted the naming errors. The participants' responses were also recorded to allow further analyses of errors and control of stress pronunciations.

Results and discussion Analyses were run only on naming errors (8.67 % of all data points), which included mispronunciation errors, phonemic errors, and stress errors (see Table 7 for the relative proportion of error types). Reaction times were not analyzed since stimuli in different conditions were not

Table 7 Mean relative percentage of each type of error, compared to the total error, for each condition, in Experiment 4 (standard deviations in parentheses)

Stress type	Stress neighborhood consistency			
	Consistent		Inconsistent	
	% Phonemic errors	% Stress errors	% Phonemic errors	% Stress errors
Dominant stress	1.55 (3.33)	2.98 (3.31)	4.65 (4.74)	10.12 (4.01)
Non-dominant stress	2.50 (3.22)	2.03 (2.91)	4.05 (4.91)	8.58 (5.01)

matched on initial phonemes, which are well known to affect naming times (e.g., Kessler, Treiman, & Mullenix, 2002). Pseudowords were only used as fillers and were not analyzed. Results are reported in Table 8.

Statistical analyses, based on mixed-effects models (Baayen et al., 2008), were carried out combining all error types (but the analyses carried out for each type of error separately were consistent). Accuracy was entered as the dependent variable and stress type (dominant vs. non-dominant) and stress neighborhood consistency (stress consistent vs. stress inconsistent) as fixed factors. Words bigram frequency was also entered as fixed factor. Participants and items were treated as random factors. The model showed a main effect of stress neighborhood consistency ($z = -2.67$, $\beta = -1.212$, $SE = 0.453$, $p = .007$), with participants being less accurate when reading stress inconsistent than when reading stress consistent words. No further effect reached significance (stress type: $z < 1$; stress type × stress neighborhood consistency: $z < 1$; bigram frequency: $z < 1$).

The data of the naming experiment confirmed the results of former studies (e.g., Burani & Arduino, 2004; Burani et al., 2014; Paizi et al., 2011), with an advantage for consistent over inconsistent stress neighbors, and no stress dominance effect, thus suggesting that the results obtained in our lexical decision experiments were not due to the particular nature of the stimuli.

Table 8 Percentage of errors by condition (with standard deviations), in Experiment 4 (Reading aloud)

Stress type	Stress neighborhood consistency	
	Consistent	Inconsistent
Dominant stress	4.54 (4.90)	13.93 (4.89)
Non-dominant stress	4.17 (4.12)	12.03 (6.37)

General discussion

The present study aimed to investigate the role of lexical stress and stress neighborhood consistency in word recognition in a transparent orthography. To summarize, Experiments 1–3 showed an advantage for dominant stress over non-dominant stress words, despite changes in nonword context and list composition. This stress effect was significant in each lexical decision experiment but tended to decrease with the change in nonword context. When nonwords became less similar to experimental words, latencies became shorter, but error rates increased. Finally, the dominant stress advantage was no longer apparent in the reading-aloud task.

The stress effect we found partially replicates Colombo (1992), who also found an advantage for dominant stress, although only in the measurement of errors. In contrast, Burani and Arduino (2004) did not find it. Several factors may be responsible for this difference, probably the most important of which is that in the latter study selected materials had a higher number of stress friends for non-dominant than for dominant stress words. We note, however, that Burani and Arduino (2004) did not find any effect of stress neighborhood consistency in lexical decision. Other factors may include the nonword type, as the present experiments show that changes in nonword context may provide slight differences in the results.

Phonology and the dominant stress advantage

In the introduction, we expected an effect of the dominant stress because overall activation within the phonological lexicon would be greater for word types with the dominant pattern, and the decision process would be able to monitor this activity and produce an advantage for words with dominant stress.

Although in principle the dominant stress advantage might be driven just by faster access to the lexical phonological representations, it is also possible that the sublexical level also contributes to the computation of phonology. Segmental phonology would be activated very fast in Italian and is not error prone, and its output, maintained in the buffer, would feed the phonological lexicon in addition to the activation from orthography. Words with dominant stress would receive more feedback activation from the phonological lexicon and therefore would be recognized faster than words with non-dominant stress. This view is consistent with simulations of the reading process in the computational model of Italian by Perry et al. (2014). Moreover, the model includes two pathways, for lexical and sublexical phonology, and easily lends itself to the possibility of relatively independent manipulations of either pathway to explain nonword context effects, as shown by the authors.

As apparent from the analyses, the change of nonword context in the three lexical decision experiments slightly but

significantly affected the results. Latencies were significantly faster in Experiment 3, where words and nonwords were more dissimilar, than in Experiment 1. The inclusion of nonwords with rare or ambiguous endings in Experiment 3 produced a higher error rate, and the overall dominant stress advantage decreased in the pattern of errors from 21 % to 9 %. This result may have been due to differences in the nonword types (in bigram frequency, for example). However, considering the three-way interaction experiment by stress by consistency in the joint analysis of Experiments 1 and 3, it seems likely that word endings were at least in part responsible for the differences. Possibly, they were indeed processed by our participants and, to some extent, affected the way they performed lexical decision, thus supporting the idea that both sublexical and lexical processes were involved in the experiments. Also supportive of this interpretation is the difference between nonword types in the analyses of both latencies and errors.

Our results are partially consistent with findings reported by Jouravlev and Lupker (2014), who manipulated stress type and neighborhood consistency in lexical decision. For Russian adjectives, the only grammatical category with strong asymmetries in the relative proportion of initial vs. final syllables, the authors reported an advantage for initial syllables stress (the most frequent stress type), no effect of consistency, and, in the pattern of errors only, an effect of consistency affecting just the less common stress pattern. Our results overall confirmed the advantage for the dominant stress pattern.

Context effects and task differences

The results of Experiment 4 showed that the same stimuli that produced a clear dominant stress advantage and no effect of stress neighborhood consistency in lexical decision (Experiment 1) showed exactly the reverse pattern (a stress neighborhood consistency effect, but no stress dominance effect) in reading aloud. The dissociation suggests that different mechanisms were at work in the two tasks. Therefore, we are confident that the effects we reported in lexical decision are due to how the system recognizes the stimuli and not to the nature of stimuli.

The different involvement of processes in reading aloud and lexical decision has been thoroughly investigated and accounted for in different ways. Balota and collaborators (Balota & Chumbley, 1984; Balota & Spieler, 1999; Colombo, Pasini & Balota, 2006; Yap et al., 2006) claimed that word–nonword discrimination involves two different processes, a familiarity evaluation that may drive responses in addition to lexical activation rate, and an attentional process, required when nonwords are very similar to words. According to the two-process model, when words and nonwords are very different, an accurate orthographic-phonological check is bypassed, and familiar stimuli may be easily accepted as words. However, when nonwords are similar to words,

differing, for example, by one letter, their discrimination from words, in particular low-frequency and less familiar words, requires an in-depth processing before a response is given.

In the present study, nonwords were more similar to words in Experiments 1 and 2 than in Experiment 3. In the latter, the greater dissimilarity may have induced participants to avoid an accurate check, and to give a fast, but often inaccurate response, thus explaining the 12 % increase in overall error rate. The much greater frequency of the penultimate syllable stress over all words in Italian makes this type of stress more familiar. However, when these words have inconsistent endings, this makes them comparatively less familiar, and this might explain the greater error rate increase (23.14 %) in Experiment 3 compared to Experiment 1 for words with dominant stress but inconsistent endings ("seNIle"). This is less of a problem when a relatively fuller processing of the words is carried out, as in Experiment 1, but induces more errors when processing is made faster by the dissimilarity of nonwords. This interpretation rests on the idea that words were distinguished from nonwords on the basis of the familiarity of their phonological representation, and that words with dominant stress have a more familiar representation because dominant stress is more frequent. This interpretation is more suitable to account for the results, compared to one purely in terms of orthography, which would not be able to account for the presence of the stress effect in Experiment 3, suggesting that phonology was active.

The present results might be explained in a slightly different framework. Stone and Van Orden (1993; see also Yap et al., 2006) used a random-walk model to account for the variation in the size of the frequency effect as a function of nonword type. Specifically, they found that the size of the frequency effect (a marker of lexical involvement) increased with the increase in word/nonword similarity (e.g., going from illegal strings, to legal nonwords, and to pseudohomophones). In contrast to the two-process model, this framework assumes that only signal strength, an evidence accumulating process, is responsible for the effects. Within this process, high frequency words have a stronger signal than low frequency words, since signal strength is greater for stimuli that are processed more efficiently. When nonwords are very similar to words (e.g., with pseudohomophones, or with shared endings), signal strength decreases for both high- and low-frequency words, increasing overall latencies. The frequency effect increases as well, with the relation between signal strength and the time to give a response following a nonlinear-concave function (Stone & Van Orden, 1993). This means that the same change in the rate of evidence accumulation for a signal has a greater impact on processing times of stimuli that are processed less efficiently (e.g., low-frequency words).

To extend this interpretation to the present data, we might assume that dominant stress words would have higher signal strength because they are more frequent as a type compared to

words with non-dominant stress. According to the random-walk model, changes in the nonword context produce changes in the response criterion, which is the distance of the decision boundary (i.e., word/nonword) from the start point and indicates how easy it is to take a decision in terms of processing involvement. With a decrease in word/nonword similarity, response boundaries become less conservative, producing faster responses but becoming more error prone. Thus, with the decrease in word/nonword similarity from Experiment 1 to Experiment 3, response boundaries became less conservative and responses were faster but more prone to errors. As a result, the size of the stress effect was larger (significantly for accuracy and numerically for RTs) in Experiment 1 than in Experiment 3. Moreover, error rates increased with the decrease in word/nonword similarity. Our overall pattern is similar to that reported by Stone and Van Orden's model. Note that the model predicts that the nonword manipulation should impact more on stimuli with lower signal strength (i.e., low-frequency words in Stone and Van Orden's study). This being the case, we would have expected the manipulation to have a stronger impact on non-dominant stress words, which should be, prima facie, the stimuli with lower signal strength. Differently, in Experiment 3, there was a larger decrease in accuracy for words with dominant stress, in particular for those with inconsistent endings ("seNIle"). In contrast, the nonword manipulation in Experiments 1–3 had a smaller impact on those stimuli that showed the lowest performance overall. Thus, neither the two-process model nor the random-walk model can completely explain the whole set of results of the present study. To summarize, the process of word recognition produced a pattern of results quite different from those exhibited in reading aloud, suggesting that the nature of processes involved in lexical decision are quite dissimilar from those involved in reading aloud, where perhaps production mechanisms are more relevant.

Overall, the present results show an effect of stress in lexical decision that supports the idea of automatic phonological activation. The decrease in the stress effect in Experiment 3 may have been related to a decrease in lexical effects, and a simultaneous increase in sublexical effects, as in Stone and Van Orden's study. This is not to say that phonological effects cannot vanish, under the appropriate conditions: for example, in Peressotti and Colombo (2012; Experiment 4) no effect of pseudohomophones was found in lexical decision, since, because of the type of nonwords included, participants were able to perform the task based solely on orthography.

To conclude, our study investigated processing of the same stimuli, requiring the same response, under different processing conditions, determined by either a different context (Experiments 1 and 3) or a blocking of stimuli (Experiment 2), and a further comparison with a different task (Experiment 4). The results showed a robust effect of prosodic

manipulation, showing that stress information may play an important role during word recognition.

Appendix

Words used in Experiments 1–4.

Stimuli with dominant stress and consistent stress neighborhood: arcana, avaro, balena, canora, carota, castoro, cicuta, decoro, dimora, fachiro, forati, fulgore, granita, icona, mulino, papiro, pedine, pepita, pulcino, querela, ristoro, scolaro, silicio, sirena, sonoro, sottane, spremuta, tapiro, vaccino, vampiro.

Stimuli with dominant stress and inconsistent stress neighborhood: alfiere, asilo, barile, baule, concime, cratere, cupido, enzima, fienile, indiano, infido, intrico, macaco, mangime, metano, monile, moviola, ortica, ostile, paiolo, pianola, raviolo, senile, sultano, tagliola, tritolo, tucano, tutela, vaiolo, vescica.

Stimuli with non-dominant stress and consistent stress neighborhood: bilico, bolide, buttero, cantico, capsule, celtico, colica, cresima, distico, duttile, esule, flebile, fodero, fossile, futile, infimo, ionico, labile, lacero, ludico, missile, porfido, raffica, sciatica, sferica, spasimo, tattile, tunica, vigile, zenzero.

Stimuli with non-dominant stress and inconsistent stress neighborhood: afona, alluce, anfore, argine, barbaro, bibita, canfora, concavo, darsena, despota, diafana, domino, folgore, fosforo, fradicio, lapponi, logoro, mescita, orbita, orfane, porpora, recluta, redini, ruggine, satiro, suddito, timpani, zattera, zefiro, zingaro.

Nonwords used in Experiments 1, 2, and 4: adimo, aldume, astola, astone, ballido, bedule, begano, berino, bettuce, bildese, biluta, birume, bistone, bodune, bollice, bovero, camoni, castubo, ceberi, celido, ceraso, chiroga, cobota, comilo, corafo, dediro, denora, derule, dirloni, dirtola, dolame, dorreta, ellate, empomi, eperi, etuce, faboga, facero, faride, feluge, fiocimi, fispuce, fraboro, gambura, ganoci, gecana, ginido, gofano, gospilo, govato, gramulo, gurafo, iruta, lagule, lemana, leneri, licero, lidame, lighena, linata, lirdane, mavena, mepela, mepuce, mulica, nalafo, necile, ninoro, nosela, nostubo, oddimo, ofide, onese, paghite, paroga, piedomi, pirtubo, pitano, pivida, polaso, pudela, rastole, reluge, remoni, rincolo, rolide, rudomi, saccubo, saloci, sbaccole, scelate, settame, sintura, sorato, sperilo, spisida, stipimi, stoleta, strimole, tenuge, terpico, tevone, tolebre, trofulo, tugile, umbica, urfina, valona, vepela, visima, vistena, voberi, zaloci, zefomi, zelido, zifilo, zilota, ziltica, zipero, zircona.

Nonwords used in Experiment 3: badusa, balusa, batila, bicopa, bidaro, bilango, bofengo, bogada, bostuda, bunesi, burnado, cambupo, catrofa, celango, cevaro, cimpofe, copifa, cudige, dabefo, dabengo, dalise, darila, delaco, dereto, dibipa,

dorango, duntesi, fazubo, fenula, fenzada, fonnira, funzaca, gafodo, galifo, galova, galtodo, gampesi, ganeca, gatala, gavira, grenofo, gulefo, gurife, labofo, lebifo, lenubo, lintaro, lofada, lopiro, lubego, madife, mefaca, midefa, midego, midesa, mindofe, motipa, muntifo, nabice, nafeno, naltige, naltipe, nelluco, neltofe, nerise, nesupo, nizzeto, paduco, paltefa, pecila, peviro, pirtefa, pivaco, purofe, ragifa, ravesa, rilofa, rimpaca, rinzapo, ronise, rucodo, rufipe, saveca, sefoma, selapo, sevipo, sicova, simbego, soveto, supala, tadrofa, tefopa, tellusa, tobefo, tolado, tolice, tolipe, tozife, truvapo, turipo, tuveno, vamige, vedoma, vibala, vibofe, vintado, vopesa, vostife, vusteca, vutopa, zartipa, zatula, zavuda, zecife, zelaco, zibofo, zolubo, zonupo, zufice.

References

Arciuli, J., & Cupples, L. (2006). The processing of lexical stress during visual word recognition: Typicality effects and orthographic correlates. *The Quarterly Journal of Experimental Psychology, 59,* 920–948.

Arciuli, J., Monaghan, P., & Seva, N. (2010). Learning to assign lexical stress during reading aloud: Corpus, behavioural and computational investigations. *Journal of Memory and Language, 63,* 180–196.

Ashby, J., Sanders, L. D., & Kingston, J. (2009). Skilled readers begin processing sub-phonemic features by 80 ms during visual word recognition: Evidence from ERPs. *Biological Psychology, 80,* 84–94.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 340–357.

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128,* 32–55.

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thorton, A. M. (2005). Corpus e Lessico di frequenza dell'italiano scritto (CoLFIS) [CoLFIS: Corpus and frequency lexicon of written Italian]. Retrieved from http://www.istc.cnr.it/material/database/colfis/index_eng.shtml

Brown, P., Lupker, S. J., & Colombo, L. (1994). Interacting sources of information in word naming: A study of individual differences. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 537–554.

Burani, C., & Arduino, L. S. (2004). Stress regularity or consistency? Reading aloud Italian polysyllables with different stress patterns. *Brain and Language, 90,* 318–325.

Burani, C., & Cafiero, R. (1991). The role of subsyllabic structure in lexical access to printed words. *Psychological Research, 53,* 42–52.

Burani, C., Paizi, D., & Sulpizio, S. (2014). Stress assignment in reading Italian: Friendship outweighs dominance. *Memory & Cognition, 42,* 662–675.

Carreiras, M., Alvarez, C. J., & Devega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32,* 766–780.

Colombo, L. (1992). Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 987–1003.

Colombo, L., Deguchi, C., & Boureux, M. (2014). Stress priming and statistical learning in Italian nonword reading: Evidence from children. *Reading & Writing, 27,* 923–943.

Colombo, L., Pasini, M., & Balota, D. A. (2006). Dissociating the influence of familiarity and meaningfulness from word frequency in naming and lexical decision performance. *Memory & Cognition, 34,* 1312–1324.

Colombo, L., & Tabossi, P. (1992). Strategies and stress assignment: Evidence from a shallow orthography. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 319–340). Amsterdam: Elsevier.

Colombo, L., & Zevin, J. (2009). Stress priming in reading and the selective modulation of lexical and sub-lexical pathways. *PLoS ONE, 4,* e7219.

Conrad, M., Carreiras, M., Tamm, S., & Jacobs, A. M. (2009). Syllables and bigrams: Orthographic redundancy and syllabic units affect visual word recognition at different processing levels. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 461–479.

Conrad, M., Stenneken, P., & Jacobs, A. M. (2006). Associated or dissociated effects of syllable frequency in lexical decision and naming. *Psychonomic Bulletin & Review, 13,* 339–345.

Frost, R. (1998). Towards a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99.

Halderman, L., Ashby, J., & Perfetti, C. (2012). Phonology: An early and integral role in identifying words. In J. Adelman (Ed.), *Visual word recognition, Volume 1: Models and methods, orthography and phonology* (pp. 207–228). East Sussex: Psychology Press.

Jacobs, A. M., & Grainger, J. (1994). Models of visual word recognition: Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1311–1334.doi:10.1037/0096-1523.20.6.1311

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 130–136. doi:10.1037/0096-1523.1.2.130

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

Jouravlev, O., & Lupker, S. L. (2014). Stress consistency and stress regularity effects in Russian. *Language and Cognitive Processes, 29,* 605–619.

Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 389–403.

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

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language, 47,* 145–171.

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

Monaghan, P., Arciuli, J., & Seva, N. (in press). Cross-linguistic evidence for probabilistic orthographic cues to lexical stress. In J. Thomson & L. Jarmulowicz (Eds.), *Linguistic rhythm and literacy. Trends in Language Acquisition Research series (TiLAR)*. Amsterdam: John Benjamins.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43–71.

Mundy, I. R., & Carroll, J. M. (2013). Spelling–stress regularity effects are intact in developmental dyslexia. *The Quarterly Journal of Experimental Psychology, 66,* 816–826.

Pagliuca, G., & Monaghan, P. (2010). Discovering large grain sizes in a transparent orthography: Insights from a connectionist model of Italian word naming. *European Journal of Cognitive Psychology, 22,* 813–835.

Paizi, D., Zoccolotti, P., & Burani, C. (2011). Lexical stress assignment in Italian developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 24,* 443–461.

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

Peressotti, F., & Colombo, L. (2012). Reading aloud pseudohomophones in Italian: Always an advantage. *Memory and Cognition, 40,* 466–482.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology, 61,* 106–151.

Perry, C., Ziegler, J. C., & Zorzi, M. (2014). CDP++. Italian: Modelling sublexical and supralexical inconsistency in a shallow orthography. *PLoS ONE, 9,* e94291.

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

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

Seva, N., Monaghan, P., & Arciuli, J. (2009). Stressing what is important: Orthographic cues and lexical stress assignment. *Journal of Neurolinguistics, 22,* 237–249.

Shulman, H. G., & Davison, T. C. B. (1977). Control properties of semantic codes in a lexical decision task. *Journal of Verbal Learning and Verbal Behavior, 16,* 91–98.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 744–774.

Sulpizio, S., Arduino, L. S., Paizi, D., & Burani, C. (2013). Stress assignment in reading Italian polysyllabic pseudowords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 51–68.

Sulpizio, S., Burani, C., & Colombo, L. (2015). The process of stress assignment in reading aloud: Critical issues from studies on Italian. *Scientific Studies of Reading.* doi:10.1080/10888438.2014.976340

Sulpizio, S., & Colombo, L. (2013). Lexical stress, frequency and stress neighborhood effects in the early stage of Italian reading development. *The Quarterly Journal of Experimental Psychology, 66,* 2073–2084.

Sulpizio, S., & Job, R. (2010). Effect of syllable frequency in speech production and visual word recognition: Evidence from Italian. *Giornale Italiano di Psicologia, 37,* 707–718.

Sulpizio, S., Spinelli, G., & Burani, C. (2015). Stress affects articulatory planning in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance.* doi:10.1037/a0038714

Thornton, A. M., Iacobini, C., & Burani, C. (1997). *BDVBD Una base di dati sul vocabolario di base della lingua italiana [BDVDB: A database for the Italian basic dictionary].* Roma: Bulzoni.

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

Wheat, K. L., Cornelissen, P. L., Frost, S. J., & Hansen, P. C. (2010). During visual word recognition, phonology is accessed within 100 ms and may be mediated by a speech production code: Evidence from magnetencephalography. *The Journal of Neuroscience, 30,* 5229–5233.

Yap, M. J., Balota, D. A., Cortese, M. J., & Watson, J. M. (2006). Single-versus dual-process models of lexical decision performance: Insights from response time distributional analysis. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1324–1344.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131,* 3–29.