University of Trento
Department of Cognitive Science and Education

Doctoral school in Psychological Science and Education
XXIV Cycle

Converging evidence on the autonomy and abstractness of the representation of lexical stress

Advisor: Prof. Remo Job
Co-advisor: Dr. Cristina Burani

PhD candidate: Simone Sulpizio

Academic year 2010-2011

To my father

who taught me the sound of words and the meaning of silence

Contents

Chapter 3

Lexical stress in reading aloud

Chapter 4

Priming lexical stress in reading Italian aloud

Chapter 5

Syllable frequency and stress priming interact in Reading aloud

Chapter 6

Overview

The aim of this thesis was to investigate the nature of word stress and its lexical representation. Lexical stress is a suprasegmental phenomenon: It pertains to the prosodic components, which differs from the phonemic-segmental component. Although the prosodic and the phonemic components pertain to two different levels, they work simultaneously.

The present work is focused on Italian, which is a polysyllabic language with free-stress position: In Italian, stress may appear on one of the last three syllables[1] of a word and, in the majority of cases, there are no rules that establish stress position. Although the three stress patterns are equally acceptable, they appear in different proportions within the lexicon (Thorton, Iacobini, & Burani, 1997). The Italian stress distribution is markedly asymmetrical: About 80% of three-syllable words bear stress on the penultimate syllable (*e.g.*, *paROla*, ‘word’; capital letters indicate the stressed syllable); about 18% of three-syllable words bear stress on the antepenultimate syllable (*e.g.*, *TAvolo*, ‘table’); and only 2% of three-syllable words bear stress on the ultimate syllable (*e.g.*, *coliBRÌ*, ‘hummingbird’). Only in the last case,

---

[1] In some inflected forms, we may find stress on the fourth last syllable (*e.g.*, *imMAginano*, ‘they imagine’).

stress is graphically marked. Stress position is not governed by rule. The only rule to assign stress to three-syllabic words refers to the weight of the penultimate syllable: If it is heavy – that is, if it ends with a consonant (*e.g.*, *biSONte*, bison) – then the syllable attracts stress (Krämer, 2009)[2]. However, in many cases the syllabic weight is not informative about the word's stress pattern and one must know which syllable has to be stressed. Therefore, observing the nature of word stress in Italian – no fixed position and no rules – we may assume that stress has to be lexically specified and that the stress pattern of a word is part of the knowledge stored in the lexicon.

The description of word stress in Italian paves the way to two basic questions: How is lexical stress represented? How does this representation intervene in perceiving or producing a word? In this thesis we address, at least in part, the latter questions in a series of experimental studies. To better understand representation and processing of word stress, we investigated lexical stress in different domains, *i.e.*, spoken-word recognition (Chapter 2) and reading aloud (Chapters 4 and 5). The results of our experiments shed new light on the role of lexical stress in these two linguistic processes, and on whether processing of lexical stress is similar in spoken-word recognition and reading aloud. In synthesis, lexical stress appears as part of the abstract prosodic knowledge stored in the lexicon: It pertains to the suprasegmental level of word representation and it is dissociable

---

[2] There are few exceptions to the rule (*e.g.*, *MANdorla* 'almond', *LEpanto*).

from the information pertaining to the segmental level, *i.e.*, the representation of a word's phonemes.

Lexical stress in spoken-word recognition

Chapter 1

1.1 Introduction

Lexical stress is an acoustic accentuation of a syllable within a word which leads to a relative prominence of the stressed syllable in the word. Listeners perceive the stressed syllable as more salient than the other syllables of the word. Stressed syllables differ from unstressed syllables both at the phonological level – stressed syllables tend to be heavier than unstressed syllables (Prince, 1990) – and at the phonetic level – stressed syllable physically differ from unstressed syllable in their acoustic correlates. Languages may be classified as having free stress – stress may appear on different syllables within a word, as in English or Italian – or having fixed stress – stress always appears in the same place within the word, as in French or Polish (Garde, 1972). The present work focuses on Italian and, more in general, on free-stress languages.

In Italian polysyllabic words, only one syllable in each word bears stress. Consider the following Italian words: *SEmino* (I sow) bears antepenultimate stress (stressed syllable is in upper case letters), with the first syllable being the more prominent; *seMIno* (small seed) bears penultimate stress and the second syllable is the most prominent; *semiNO* (he/she sowed) bears final stress, with the last syllable being the most prominent one. These three words (SEmino, seMIno, and semiNO) are identical at the segmental level and differ from each other only at the suprasegmental level[3]. This fact has potentially interesting consequences at the processing level as stress may bring useful information for lexical access: Listeners perceive the relative prominence of the stressed unit and they can use it to optimize the spoken-word recognition process.

In the present introductory chapter, we briefly address three related issues. First, what kind of acoustic information is related to lexical stress and what part of this information listeners exploit to detect stress. In a cross-linguistic perspective, we will note that native speakers of different languages exploit different types of acoustic information to detect lexical stress. Second, we discuss whether and how listeners use lexical stress in the recognition of spoken words. Anew, we will note that there is a difference in how listeners of different languages use lexical stress during word recognition. Third, we will briefly touch upon the theoretical debate on how words are accessed and stored in the lexicon. According to some theories, listeners can recognize spoken words through pre-lexical abstraction, which allows listeners to map different acoustic events into the same lexical representation. We will discuss whether pre-lexical abstraction may refer not only to the segmental features, but also to the suprasegmental features and thus to lexical stress.

---

[3] Note that minimal stress pairs exist in Italian (*e.g., ANcora - anCOra*, ‘anchor - again’), but they are few in number.

1.2 The perception of lexical stress

Lexical stress acoustically modifies the realization of a vowel, with stressed and unstressed syllables differing mainly on three acoustic cues: Amplitude, duration, and fundamental frequency (henceforth F0). All the three cues may contribute to the stress realization, thought in different ways (Fry, 1955; Ladefoged, Draper, & Whitteridge, 1958; Lieberman, 1960). At the perceptual level, the acoustic differences in amplitude (stressed vowels are higher in intensity), duration (stressed vowels are longer), and F0 (stressed vowels are louder) between stressed and unstressed vowels allow listeners to perceive which unit bears stressed within a word.

Although lexical stress is characterized by differences in amplitude, duration, and F0, different languages may rely on sub-sets of these acoustic cues to mark stress. Thus, some languages base the distinction between their stressed and unstressed syllable more on F0 differences, other languages more on duration differences, others more on amplitude differences. Moreover, in some cases, the selection of one or more cues to detect stress may also vary according to other features of the languages' phonological systems. In a tone language as Thai, for example, listeners perceive stress using duration alone (Potisuk, Gandour, & Harper, 1996), because F0 is used to realize tones.

In most cases the language specific cues are not rule based. To illustrate, in English, listeners' stress perception is mainly driven by F0 excursion (Fry, 1958), but also by syllable duration (Fry, 1955). Differently, in Dutch, stress perception is driven by duration (Reinisch, Jesse, & McQueen, 2010; Sluijter & van Heuven, 1996) and amplitude

(Sluijter & van Heuven, 1996). In Spanish, listeners perceive stress exploiting F0 and duration or F0 and amplitude (Llisterri, Machuca, de la Mota, Riera, & Rios, 2003).

As for lexical stress in Italian, recent research has shown that Italian listeners use duration to detect stress (Alfano, 2006; Alfano, Savy, & Llisterri, 2009). Stressed vowels are longer than unstressed vowels and this difference indexes the stress position. Although these studies have provided important results, their findings are partially incomplete, as these studies did not take into account the role of intensity. Whether amplitude plays a role or not in perceiving lexical stress in Italian has not been investigated yet, although linguists have claimed that intensity might be the main acoustic correlate of Italian stress (*e.g.*, Albano Leoni & Maturi, 1998). In synthesis, the relative roles of amplitude, duration, and pitch during stress recognition in Italian will be part of the investigation presented in Chapter 2.

1.3 Lexical stress in spoken word recognition

Spoken-word recognition starts with a pre-lexical processing in which listeners extract the acoustic-phonetic information from the speech signal. In such a way, the speech input activates those candidates that match, in their structures, the information in the signal (see, *e.g.*, McQueen, 2007). Inside this process, both the segmental and the suprasegmental acoustic-phonetic information may play the same role: Words that do not match in segments or stress are not taken into consideration as possible candidates. Accordingly, stress information could intervene during lexical selection. For example, consider the words *COdice* (codex) and *coDIno* (ponytail). Their first two syllables are identical at the segmental level, but the two words have antepenultimate and penultimate stress respectively, that is they differ at the suprasegmental level. The two syllables differ phonetically from each other and this difference may be exploited by listeners during word recognition. Thus, hearing the utterance *CO*- might activate words like *COdice* (codex), *COtica* (rind), and *COmico* (comic), but not words like *coDIno* (ponytail), *coROna* (crown), or *coLIte* (colitis). In this view, stress information may allow listeners to reduce the number of possible candidates that compete during word recognition.

Empirical data show that lexical stress does indeed play a role in lexical activation. Most studies addressing this issue used the cross-modal priming paradigm: Participants perform a task on a visual target preceded by an auditory prime. Using this paradigm with a lexical decision task, Cooper, Cutler and Wales (2002) investigated the role of lexical stress in English. Participants were presented with a visual target (*e.g.*, ADmiral) preceded by a two-syllable long spoken prime (*e.g.*, *admi*-). When prime and target had the same stress pattern, participants were faster in their responses. The same result was found using a one-syllable prime (*e.g.*, *mu*-) that could or could not match the target's stress (*e.g.*, MUsic or muSEUM). Similar results were obtained in Dutch (Cutler & Donselaar, 2001; Donselaar, Koster, & Cutler, 2005). Again, when the target word (*e.g.*, OCtopus 'octopus') was preceded by a two-syllable prime that matched the target's stress (*e.g.*, *OCto*-), then participants' responses were faster. The same facilitation effect was found using one-syllable primes (*e.g.*, *oc*-). Moreover, Donselaar and colleagues (2005) found that responses preceded by a mismatching stress prime (*e.g., ocTO-*) were slower than the control condition, but only if the prime was two-syllable long.

Studies conducted in romance languages show the same results. Spanish participants performed a lexical decision task on a target (*e.g.*, PRINcipe ‘prince’) preceded by a two-syllable auditory prime (Soto-Faraco, Sebastian Gallés, & Cutler 2001). Results showed that participants were faster in their response when the prime matched the target in its stress pattern (*e.g., PRINci-*), than in the control condition (a segmentally different prime fragment as *manti-*). Moreover, participants’ responses were slower when the target was preceded by a mismatching prime (*e.g., prinCI-*), than when it was preceded by the control fragment. Tagliapietra and Tabossi (2005) reported similar results in Italian: Listeners’ responses were faster when prime and target had the same stress pattern (*e.g., GOmi-* and GOmito ‘elbow’), than when prime and target had a mismatching stress (*e.g., goMI-* and GOmito ‘elbow’).

All the previous findings show that listeners of different languages use lexical stress during word recognition. However, while all the languages show the facilitation effect due to the prime and target stress’ congruency, results obtained in English do not show any inhibition effect when prime stress and target stress mismatch, with no difference between the mismatch stress condition (*adMI-* and ADmiral) and the control condition (Cooper *et al.*, 2002). This suggests that the contribution of lexical stress varies among languages. It has been argued that this difference is motivated by the likelihood that stress has to reduce the possible candidates during word-recognition. In Dutch and Spanish – and probably also in Italian – when listeners take into account lexical stress, they can greatly reduce the number of possible embedded competitors words (Cutler, Norris, & Sebastián-Gallés, 2004; Cutler & Pasveer, 2006). Differently, in English, the number of embedded competitors words that compete with the target is smaller than in other languages. In fact, in English unstressed syllables usually contain a reduced vowel and thus they differ from their phonetically identical stressed version that contain the full vowel. This means that stressed and unstressed syllables differ not only at the suprasegmental level, but also at the segmental level. This fact contributes to reduce the number of possible embedded words and, as a consequence, lexical stress becomes less important in word recognition (Cutler, 2005; Cutler & Pasveer, 2006).

The above reviewed studies have shown that listeners use lexical stress during spoken-word recognition. However, they did not investigate when exactly listeners exploit stress information during the recognition process. This issue was addressed by Reinisch and colleagues (Reinisch, Jesse, & McQueen, 2010) in Dutch. They used the printed word eye-tracking paradigm (Heuttig & McQueen, 2007; McQueen & Viebahn, 2007), displaying on the screen minimal pairs of words that were identical for the first two syllables, but different for the stress pattern (*e.g.*, OCtopus *vs.* okTOber). Participants heard a target word at the end of a carrier sentence and they had to click the mouse button on the right printed word. Reinisch *et al*. (2010) found that listeners were able to select the correct word (*e.g.*, okTOber) and discard its competitor (OKtopus) before the beginning of the third syllable, when the diverging segmental material occurs. This result shows that participants could use stress information as soon as it becomes available. A similar study using the same paradigm will be presented in Chapter 2, where we investigate when Italians use lexical stress during spoken-word recognition and whether stress information is used as soon as it becomes available to optimize the spoken-word recognition process.

1.4 How much prosodic information is abstract and stored?

During spoken word recognition listeners extract phonetic information from the acoustic signal and map this information onto cognitive representations (McQueen & Cutler, 2010). Different theoretical approaches converge on the existence of a mental lexicon, which contains a variety of information in its entries, such as phonological information, morphological structure, semantic and syntactic information. However, there is no agreement on how words in the mental lexicon are accessed and stored.

On the one hand, some theories postulate that the mental lexicon is composed of episodic traces. Each word would be represented by multiple different traces that consist of detailed acoustic representations of episodic encounters with those words (Goldinger, 1996, 1998; Pierrehumbert, 2002). On the other hand, other theoretical positions assume that the mental lexicon is composed of phonologically abstract forms that are accessed through a prelexical stage of computation, matching the acoustic signal with stored lexical knowledge (McClelland & Elman, 1986; Gaskell & Marslen-Wilson, 1997; Norris & McQueen; 2008). Both approaches show some limits in explaining experimental data: Episodic models cannot explain prelexical abstraction about speech segments (see below, McQueen, Cutler, &

Norris, 2006), whereas abstractionist models cannot explain evidence that episodic details are maintained in long-term memory (Goldinger, 1998). However, the more recent models try to take into account both the abstract and the episodic components. These hybrid models (Cutler, Eisner, McQueen, & Norris, 2010; Goldinger, 2007) propose abstract prelexical and lexical representations combined with an episodic memory component that can deal with those aspects more related to specific episodes (as, for example, talker variability).

In the framework of the hybrid models, an important question to ask concerns the division of labor between the abstractionist and the episodic components in the word-recognition process. We need to clarify what information listeners have stored as abstract knowledge and what information they have stored as episodic knowledge. At this regard, a crucial study was conducted by McQueen and colleagues (2006). Using a perceptual learning paradigm, they trained participants in a lexical decision task in which listeners heard an ambiguous [f-s] sound replacing an [f] or [s] within words. Then, listeners heard an auditory prime and performed a lexical decision task on a visual target. The critical material was composed of minimal pairs that could be a word both with [f] or [s] (*e.g.*, *knife* or *nice*). Neither of these words had been heard during training. Results showed that listeners trained to interpret the ambiguous sound as [f] tended to interpret the ambiguous words as containing an [f] sound (*e.g.*, [doo?] was interpreted as *doof*); on the opposite, listeners trained to interpret the ambiguous sound as [s] tended to interpret the ambiguous words as containing an [s] sound (*e.g.*, [doo?] was interpreted as *doos*). These results show that spoken-word recognition is based on pre-lexical representations that are abstract and flexible: Speech signal is initially processed in terms of pre-lexical abstract representations that listeners use to perform lexical access (see also, Cutler, 2008; Kouider & Dupoux, 2005; Norris, McQueen, & Cutler, 2003; Obleser & Eisner, 2009).

But what about suprasegmental information? Is it stored as abstract knowledge – as well as segmental information – or not? A first study conducted in Dutch (Shatzman & McQueen, 2006) has shown that listeners have abstract knowledge about syllable duration – syllables with longer duration tend to be interpreted as monosyllabic words, while syllables with shorter duration tend to be interpreted as the initial syllables of polysyllabic words – and they are able to use this knowledge when recognizing new words. The first evidence obtained in Dutch enforces us to go further in this direction to explore further what kind of word's prosodic knowledge is abstract and available at the pre-lexical level. Lexical stress in Italian may be a good test. The asymmetric distribution of the two main Italian stress patterns (80% of words bear penultimate stress whereas 18 % of words bear antepenultimate stress) allows to test whether listeners have stored knowledge about lexical stress not only at the acoustic level – which cues listeners use to recognize the word's stress – but also at the distributional level: Listeners might consider penultimate stress as a default and they would only detect the antepenultimate stress by using the phonetic cues they found in the signal. Thus, lexical stress might be part of the abstractionist component that is at work during spoken-word recognition.

To conclude, the investigation of lexical stress involves more than one dimension: It is related to the discovery of the acoustic-phonetic cues that allow us to perceive stress; it may help to understand how listeners recognize words; it might reveal that listeners have abstract knowledge about lexical stress. Thus, investigating how Italian listeners use lexical stress during spoken-word recognition may help to investigate three related questions. First, when do Italian listeners make use of lexical stress information during spoken-word recognition? Second, does the distributional bias affect the recognition process? Third, which acoustic cues do Italian listeners pick up to detect the word’s stress pattern, and how do these cues interact with the distributional bias? The answer to these questions will allow us to shed new light on the nature of lexical stress in Italian, and, more generally, on the nature of lexical representation: Is lexical stress knowledge – both at the acoustic and distributional level – stored abstractly, and is it able to assist listeners during spoken-word recognition? Some of these issues will be addressed in Chapter 2, where, in two eye-tracking experiments, we explored how Italian listeners use lexical stress in recognizing spoken words and whether the recognition process is affected by stored prosodic knowledge that listeners have on lexical stress.

Italians use abstract knowledge about lexical stress during spoken-word recognition

---

Chapter 2

Sulpizio, S. & McQueen, J. M. (in press). Italians use abstract knowledge about lexical stress during spoken-word recognition. *Journal of Memory & Language.*

In two eye-tracking experiments in Italian, we investigated how acoustic information and stored knowledge about lexical stress are used during the recognition of tri-syllabic spoken words. Experiment 1 showed that Italians use acoustic cues to a word's stress pattern rapidly in word recognition, but only for words with antepenultimate stress. Words with penultimate stress – the most common pattern – appeared to be recognized by default. In Experiment 2, listeners had to learn new words from which some stress cues had been removed, and then recognize reduced- and full-cue versions of those words. The acoustic manipulation affected recognition only of newly-learnt words with

antepenultimate stress: Full-cue versions, even though they were never heard during training, were recognized earlier than reduced-cue versions. Newly-learnt words with penultimate stress were recognized earlier overall, but recognition of the two versions of these words did not differ. Abstract knowledge (*i.e.,* knowledge generalized over the lexicon) about lexical stress – which pattern is the default and which cues signal the non-default pattern – appears to be used during the recognition of known and newly-learnt Italian words.

Chapter 2

Introduction

As listeners recognize spoken words, they must combine acoustic-phonetic information in the speech signal with stored knowledge about the sound patterns of words. This much is uncontroversial. But which sources of information do listeners rely on, what knowledge do they have about how words sound, and when do they integrate information that has been extracted from the speech signal with stored knowledge? We ask here when and how Italian listeners recognize polysyllabic Italian words that differ in their stress patterns. Answers to these questions provide constraints on the nature of the lexical access process, and on the nature of the knowledge stored in the mental lexicon.

How words are accessed and stored in the lexicon is a matter of ongoing debate. Two extreme theoretical positions can be defined. According to the first approach, the mental lexicon consists of episodic traces. Each word is represented by multiple traces that consist of detailed acoustic representations of episodic encounters with those words (Goldinger, 1998; Pierrehumbert, 2002). Word recognition entails comparison of the current acoustically detailed input with those stored traces. There thus needs to be no phonological abstraction prior to lexical access. The second approach assumes that the mental lexicon contains phonologically abstract forms (McClelland & Elman, 1986; Gaskell & Marslen-Wilson, 1997; Norris & McQueen, 2008). Word recognition again entails comparison of the current input with stored lexical knowledge, but this requires a prelexical stage of phonological abstraction so that contact can be made with the abstract representations in the lexicon.

Neither of these extreme positions is tenable. Strictly episodic models cannot explain evidence of prelexical abstraction about speech segments (McQueen, Cutler & Norris, 2006), and strictly abstractionist models cannot explain evidence that episodic details are maintained in long-term memory (Goldinger, 1998). What is required, therefore, is a hybrid model with both episodic and abstractionist components (Cutler, Eisner, McQueen, & Norris, 2010; Goldinger, 2007). An important question to ask, therefore, is what the division of labor is between these two components in the word-recognition process. For example, do listeners have abstract knowledge not only about speech sounds (McQueen *et al*., 2006) but also about the prosodic structure of words (that is, about their lexical stress patterns and about other aspects of lexical prosody)? Is that knowledge the result of forming generalizations over the lexicon? Furthermore, can listeners use that knowledge during the lexical access process? We asked these questions here, with respect to knowledge about stress in Italian words.

Italian offers an especially interesting test of whether abstract prosodic knowledge is used in word recognition because it has a strongly asymmetrical distribution of lexical stress patterns. Consider three-syllable words. There are two main stress types (Krämer, 2009): An antepenultimate stress pattern (*i.e.*, the first syllable bears stress, *e.g.*, TAvolo ‘table’; capital letters indicate stress), and a penultimate stress pattern (*i.e.*, stress appears on the second syllable, *e.g.*, coLOre ‘color’). The only rule to assign stress in trisyllabic words refers to the weight of the penultimate syllable: If it is heavy – that is, if it ends in a consonant – then it must be stressed (Krämer, 2009). Nevertheless, there is a strong distributional bias toward the penultimate stress pattern. In fact, 80% of Italian tri-syllabic words have penultimate stress, 18% have antepenultimate stress, and 2% have stress on the last syllable (*e.g.*, serviTU, 'servitude'; Thornton, Iacobini, & Burani, 1997). This distributional asymmetry may be reflected in how Italians recognize spoken words. If they have abstracted the knowledge (generalized over the relevant entries in the Italian lexicon) that a trisyllabic word will usually have penultimate stress, then they may assume (in the absence of evidence to the contrary) that this is the stress pattern of any trisyllabic word they hear. This assumption that there is a default stress pattern may apply both when Italians are recognizing known Italian words, and when they are recognizing newly-learnt words. We tested both these possibilities in the present experiments.

Prior research has already indicated that Italian listeners are sensitive to lexical stress information (Tagliapietra & Tabossi, 2005). In a cross-modal priming paradigm, listeners performed a lexical decision task on visual targets preceded by spoken bi-syllabic primes. Responses were facilitated when the target (*e.g.*, GOmito, 'elbow') was preceded by a fragment-prime with the same stress pattern (*e.g.*, GOmi), in line with previous findings for Dutch (Cutler & Van Donselaar, 2001; Van Donselaar, Koster, & Cutler, 2005) and Spanish (Soto-Faraco, Sebastian-Galles, & Cutler, 2001). Italian listeners thus appear to use lexical stress cues to recognize spoken words. It is not clear, however, how early in the recognition process knowledge and information about stress in Italian are brought to bear. Dutch listeners use stress information very early (*i.e.*, in words that are segmentally identical in their initial syllables, such as OCtopus, 'octopus', and okTOber, 'October', stress information is used prior to the segmental disambiguation point; Reinisch, Jesse & McQueen, 2010). Since in Italian, as in Dutch, the difference between stressed and unstressed syllables is at the suprasegmental rather than the segmental level, we expect that Italian listeners can also take advantage of stress cues early in the recognition process. An open question, however, is whether the distributional bias toward the penultimate stress pattern in Italian can affect the earliest stages of word recognition.

Furthermore, although Tagliapietra and Tabossi's (2005) findings suggest that the word-recognition process in Italian benefits from stress information, it remains unclear what exactly that information is. Which acoustic cues specify the stress patterns of Italian words? In general, stressed vowels differ acoustically from unstressed vowels in pitch, duration, and intensity (Albano Leoni & Maturi, 1998). But it is not clear which of these acoustic cues Italian listeners pick up on. Some authors consider amplitude to be the main stress correlate (Albano Leoni & Maturi, 1998). Others argue that duration plays the main role (Alfano, 2006; Alfano, Savy, & Llisterri, 2009). An additional aim of the present study was therefore to establish which stress cues Italian listeners use during word recognition. We were especially interested in whether the bias toward the penultimate syllable stress pattern modulates the way the acoustic information that signals stress is processed. In fact, if Italian listeners have stored knowledge about the acoustic correlates of stress and about the asymmetrical distribution of the two stress patterns, then it is possible that their use of acoustic information about stress may also be asymmetric. In particular, they should be more sensitive to the acoustic cues specifying an antepenultimate stress pattern than to those specifying the penultimate pattern – because the latter pattern can be assumed to occur by default.

In summary, the present study investigated three related questions. First, when do Italians use knowledge and information about lexical stress in spoken-word recognition? Second, how does the distributional bias favoring penultimate stress in Italian affect the recognition process? Third, which acoustic cues are picked up by Italians as they detect stress position, and how do these cues interact with the distributional bias? Answers to these questions should inform the debate on the nature of lexical representation. Is lexical stress knowledge stored in an abstract way (*i.e.*, are there generalizations made across the Italian lexicon), and is that knowledge available to assist in word recognition?

To address these questions, we examined how Italian listeners use lexical stress to recognize known and newly-learnt words. In Experiment 2, an artificial-lexicon study, we examined recognition of newly-learnt words. This allowed us to control for the amount of exposure to specific episodes of those words and test whether prior knowledge about prosodic structure (abstracted from earlier experience with real Italian words) can nonetheless be brought to bear during word recognition. Shatzman and McQueen (2006) used the same paradigm to test whether Dutch listeners have abstract prosodic knowledge about syllable duration and whether they can use it in the recognition of new words. Shatzman and McQueen trained participants to associate spoken non-words with novel shapes (displayed on a computer screen). The critical materials were pairs of monosyllabic non-words (*e.g.*, *bap*) and bisyllabic non-words which had the same syllable embedded in onset position (*e.g.*, *baptoe*). The initial syllables in each pair (*e.g.*, *bap*) had the same ambiguous duration during the training phase of the experiment. In the subsequent test phase, syllable duration was manipulated: It was longer, shorter or equal to the duration used during training. The results showed that participants tended to interpret shorter syllables as bisyllabic word onsets and longer syllables as monosyllabic words, as indeed tends to be the case in real Dutch words (Salverda, Dahan & McQueen, 2003), even though the participants had heard the novel words with only ambiguous durations during the training phase. Dutch listeners thus appear to have abstract prosodic knowledge about syllable duration and they appear to be able to use this knowledge during the recognition of newly-learnt words. Experiment 2 is based on Shatzman and McQueen (2006). We test there whether Italian listeners have abstract prosodic knowledge about lexical stress (about which pattern is the default and about the cues which specify a word's stress pattern) and whether they can use this knowledge to improve their ability to recognize novel words. Experiment 2 thus provides the critical test of whether stress knowledge in Italian is abstract.

In Experiment 1, however, we first use real words to examine when Italian listeners use lexical stress information in spoken-word recognition and whether the distributional bias favoring penultimate stress affects the recognition process. Moreover, we investigated which acoustic cues Italians used to detect stress. The answers to these questions provide the basis for the further investigations in Experiment 2. Before asking if Italians use abstract knowledge about lexical stress in recognizing new words, we have to establish whether this knowledge exists and, if so, how it is used in the recognition of known words.

Experiment 1

In Experiment 1 we used the printed-word eye-tracking paradigm (Huettig & McQueen, 2007; McQueen & Viebahn, 2007). Italian listeners heard spoken target words (*e.g., CAnapa*, ‘hemp’) and had to identify the printed forms of those words on a computer screen from among an array of four alternative words which included a competitor with overlapping onset segments but a different stress pattern (*e.g., caNAle*, ‘channel’). Previous findings with this paradigm have shown that Dutch listeners use stress information as soon as it becomes available: The listeners preferred to fixate the targets before their spoken forms diverged segmentally from the competitors (Reinisch *et al.*, 2010). We assume that a similar pattern of results will emerge for Italians. Two reasons make it plausible that Dutch and Italian will be treated similarly. First, lexical stress does not modify the segmental material in either language. In particular, unstressed vowels are not reduced (as occurs, *e.g.*, in English). Second, studies conducted in these two languages using the identity priming paradigm have shown similar results: Listeners benefit from stress information during word recognition (for Italian, see Tagliapietra & Tabossi, 2005; for Dutch, see Cutler & Van Donselaar, 2001; Van Donselaar *et al*., 2005).

We thus hypothesize that Italian listeners will use stress information to constrain lexical access as soon as that information becomes available. Moreover, we expect that the asymmetry in the penultimate and antepenultimate stress distribution will affect word recognition. We hypothesize that listeners have knowledge about this distributional asymmetry and that they use this knowledge to optimize word recognition. If listeners know, when they hear a trisyllabic word, that it will usually have a penultimate-stress pattern (*i.e.*, 80% of the time), then they can consider this pattern as a default. Acoustic cues may therefore not play a large role in the recognition of penultimate-stress words. In contrast, antepenultimate stress detection may be driven by the acoustic cues in the speech signal. Listeners could use this information to detect antepenultimate stress and hence to discard the default pattern. Acoustic cues may therefore be more important in the recognition of antepenultimate-stress words than in the recognition of penultimate-stress words.

In summary, we tested the following predictions. First, listeners should use the words' stress patterns to disambiguate segmentally identical fragments. They should tend to fixate targets (*e.g., CAnapa*) and tend to ignore their segmentally overlapping competitors (*e.g., caNAle*) before segmental disambiguation (*e.g.*, the /p/ of *CAnapa*) is available. Second, the distributional bias in Italian should affect performance. If Italians indeed assign penultimate stress by default, they should need to use acoustic cues to stress actively only when recognizing words with antepenultimate stress. Acoustic markers of stress should thus correlate with eye-movement behavior only for antepenultimate-stress targets. Testing this latter prediction should also allow us to identify which acoustic cues drive antepenultimate-stress detection; that is, we should be able to establish whether Italian listeners depend more on duration (Alfano, 2006; Alfano *et al.*, 2009) or on amplitude (Albano Leoni & Maturi, 1998) in stress recognition.

Method

Participants

Thirty-two students (mean age: 26.3, sd: 6.2) from the University of Trento took part in the experiment. They received course credit for their participation. All participants were Italian native speakers with no known hearing problems and normal or corrected-to-normal vision.

Materials

Thirty-two pairs of trisyllabic words were selected as experimental targets (see Appendix). The words in each pair were segmentally identical in their first two syllables, but they differed in stress location. One word in each pair had stress on the penultimate (second) syllable and the other had stress on the antepenultimate (first) syllable. All pairs could be segmentally distinguished at the beginning of the third syllable (*e.g.*, *CAnapa* and *caNAle*). Thirty-two distractor pairs were then selected, each coupled to one of the experimental stress pairs. Words in each distractor pair overlapped orthographically and phonologically on their first two syllables, and they did or did not differ in stress pattern (*e.g.*, *GEnero* 'son-in-law', *GEnesi* 'genesis' for a stress-matched pair; *RUGine* ‘rust’, *rugGIto* ‘roar’ for a stress-mismatched pair). There were no semantic relationships among the four words in each set (*i.e.*, an experimental pair plus a distractor pair). Twelve additional pairs were selected to be used in practice trials. Stress pairs as well as distractor pairs were matched on frequency (both t’s < 1) (CoLFIS database, Bertinetto *et al*., 2005), length in syllables (all were trisyllabic) and length in letters (t (31) = 1.13, p = .14). Acoustic measures of the stimuli are given below in Table 1.

A female native Italian speaker, naïve about the experiment’s purpose, recorded the stimuli in a sound-attenuated room (sampling at 44 kHz, 16 bit resolution, mono). Each word (*i.e.*, each member of each of the 32 sets) was spoken at the end of the sentence "Clicca sulla parola" ("Click on the word"), with the sentence accent on the target word. Speaking rate was measured as the ratio per stress-type condition between total sentence duration and the number of syllables in the sentence. These ratios were matched across conditions (penultimate stress: 5.59 syllables per second; antepenultimate stress: 5.61 syllables per second; $t < 1$).

Procedure

Participants were seated about 50 cm in front of a computer screen (screen size 360 mm x 270 mm). The experiment had two parts: A familiarization task followed by the main eye-tracking experiment. During the first part, participants were familiarized with the stimuli. Because stimuli were all low-frequency words, all 64 experimental and distractor words were shown in lower-case letters in the middle of the screen, one by one in random order. Participants had to read them aloud: No word was found to be unknown to any participant, and all participants performed the task very well.

After this familiarization task, the eye-tracking experiment was run. Eye movements were recorded using a head-mounted Eyelink II System, at a sampling rate of 500 Hz, recording both eyes. The experimental section was composed of four blocks of 32 displays, each combined with a spoken instruction. In each display, four printed words were shown, one pair of experimental words plus one pair of distractor words. Each display of four words was shown in each block. Across blocks, different words from within each set of four were targets (*i.e.*, were the words mentioned in the spoken instructions). In the first block, however, only words from the experimental stress pairs were selected as targets; half of them had penultimate stress and half had antepenultimate stress. In the subsequent blocks, the target could be the same word that was seen in the first block, its experimental competitor, or one of the distractor words. In this way, participants could not know which word they would hear when a given display was presented in any given block, because all four alternatives could occur as targets. Block order was counterbalanced across participants, and within each block trial order was randomized. The experiment was preceded by a small practice session using six displays; each display was shown two times, for a total of twelve practice trials. There were no breaks between the blocks.

Each trial started with a fixation cross in the center of the screen, displayed for 500 ms. Four words then appeared on the screen and remained there either until participants clicked the mouse button or for a maximum of 5000 ms. A white screen was used during the inter-stimulus interval of 480 ms. All words were presented in lowercase Lucida Sans Typewriter font, size 20. The four words were centered in the four quadrants of the screen. The auditory instructions (*i.e.*, the carrier sentence plus each target word, *e.g.*, "Clicca sulla parola *canapa*") were played over headphones; the instructions began at the same time as the printed words appeared. Participants had to click the mouse on the target word that they heard at the end of the carrier sentence. Every eighth trial there was a drift correction to adjust for possible small head movements.

Results

Three analyses of fixation behavior were performed. First, to test whether lexical stress information and/or knowledge is used early to distinguish between possible candidate words, a comparison between target and competitor fixations within each stress pattern (penultimate and antepenultimate) was run. Second, to test whether there was a distributional bias in the data, we performed an analysis comparing performance across the two stress patterns. Third, correlations between acoustic measures and behavioral data were run to establish which acoustic cues, if any, were used by listeners to detect the words' stress patterns, and to ascertain whether this information was mainly used in the recognition of words with antepenultimate stress (*i.e.*, the words with the non-default pattern), and less so (or not at all) in the recognition of penultimate-stress words.

Only trials in which participants clicked on the correct word were considered in the analyses (1% of all the trials were discarded for this reason). If a target was repeated during the experiment, only data from its first presentation were used. We considered fixations on a word as being all those that fell within a 6.3 cm square centered of the middle of each word: Thus, each fixation was coded as pertaining to the target, to the competitor, or to one of the two distractors. The proportion of fixations to each word over time (in 10 ms time intervals) was computed in each condition, by summing the number of fixations to each type of word and dividing it by the total number of fixations in the same time interval.

In all eye-tracking analyses, time windows were defined considering a delay of 200 ms as an estimate of the time needed to program and launch a saccade (Matin, Shao, & Boff, 1993). Thus, for example, when considering fixations in response to the first syllable of the words, a time window was defined as starting 200 ms after the acoustic onset of the syllable and ending 200 ms after the syllable's acoustic offset. Figure 1 shows fixations on target, competitor and the two distractors over time for each stress pattern.

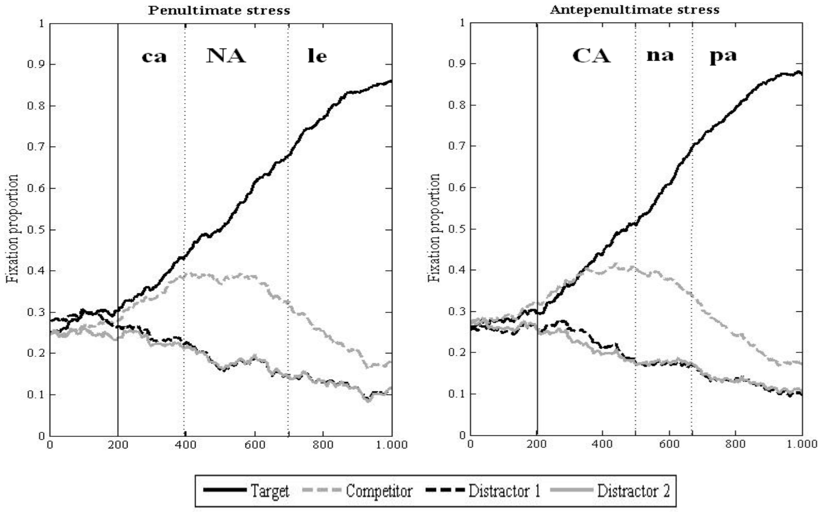

Figure 1. *Experiment 1:* Fixation proportions to targets, competitors, and distractors over time (in ms on the abscissa). The solid vertical lines show the beginnings of the time windows starting 200 ms after the words' average onsets; the dotted lines indicate the ends of the time windows aligned to the average offsets of the first and the second syllables respectively, each again delayed by 200 ms.

*Comparing target and competitor fixation*s

Within each stress condition, a mixed-effects analysis (Baayen, Davidson, & Bates, 2008) was performed comparing fixation proportions on target and competitor words (*e.g.*, *canapa* and *canale* when the spoken target was *canapa*). Fixation proportions were log transformed (Barr, 2008). Participants and items were treated as random factors, and stimulus type (target *vs*. competitor) was treated as a fixed factor. Models were fitted using R software (version 2.11; The R foundation for statistical computing) and *p*-values were calculated using the MCMC procedure, sampling 10,000 times (Baayen *et al*., 2008).

We ran three separate analyses considering the following time windows: First syllable, first 1.5 syllables, and first two syllables. Figure 2 shows fixations proportions for targets and competitors in each of these three time windows. The first syllable and first two syllable windows were defined relative to the acoustic syllable boundaries, offset by 200 ms (*i.e.*, 200-396 ms and 200-699 ms for penultimate-stress words, respectively, and 200-499 ms and 200-669 ms for penultimate-stress words). But the 1.5 syllable window was defined in absolute terms, and thus was of the same fixed length in both stress conditions (200-566 ms). The average duration of the first syllable (of both types of word) plus half of the average duration of the second syllable (again of both types of word) was 366 ms. The use of this time window thus allowed us to control for the differences across stress conditions in syllable duration, and thus also equated the amount of data used in the analysis in each condition. Although these syllable-duration differences are already controlled in the current within-item comparisons (target *vs*. competitor within stress type), they are not in the subsequent comparisons across stress types. Analyses of behavior in the 1.5 syllable window (along with those for the first syllable alone) also allowed us to ask whether eye movements were modulated by stress cues alone (*i.e*., before effects of the first consonant of the word's third syllable could influence behavior).

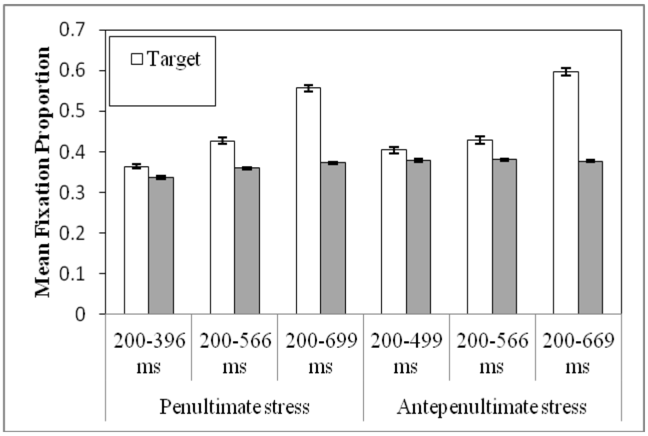

Figure 2. *Experiment 1*: Mean fixation proportions to targets and competitors in each time window: the first syllable (200-396 ms or 200-499 ms, respectively, for penultimate and antipenultimate stress words), the first syllable plus half of the second syllable (200-566 ms for both types of word), and the first and second syllables (200-699 ms or 200-669 ms, respectively, for penultimate and antipenultimate stress words). Error bars are standard errors.

In the analyses of fixations in response to the first syllable, no differences between target and competitor were found (for penultimate stress, t < 1; for antepenultimate stress, t = 1.01). The analysis on fixations in response to the first 1.5 syllables revealed a difference between target and competitor fixations, for words with penultimate stress (β = 0.346, t = 5.91, p < .01) and for words with antepenultimate stress (β = 0.467, t = 6.96, p < .01). Participants looked at the target more than at the competitor before segmentally disambiguating information (at the onset of the third syllable) became available. The analysis on fixations in response to the first and second syllables showed the same pattern for both penultimate (β = 0.597, t = 10.87, p <.01) and antepenultimate stress (β = 0.466, t = 8.39, p < .01).

Comparing penultimate- and antepenultimate-stress words

A mixed-effects analysis was run to see whether there was a difference in the proportion of fixations between penultimate- and antepenultimate-stress targets. Fixation proportions were again log transformed. The time windows included the first 1.5 syllables (as already noted, this window of fixed duration controls for durational differences across stress types) and the first two syllables (where amount of information in terms of the number of segments is controlled, and the durational difference across stress types is only 30 ms on average). Analyses of the first syllable alone were not included because of the large durational difference between antepenultimate- and penultimate-stress words (103 ms on average). Fixations on target words were used as the dependent variable, with stress type (penultimate *vs*. antepenultimate) as fixed factor and participants and items as random factors. No effect of stress type was found in either time window (t's <1). To test whether the amount of competition varied across stress patterns, we conducted further analyses using the difference in the proportion of fixations to target and competitor as dependent variable and the stress type (penultimate *vs*. antepenultimate) as fixed factor. Again, no effect of stress type was found in either time window (t's <1).

Correlation analyses

Acoustic measures of the first two vowels of the target words were performed in order to explore which information listeners used to determine the words' stress patterns. For the first and the second vowel of each target word, we measured pitch (in Hz), duration (in ms), amplitude (in Pascal), and spectral tilt (calculated as a ratio between energy in high and low frequency band; see Cutler, Wales, Cooper, & Janssen, 2007). These analyses revealed that, in antepenultimate-stress words, the first vowel (*i.e.*, the stressed vowel) was longer, higher, and louder, and had more high frequency energy than the unstressed second vowel. In contrast, in the penultimate-stress words, the second (stressed) vowel was longer but it was also lower and weaker than the unstressed initial vowel, and the spectral tilt of the two vowels did not differ (see Table 1).

Antepenultimate stress

	First vowel	Second vowel	t (29)	p value
Duration (ms)	165	81	16.14	<.01
Pitch (Hz)	219	177	10.54	<.01
Amplitude (Pascal)	.09	.04	8.03	<.01
Spectral Tilt	.7	.03	5.19	<.01

Penultimate stress

	First vowel	Second vowel	t (29)	p value
Duration (ms)	75	180	-18.2	<.01
Pitch (Hz)	238	204	4.43	<.01
Amplitude (Pascal)	.09	.06	4.42	<.01
Spectral Tilt	.3	.3	<1	n.s.

Table 1. Mean acoustic measures and t-test comparisons for the first and second vowel of the words with each stress pattern in Experiment 1
Note. Spectral tilt is expressed as a unitless ratio.

Two types of correlations were then performed comparing the acoustic measures of the first and second vowels with fixation behavior: within and between stress types. The within stress-type comparison indicates whether listeners used the differences between the vowels within words to detect the words' stress pattern. The between stress-type comparison provides an index of whether use of cues in the recognition of penultimate-stress words differs from that in the recognition of antepenultimate-stress words. In both cases, only significant correlations are reported.

Within stress types. For each stress pattern, correlations were performed on difference measures: For each acoustic measure, the difference between fixation proportions on the stressed and unstressed vowels was compared to the acoustic difference between the first and second vowels. A significant correlation was found only between the behavioral data for antepenultimate-stress targets and the amplitude difference between the vowels of those words (r = .46, t (29) = 2.57, $p<.05$). As the difference between the first and second vowels became larger, listeners looked more at the target words. In a backward regression model, with fixation difference as dependent variable and the acoustic difference measures as predictors, amplitude was the only significant predictor in the model: t (29) = 2.57, $p<.05$. ($R2$ = .186, adjusted $R2$ = .158).

Between stress types. Correlations were also performed comparing the difference in fixation proportions between antepenultimate- and penultimate-stress words to the differences in the acoustic measures between antepenultimate- and penultimate-stress words. Based on the earlier analyses on fixation proportions, the first and second syllable time window was chosen. Behavioral and acoustic measures of the target words (*e.g.*, *CAnapa*) were subtracted from the respective measures of their competitor words, that is, the words with the opposite stress pattern (*e.g.*, *caNAle*). The data showed a significant correlation between the difference in fixations between the antepenultimate-stress targets and their penultimate-stress competitors and the corresponding difference in spectral tilt (r = -.46, t (29) = -2.81, p<.01). As the difference in spectral tilt between the second vowels of penultimate- and antepenultimate-stress words decreased, listeners looked less to antepenultimate-stress targets. A backward regression model with the fixation difference between antepenultimate targets and their competitors as dependent variable and the acoustic difference measures as predictors revealed that spectral tilt was the only significant predictor: t (29) = -2.18, p<.01 (R2 = .215, adjusted R2 = .188).

Discussion

The results of Experiment 1 show that Italians use acoustic cues to lexical stress and lexical-stress knowledge during spoken-word recognition. In line with results obtained in Dutch (Reinisch *et al*., 2010), Italian listeners use the acoustic information about stress in the speech signal as soon as it becomes available. They can thus distinguish between two trisyllabic words with segmentally-identical first and second syllables but different stress patterns (*e.g.*, *CAnapa* and *caNAle*) before the segmental disambiguation (the [p] or the [l]) is available to them.

The correlation analyses, however, suggest that Italians pick up on acoustic cues only when detecting antepenultimate stress: They used primarily intensity information (Albano Leoni & Maturi, 1998) to identify antepenultimate stress patterns, but appeared not to use acoustic cues when recognizing words with penultimate stress (despite the presence of such cues in the speech signal). The primary acoustic signal that listeners detect appears to be a marked decrease in amplitude of the second vowel compared to the first vowel of words with antepenultimate stress (note that there is a smaller decrease in amplitude across the first two vowels in words with penultimate stress, see Table 1). It might be assumed that listeners could use a further criterion, based on the amplitude of the first syllable: If the amplitude is higher than a threshold value, listeners could start to assume the word has antepenultimate stress. The analyses on fixations in response to the first syllable, however, did not show that the competition between target and competitor was already resolved at this point in time. Furthermore, as Table 1 shows, there is no difference in mean first syllable amplitude across conditions. These observations suggest that first syllable amplitude alone is not enough to recognize antepenultimate stress. But it is possible that listeners might use both criteria – the amplitude of the first vowel and the amplitude difference between the first and second vowels – to identify that a word has antepenultimate stress.

Even though there was no evidence of listener sensitivity to the acoustic cues signaling the penultimate stress pattern, penultimate-stress words were recognized just as quickly as antepenultimate-stress words. This suggests that Italians were using knowledge that penultimate stress is the much more frequent pattern, and so were recognizing penultimate-stress words by default. In short, it appears that Italian listeners assume that trisyllabic words will have stress on their penultimate syllables and hence will recognize sequences such as *cana*- as being the onset of *canale*, unless acoustic evidence (primarily a more marked decrease in amplitude in the second vowel relative to the first vowel) indicates that the antepenultimate syllable is stressed, and hence that they must be hearing *canapa*.

These findings thus indicate that Italians have knowledge about the stress-pattern distribution in the Italian lexicon. They know that penultimate stress is the most frequent pattern in trisyllabic Italian words, and they exploit this knowledge to optimize word recognition. They assign this more frequent pattern by default, and detect words with antepenultimate stress using the intensity information contained in the signal. To further test Italians’ knowledge about the use of this distributional bias, and the interaction between this knowledge and the use of acoustic cues to stress, we ran a second experiment using an artificial lexicon. Critically, the use of newly-acquired words allowed us to test if stored prior knowledge about lexical stress can be used by Italians when the number of exposures to those new words was controlled, and hence whether that knowledge is abstracted away from memories of specific lexical episodes.

Experiment 2

In Experiment 2 we investigated how listeners use prosodic knowledge about stress to recognize newly-acquired words. We tested whether Italians apply their stored knowledge about default stress patterns and about the acoustic cues to stress when they are recognizing words that they have never heard before the experiment began. As discussed above, Shatzman and McQueen (2006) found evidence that Dutch listeners use stored prosodic knowledge about word duration to recognize newly-learnt words. In keeping with this finding, we hypothesize that Italian listeners will use prior knowledge about lexical stress when recognizing new words. Following the distributional bias, Italian listeners may assign penultimate stress by default and identify only novel words with antepenultimate stress on the basis of the acoustic information in the speech signal.

To test these assumptions, we used an artificial-lexicon eye-tracking paradigm (Creel, Tanenhaus, & Aslin, 2006; Magnuson, Tanenhaus, Aslin, & Dahan, 2003; Shatzman & McQueen, 2006). Previous studies have shown that lexical access with an artificial lexicon works in a similar way to lexical access with a real lexicon: Participants' eye movements show the same kinds of effects as those observed with real words (Magnuson *et al.*, 2003). In addition, the recognition of artificial-lexicon words appears to be relatively unaffected by their similarity to specific real words (*i.e.*, there is effectively no competition from words belonging to the lexicon of the participant's native language; Magnuson *et al.*, 2003). The use of an artificial lexicon therefore allowed us to investigate, in a controlled fashion, the involvement of stored prosodic knowledge and signal-based prosodic information during word recognition.

We trained participants to associate non-objects (nonsense shapes) with spoken non-words. The non-objects' names formed minimal pairs that were segmentally identical and differed only in stress placement (*e.g.*, TOlaco *vs*. toLAco). In the training phase, participants heard acoustically reduced versions of the non-words as they learned the object-word associations. Differences in two acoustic stress cues in the original natural utterances – amplitude and duration differences – were neutralized. In the test phase, participants heard stimuli in both reduced- and full-cue versions (*i.e.*, with and without the acoustical manipulations, though note that the full-cue versions were also edited tokens and hence were not the original recordings). In the test phase, participants had to recognize the corresponding objects. If participants use their stored prosodic knowledge about lexical stress – that penultimate stress is the default, and that primarily amplitude cues signal words with antepenultimate stress – then there should be a difference in fixation behavior between the reduced- and full-cue versions only for antepenultimate stress words. If penultimate stress words such as toLAco are recognized by default, the addition of amplitude (and duration) cues should not influence their recognition. But the addition of these cues in the test phase should allow participants to perform better when they hear the full-cue versions of antepenultimate non-objects' names such as TOlaco. Critically, if this benefit for the full-cue versions is found, it must reflect prior abstract, not word-specific knowledge about antepenultimate stress cues (*i.e.*, knowledge that is generalized over the real Italian lexicon and hence is not specific to the newly-learnt words). It cannot reflect memories for specific episodic encounters with the newly-learnt words, since, prior to the test phase, the participants will never have heard these words with amplitude (or duration) differences between their first two vowels.

In summary, in Experiment 2 we investigated whether Italian listeners exploit stored abstract knowledge about lexical stress to optimize the recognition of newly-learnt words. Such a finding would suggest that prosodic knowledge should be considered part of the listener’s abstract phonological knowledge about spoken words, knowledge which, alongside that about individual segments, is used during lexical access (Cho, McQueen & Cox, 2007; Gaskell & Marslen-Wilson, 1997; McQueen *et al*., 2006; Norris & McQueen; 2008; Shatzman & McQueen, 2006).

Method

Participants

Twenty-two students (mean age: 27.9, sd: 5.1) from the University of Trento took part. They received course credit for their participation. They were all Italian native speakers with no known hearing problems and normal or corrected-to-normal vision. None had participated in the previous experiment.

Materials

Twelve trisyllabic non-words were created (*binulo, canvilo, confuro, curfino, desico, goliso, patuco, pencilo, pindumo, tefubo, tolaco*, and *tudero*). Each non-word was recorded twice, once with penultimate stress (e.*g*., toLAco), and once with antepenultimate stress (*e.g.*, TOlaco) by a female Italian speaker in a sound-attenuated room (sampling at 44 kHz, 16 bit resolution, mono). Each non-word was spoken at the end of the sentence “Clicca sul” (“Click on the”). As in Experiment 1, speaking rate was controlled across conditions (penultimate stress: 5.06 syllables per second; antepenultimate stress: 4.92 syllables per second; t = 1.16, p = .25). In this way, we obtained twelve critical pairs. Each critical pair was composed of two segmentally identical non-words that differed only in stress placement (toLAco *vs*. TOlaco). Twenty-four line drawings of nonsense objects were randomly selected from a database of non-objects (Non-existing Objects Database, www-server.mpi.nl/experiment-pictures/production-pictures/; see Figure 3). The nonsense objects were randomly assigned to the non-words.

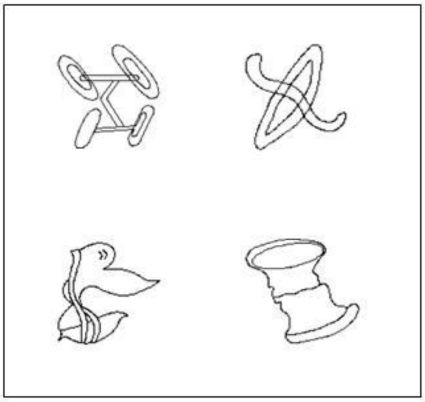

Figure 3. Experiment 2: Examples of non-objects displayed in a 4-alternative trial.

We created a modified version of each non-word. Based on the results of Experiment 1, we neutralized one main stress cue (the amplitude of the first two vowels) and one secondary stress cue (the duration of these vowels). For each non-word, we calculated the average amplitude and the average duration of its first and second vowels. Then, using the PSOLA algorithms in Praat (Boersma & Weenink, 2001), we replaced the original duration values of both vowels in each non-word with the average duration of those two vowels in each non-word. The amplitude of the first two vowels in each nonword was set to the average value of those vowels. In this way, we replaced the original acoustics of the two first two vowels of each non-word. Note that the full-cue versions were also obtained by editing the natural versions of the stimuli. That is, we applied the same adjustment procedures used when creating the reduced-cue versions, but replaced the original values with those same values. This meant that the overall duration and amplitude values in the full-cue materials remained the same as in the original recordings, but also ensured that the stimuli had nonetheless been passed through the same procedures, so that the reduced- and full-cue versions did not differ in their overall quality. In this way, we had two versions of each non-word: The full-cue version and the reduced-cue version, in which the acoustic cues to lexical stress pattern were partially neutralized (see Table 2). Both the full- and reduced-cue versions of all non-words were spliced back into the carrier sentence (“Clicca sul”). The same token of this sentence was used throughout.

Antepenultimate stress					
	First vowel	Second vowel	t (11)	p value	Mean
Duration (ms)	165	82	8.26	<.001	121
Amplitude (Pascal)	.05	.02	5.19	<.001	.03
Pitch (Hz)	196	181	4.36	<.001	-
Spectral Tilt	.3	.03	2.41	<.05	-

Penultimate stress

	First vowel	Second vowel	t (11)	p value	Mean
Duration (ms)	64	165	-10.69	<.001	211
Amplitude (Pascal)	.04	.02	3.36	<.005	.03
Pitch (Hz)	244	186	20.23	<.001	-
Spectral Tilt	.1	.1	<1	n.s.	-

Table 2. Original values of duration, amplitude, pitch and spectral tilt and their t-test comparisons, for the first and second vowel of the novel words with each stress pattern in Experiment 2. *Notes.* The mean values used in creating the manipulated versions of these stimuli are also reported. Spectral tilt is expressed as a unitless ratio.

For each stimulus, a feedback sentence for use in the training phase was also recorded by the same speaker, with the stimuli uttered at the end of the sentence (*e.g.*, "Ora puoi vedere di nuovo il TOlaco", “Now you can see the TOlaco again”). One token of this feedback sentence, without the final non-word, was selected and each reduced-cue non-word was spliced onto the end of it.

Procedure

The experiment was composed of three phases: Two training phases plus a test phase. Because previous research has shown that the lexicalization of newly-acquired words is associated with nocturnal sleep (Dumay & Gaskell, 2007; Davis, Di Betta, Macdonald, & Gaskell, 2009), we decided to run the experiment over two consecutive days. This choice increased the chance that the new words would be learnt well, but note that even words learned over two days with the benefits of overnight consolidation should not be considered to be equivalent to existing words. On the first day, participants completed the first training phase. On the second day, they returned to do the second training phase and the test phase. During the test phase we recorded participants' eye-movements using a head-mounted Eyelink II System, at a sampling rate of 500 Hz and recording both eyes. Participants performed both the training and the test phase sitting approximately 50 cm in front of a computer screen (screen size 360 mm x 270 mm).

The first training phase was composed of 5 blocks. Within each block, each stimulus was presented 4 times, for a total of 96 trials for each block (24 non-words x 4 repetitions each). In Block 1, participants had to choose between 2 non-objects displayed on the screen; they never saw both non-objects that formed a critical pair on the same screen (*e.g.*, we displayed TOlaco and biNUlo, but never TOlaco and toLAco). In Block 2, participants had to choose between 2 objects that did form critical pairs (*e.g.*, we displayed TOlaco and toLAco). In Block 3, participants had to choose among 4 non-objects, and, as in Block 1, no critical pairs were displayed together (*e.g.*, we displayed TOlaco, biNUlo, CANvilo, and deSIco). Block 4 was the same as Block 2 (2 objects forming a critical pair). Finally, in Block 5 participants had to choose among 4 non-objects and, as in Blocks 2 and 4, the displayed stimuli formed critical pairs. This procedure is summarized in Table 3.

Day 1				Day 2			
Blocks	Objects	Stress Pairs	Accuracy	Block	Objects	Stress Pairs	Accuracy
B1	2	No	75%	B6	2	No	98%
B2	2	Yes	65%	B7	2	Yes	86%
B3	4	No	90%	B8	4	Yes	87%
B4	2	Yes	78%	B9	4	Yes	90%
B5	4	Yes	83%				

Table 3. Training block structure in Experiment 2 and percentage accuracy per block. *Notes*. Blocks = block number; Objects = number of objects displayed per screen; Stress Pairs: Yes if stress pairs were shown in the same display; Accuracy = percentage of correct responses.

As is also shown in Table 3, the second training phase was composed of 4 blocks. The procedures for Blocks 6 and 7 corresponded respectively to Blocks 1 and 2 from Day 1. Blocks 8 and 9 corresponded to Block 5. Note that we included two blocks where both members of a critical pair appeared on the same screen because a previous study on lexical learning showed that Italians build stress information into new lexical representations only when they are explicitly encouraged to do so (Sulpizio & McQueen, in press). Including trials with minimal pairs forced participants to attend to stress differences.

We used the same timing procedures in the two training phases. Each trial started with a fixation cross in the center of the screen, displayed for 500 ms. Then two or four non-objects appeared on the screen and remained there until participants clicked the mouse button. At the same time as the visual stimuli appeared, the auditory instruction (carrier sentence plus target word, *e.g.*, "Clicca sul *TOlaco*") was played over headphones. Participants had to click the mouse on the target non-object that corresponded to the non-word they heard at the end of the carrier sentence. At the same time the mouse was clicked, a sentence was played to indicate if the response was correct (*giusto*, 'right') or not (*sbagliato* 'wrong'). Then the target non-object was displayed again, centered on the screen and the feedback sentence (*e.g.*, "Ora puoi vedere di nuovo il *TOlaco*" 'Now you can see the TOlaco again') was played. In the two training phases participants heard the non-words only in their reduced-cue versions.

The test phase followed the second training phase. Before the test, the eye-tracker was mounted and calibrated. The test phase was composed of 2 blocks. Within each block, each trial was repeated two times. For each trial, participants heard a target non-word (*e.g.*, TOlaco) and they had to select the corresponding non-object among 4 possible alternatives displayed on the screen. The four possible choices belonged to two critical pairs (*e.g.*, TOlaco and toLAco; BInulo and biNUlo). In Block 1 participants heard stimuli only in their reduced-cue versions, whereas in Block 2 they heard the non-words only in their full-cue versions. The two blocks were run one after the other, with no break between them. Stimuli were randomized within each block.

In the test phase each trial was structured as follows. First a fixation cross was displayed, centered on the screen, for 500 ms. Then four non-objects appeared on the screen (see Figure 3) and remained there either until participants clicked the mouse button or for a maximum of 5000 ms. A white screen was used during the inter-stimulus interval of 480 ms. The four non-objects were centered in the four quadrants of the screen. The auditory instructions (carrier sentence plus target non-word, *e.g.*, Clicca sul *TOlaco*) were played over headphones (starting when the non-objects appeared). Participants had to click the mouse on the target non-object whose name was heard at the end of the carrier sentence. During the test phase, participants did not receive any feedback. Every eighth trial there was a drift correction to adjust for possible small head movements.

Results

During the training phases, participants successfully learned the non-object names. At the end of the first training phase, object identification accuracy reached 83%, whereas at the end of the second training phase it reached 90% (for details, see Table 3). Training phase data were not analyzed further.

Two analyses were performed on the results from the test phase. First, we ran a 2x2 analysis, comparing performance on the two stress types (penultimate and antepenultimate) for each of the two acoustic versions of each newly-learnt word (reduced-cue and full-cue versions, with data from Blocks 1 and 2 of the test phase respectively). In this way, we tested whether there was evidence that the distributional bias favoring penultimate stress in the Italian lexicon influenced the behavioral data and more specifically whether this bias affected the recognition of the reduced- and full-cue stimuli. Second, correlations between duration and amplitude measures of the full-cue stimuli (*i.e.,* the cues that had been neutralized during training) and the fixation behavior on these stimuli were run in order to establish whether listeners, in this condition, used those cues to identify the words' stress patterns.

Only trials in which participants clicked on the correct non-object were considered in these analyses (10% of all observations were discarded for this reason). We considered fixations on a non-object as being all those that fell within a 6.3 cm square centered on the middle of each non-object. Thus, each fixation was coded as being made to the target non-object, to its competitor, or to one of the distractor non-objects. Fixation proportions were computed in the same way as in Experiment 1, and, also as before, three time windows (first syllable, an absolute window corresponding to the grand average duration of the first 1.5 syllables, and the first two syllables) were defined (again with an offset of 200 ms for programming and making a saccade). Figure 4 shows fixations on target non-objects, their competitors and the two distractors over time for each experimental condition. Figure 5 shows fixation proportions to targets in each condition in each of the time windows. As in Experiment 1, mixed-effect analyses of log transformed data were performed with participants and items as random factors.

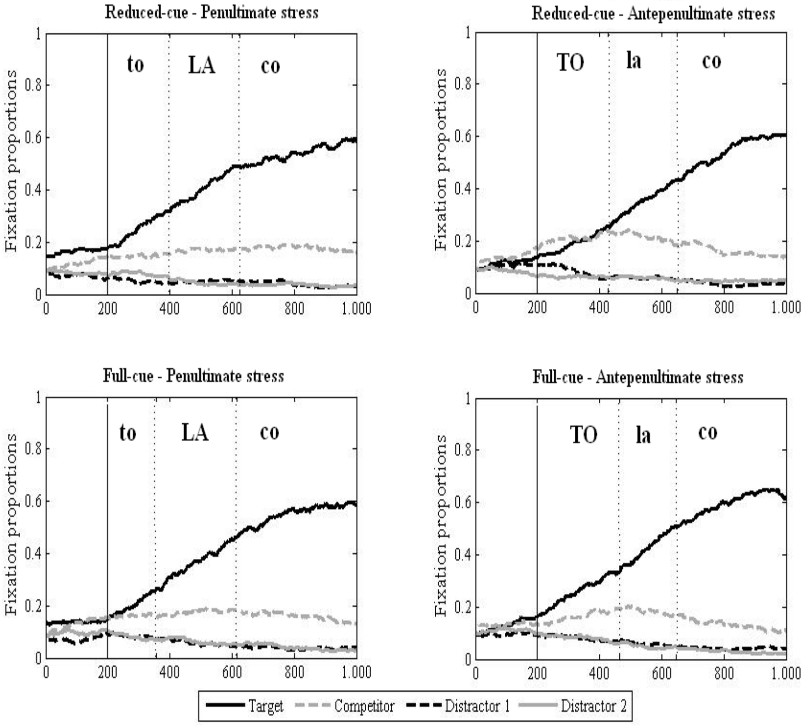

Figure 4. *Experiment 2:* Fixation proportions to targets, competitors, and distractors over time (in ms on the abscissa) in each of the four experimental conditions. Fixations in response to newly-learnt words with penultimate (*e.g.*, toLAco) and antepenultimate stress (*e.g.*, TOlaco) are shown on the left and right respectively. Fixations to reduced-cue tokens (those heard during the learning phase) are given in the upper panels; those to the full-cue tokens are shown in the lower panels. The solid vertical lines show the beginnings of the time windows starting 200 ms after the words' average onsets; the dotted lines indicate the ends of the time windows aligned to the average offsets of the first and the second syllables respectively, each again delayed by 200 ms.

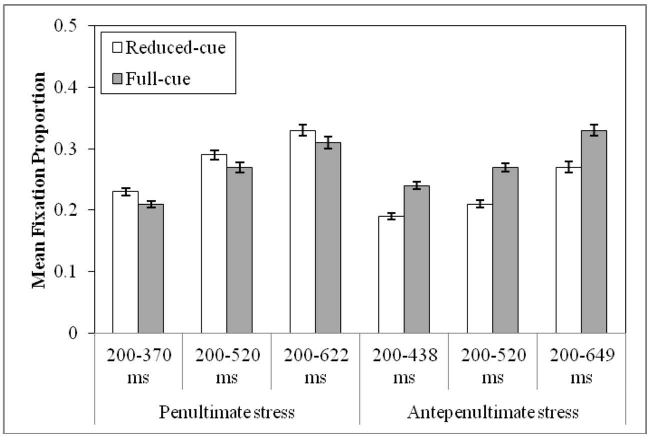

Figure 5. *Experiment 2:* Mean fixation proportions to penultimate- and antepenultimate-stress targets for both the reduced- and the full-cue conditions. The mean values are given (in ms) for all three time windows: the first syllable (200-370 ms or 200-438 ms, respectively, for penultimate and antipenultimate stress words), the first syllable plus half of the second syllable (200-520 ms for both types of word), and the first and second syllables (200-622 ms or 200-649 ms, respectively, for penultimate and antipenultimate stress words). Error bars are standard errors.

Fixation analyses

Target analysis. Fixation proportions on targets was the dependent variable, and stress type (penultimate *vs*. antepenultimate), acoustic version (reduced- or full-cue), and their interaction were fixed factors. The analysis on fixations in response to the first-syllable revealed that participants fixated more penultimate-stress than antepenultimate-stress targets (β = -0.81, t = -3.19, p < .01). An interaction between the two factors revealed that, compared to the reduced-cue version, listeners improved their performance when they heard the full-cue versions of the newly-learnt words, but this happened only for stimuli with antepenultimate stress (β = 0.77, t = 2.16, p <.05). No main effect of acoustic version was found (t = -1.6). This suggests that addition of acoustic cues benefited recognition only of newly-learnt words with antepenultimate stress. But the duration of the first syllables differed across conditions (and hence the amount of data contributing to the different cells of the analysis was not controlled). The same results were found, however, in the other two analyses, where durational differences were controlled (across stress types and acoustic versions for the 1.5-syllable window, and across acoustic versions for the 2-syllable window). For responses to the first 320 ms of the stimuli (the 200-520 ms time window, *i.e.*, the first 1.5 syllables), there was a main effect of stress type (β = -0.97, t = -3.38, p < .01), a significant interaction between stress type and acoustic version (β = 0.87, t = 2.13, p <.05) and no main effect of acoustic version (t = -1.54). For responses to the first two syllables, there was again a main effect of stress type (β = -0.98, t = -3.046, p < .01), a significant interaction between stress type and acoustic version (β = 0.86, t = 1.98, p <.05) and no main effect of acoustic version (t = -1.6).

Target-competitor analysis. Further analyses compared the amount of competition in the four conditions. We used the difference in fixation proportions on target and competitor as dependent variable and stress type and acoustic version as fixed factors. We selected the same three time windows as before. The analysis in the first-syllable time window revealed more competition for antepenultimate-stress than for the penultimate-stress targets (β = -1.20, t = -2.55, p < .01), presumably because, in the former case, the (default) penultimate stress words were the competitors. Moreover, the interaction between the two factors revealed that competition decreased when listeners heard the full-cue versions of the stimuli, but only when the targets had antepenultimate stress (β = 1.16, t = 1.78, p =.08). The other two analyses (with durational differences controlled) showed the same pattern of results, with a main effect of stress type (first 1.5 syllables: β = -1.58, t = -3.03, p < .01; first two syllables: β = -1.63, t = -2.949, p < .01) and a significant interaction between stress type and acoustic version (first 1.5 syllables: β = 1.53, t = 2.06, p <.05; first two syllables: β = 1.52, t = 2.029, p <.05).

Correlation analysis

We performed correlation analyses to test whether, in the full-cue condition, amplitude and duration cues drove the observed improvement in antepenultimate-stress detection. We did not run these correlations for penultimate-stress targets because recognition of penultimate-stress targets in the full-cue condition did not improve. Considering first the initial syllable and then the first two syllables, correlations were performed across words, comparing the difference in fixation proportions between pairs of antepenultimate- and penultimate-stress words to the duration differences between these pairs of words, and then again for the corresponding amplitude differences. Behavioral and acoustic measures of the target words were subtracted from the respective measures of their competitors (those with the opposite stress pattern). When we used the first syllable as the time window, we found a marginal correlation between fixations and duration (*r* = -.50, t (10) = -1.83, p <.1). When we used the first two syllables as the time window, we found a marginal correlation between fixations and amplitude (*r* = -.41, t (10) = -1.43, p <.1). In both correlations, as the difference in duration or amplitude increased, listeners tended to look more at the antepenultimate-stress targets.

Discussion

In Experiment 2 we found that Italian listeners applied two kinds of stored knowledge about lexical stress when recognizing newly-learnt words. First, they appeared to have knowledge about the distributional stress bias in Italian trisyllabic words, and specifically that penultimate stress is the default, because the acoustic manipulation of stress cues did not affect how they recognized non-objects with penultimate stress. Moreover, penultimate-stress targets were recognized earlier than antepenultimate-stress targets: Listeners appear to assume that the penultimate pattern is the default. Second, they appeared to know about the acoustic cues that normally signal words with antepenultimate stress. Unlike in Experiment 1, they used not only amplitude, but also duration when detecting antepenultimate stress (though both effects were statistically weak). This difference across experiments suggests that Italians have knowledge about the variety of acoustic cues that are used to signal stress, and that they can use them to different degrees in different situations. In the normal situation (*i.e.*, with the real words tested in Experiment 1), amplitude information appears to be enough to establish that the current word does not have the default stress pattern. But in the situation where listeners are attempting to recognize new words that differ only in stress (as in

Experiment 2), they may appeal to all available acoustic cues (*i.e.*, amplitude and duration information) to detect the non-default pattern.

Both of these findings reflect the use of abstract (*i.e.*, not word-specific) knowledge about lexical stress. Listeners heard the antepenultimately and penultimately stressed newly-learnt words in equal proportions, so there was nothing in their experience with these specific words that indicated that they should be treated differently. Furthermore, when listeners heard the full-cue versions of the newly-learnt words, this improved their recognition of the antepenultimate-stress targets, even though they had learned those stimuli through hearing acoustically-different (reduced-cue) versions. That is, there was nothing in their prior experience with these new words that indicated they should have particular durational or amplitude properties. There are therefore two different types of knowledge about stress that Italian listeners have abstracted and stored: The phonological patterns related to stress, and the relative frequency of those patterns in the Italian lexicon. Both of these types of knowledge appear to be used during the recognition of newly-learnt words.

General Discussion

We investigated how Italians use lexical stress in spoken word recognition and whether they use abstract knowledge about lexical stress when recognizing spoken words. In Experiment 1, in line with previous results for Dutch (Reinisch *et al.*, 2010), we found that Italian listeners used stress information in word recognition as soon as it became available, and prior to segmental disambiguation. Listeners considered penultimate stress (the most common pattern) to be the default, and picked up on acoustic cues to stress only when recognizing words with antepenultimate stress. In Experiment 2, we found two main results. First, we found further evidence for default assignment of penultimate stress: Novel words with this stress pattern were recognized more quickly than those with antepenultimate stress, and the addition of stress cues to penultimate-stress newly-learnt words did not improve their recognition. Second, we found that Italians used their knowledge about the acoustic cues normally associated with antepenultimate stress to help them recognize new antepenultimate-stress words. This prior knowledge appears to be abstract knowledge about how stress is normally cued in Italian (*i.e.*, not knowledge specific to the newly-learnt words) since, prior to the test phase, these cues had not been associated with the novel words.

These results shed new light on several issues. First, they provide information on the acoustic cues that Italians use as they detect stress position and on how these cues interact with the distributional bias. Our results show that Italian listeners use mainly amplitude to identify a word's stress pattern (Albano Leoni & Maturi, 1998), but only when those words have antepenultimate stress. We also found that duration may be involved in the detection of antepenultimate stress. This happened only in Experiment 2, during the recognition of newly-learnt words (and there only weakly). In such cases, amplitude information may not be sufficient to signal antepenultimate stress: To detect the correct stress pattern of a newly-learnt word accurately, listeners may tend to use all the acoustic information that they find in the signal. These results are partially in contrast with those of Alfano (2006). In her experiments, Alfano manipulated the vowel duration and pitch of words belonging to minimal pairs (*e.g.*, PAgano 'they pay' *vs.*

paGAno 'pagan'), but she did not investigate amplitude. She asked participants to listen to the manipulated words and to identify which syllable bore stress; then, listeners had to judge whether the two words of the minimal pairs had the same pattern or not. She showed that duration was the main acoustic cue that the listeners used to identify stress. A possible explanation for these different findings is that listeners are able to use more or fewer cues depending on the amount of information found in the signal. Thus, when amplitude information is not sufficient, Italians may also use other available cues to detect stress.

The way Italians exploited acoustic information for recognition of newly-learnt antepenultimate-stress words suggests that they have abstract knowledge about the acoustic cues related to stress. This means not only that Italian listeners are able to analyze spoken words into their component phonological parts (segments and suprasegmental attributes), but also that they are able to form abstractions about those components. In particular, they appear to have knowledge that words that have a particular stress pattern tend to have particular acoustic properties. Besides this acoustic-phonetic knowledge, however, Italians also have other knowledge available for use in the recognition of both novel and well-known words. The listeners combined their knowledge about acoustic cues with knowledge about the biased distribution of penultimate and antepenultimate stress in Italian (Thornton *et al.*, 1997). These two sources of stored knowledge appear to interact with each other to optimize stress detection. This could work very efficiently, at least in the situation where it is known that the target word will have three syllables, as in the present experiments. Italian listeners could assign penultimate stress by default (because they know that 80% of trisyllabic words will have this pattern) and then test the validity of this assumption by checking the phonetic information in the speech signal. If the first two syllables of a trisyllabic word contain antepenultimate stress cues (*e.g.*, a marked reduction in amplitude going from the first to the second syllable), then listeners would need to change this default assumption, but otherwise they could maintain the hypothesis that they are hearing a word with penultimate stress. The assumption that penultimate stress is the default pattern might extend beyond trisyllabic words, however. This is because, across word lengths, penultimate stress is the most common pattern in the Italian lexicon (Krämer, 2009). Italians might thus exploit knowledge about this bias in words of all lengths.

This study also addressed a temporal question: When do Italians use stress information in spoken-word recognition? Lexical stress is a source of information that could help to resolve the lexical competition process: Stress information can reduce the number of possible competitors that the listener needs to consider during word recognition. The results of Experiment 1 show that Italians take advantage of stress information to modulate the lexical competition process as soon as that information comes available. In this situation, they exploited lexical stress information (and knowledge about the distributional bias) before segmentally disambiguating material became available (as Dutch listeners also appear to do with respect to signal-based stress cues; Reinisch *et al*., 2010). More generally, however, Italian listeners are likely to use segmental and suprasegmental information at the same time during word recognition. For instance, in distinguishing between CAnapa and caNAle as they unfold over time, Italians appear to use the stress information in the first two syllables to help resolve the temporary ambiguity between these two candidates, but it is likely that they are also using the segmental information in these syllables to rule out other candidates – those that do not begin /kana/.

Three points should be made concerning this perspective on how stress is processed over time in Italian. First, the present findings offer further support for the view the listening to speech is incrementally optimal (Norris & McQueen, 2008; Reinisch *et al*, 2010; Tanenhaus *et al*. 1995; Warren & Marslen-Wilson, 1987), that is, that listeners use all incoming information as soon as it becomes available to form an optimal interpretation of the currently unfolding utterance. Second, the current findings are consistent with the hypothesis that segmental and suprasegmental properties of the speech signal are decoded in parallel (Cho *et al*., 2007; Tagliapietra & McQueen, 2010). Cho *et al*. (2007) proposed a Prosody Analyzer that is responsible for the computation of suprasegmental information. Working in parallel with prelexical mechanisms responsible for the extraction of segmental information, the Prosody Analyzer extracts suprasegmental information (including that which signals lexical stress) from the speech input, and builds a prosodic representation of the current utterance. This representation then constrains the word recognition process, along with segmental representations of the input. Because the segmental and suprasegmental analyzers use the same source of acoustic information, they are interconnected, the results of the two processes are inter-dependent, and the two types of representation are computed at the same time. Italian listeners thus appear to be processing segmental and stress information simultaneously.

Third, however, use of stress cues in lexical access is language specific. For example, it appears that English listeners do not depend heavily on suprasegmental stress information during spoken-word recognition (Cooper, Cutler, & Wales, 2002; Creel *et al.*, 2006). English listeners certainly make use of other prosodic cues in the speech signal (*e.g.*, information about the location of Intonational Phrase boundaries; Cho *et al.*, 2007), but suprasegmental stress information appears to be relatively unimportant in English word recognition because stress in English is signaled segmentally. Unstressed vowels in English are usually reduced, and hence English listeners focus primarily on the segmental distinction between full vowels and reduced vowels (usually schwa; Cooper *et al.*, 2002; Fear, Cutler & Butterfield, 1995). Since segmental information is enough to recognize words efficiently, English listeners do not rely on suprasegmental lexical stress cues (Cooper *et al*, 2002; Cutler, 2005; Fear *et al.*, 1995). In contrast, in a language such as Italian, where suprasegmental stress cues can be temporarily more informative than segmental cues, lexical stress information is used to optimize the recognition of spoken words.

Our findings also shed new light on the ongoing debate about how words are stored in the mental lexicon. We found that Italians have abstract knowledge about lexical stress – about the acoustic cues signaling antepenultimate stress and about the distributional bias favoring penultimate stress – and that this knowledge is used during lexical access. This first set of findings could be explained either in a model in which lexical representations are phonologically abstract (McClelland & Elman, 1986; Gaskell & Marslen-Wilson, 1997; Norris

& McQueen, 2008) or in a model in which lexical representations are episodic traces (Goldinger, 1998; Pierrehumbert, 2002). In both cases, stress knowledge could modulate the word-recognition process (*e.g.*, knowledge about the distributional bias could be derived over time either from abstract or episodic lexical representations, and then used on-line to influence lexical selection). Critically, however, the other findings from Experiment 2 cannot be explained by a purely episodic model of lexical representation. Recognition of antepenultimate-stress targets improved when full-cue versions of the newly-learnt words were presented, even though the participants had never heard those acoustic versions before. A purely episodic model of the lexicon predicts that listeners should recognize a word better if the acoustically detailed input perfectly matches previously stored traces of that word, and hence that the reduced-cue versions of the new words – which had each been heard 72 times during the exposure phase – would be recognized better than previously unheard full-cue versions. Our results show that this was not the case.

A possible response in defense of the episodic position might be that the full-cue advantage for the antepenultimate-stress words arises because listeners recognize the new words by comparing them to episodic traces of known words with the same stress pattern – traces which do contain amplitude and duration cues. But this seems to be an unlikely possibility. First, as Magnuson *et al*. (2003) have shown, recognition of newly-learnt words in the artificial-lexicon paradigm appears to be relatively unaffected by the similarity of those new words to specific well-known words. The fact that the artificial lexicon may be considered self-contained and not affected by competition from words in the listeners' native lexicon does not mean that the way newly-learned words are accessed cannot be affected by knowledge about the segmental and suprasegmental phonology of the native language (cf. Magnuson *et al.*, 2003). But the effects of specific words on the recognition of words in an artificial lexicon appear to be limited. Second, if word recognition was done through a process of comparison to episodic traces, one would still have to predict that the strongest analogies would be between the current input and previous traces of the same word. That is, one would still expect the exposure episodes of the novel words to dominate in the comparison, and hence that there would be an advantage in the recognition of the reduced-cue versions of these words.

Alternatively, one might argue that listeners do not store episodes of words at the lexical level, but instead store episodes of fine-grained, sub-lexical phonetic details. These components, if available at a prelexical level, could then be used in the recognition of both well-known and newly-learned words, and in particular it would be possible for the listener to use generalizations made over prior episodes of known words (*e.g.*, about the acoustic-phonetic properties of words with antepenultimate stress) in the recognition of new words. But this is not a theory about lexical representation.

The current results thus constitute a challenge for the view that the lexicon is composed solely of specific stored episodes of words. It would instead appear to be the case that the lexicon is comprised of phonologically abstract representations. As new words are learned, knowledge about the phonological content of those words – including their segmental make-up and their stress pattern – comes to be stored in the mental lexicon. Specifically, listeners appear to be able to label a novel word as having particular phonemes and a certain stress pattern, and then use stored knowledge about the acoustic properties of other (existing) words which have phonological components that are labeled in the same way. It is this analytic capability (the ability to form abstract representations of the components of spoken words and use those representations in word recognition) that strictly episodic models of the lexicon lack (Cutler *et al.*, 2010; McQueen *et al.*, 2006). In Italian, it appears that new words with penultimate stress (like existing words of this type) are coded as having the dominant stress pattern, and hence that they can be recognized by default. For words with antepenultimate stress, however, they are coded during learning as such, and hence, when additional cues associated with this pattern are present in the input, and even though those cues have never been heard before in those words, those cues can nevertheless be used to facilitate word recognition. On this view, knowledge about the cues signaling antepenultimate stress is abstract too – it needs to be general (*i.e.*, not word-specific) knowledge about antepenultimate stress for it to be applied to other words which share that structure.

It is important to note, however, that we are not advocating a strictly abstractionist model of spoken-word recognition. As we argued in the Introduction, models in which all episodic details are lost cannot explain the evidence that such details are retained in long-term memory (Goldinger, 1998; Nygaard, Sommers, & Pisoni, 1994). Creel, Aslin and Tanenhaus (2008), for instance, have recently shown that listeners may use information about the talker's voice in word recognition. Using the same paradigm as we adopted in Experiment 2, they found that listeners fixated targets more when a target and its competitor had been spoken, during the learning phase, by different talkers than when the same stimuli had been spoken by the same talker. These results suggest that episodic details (such as those concerning talker voice) play an important role in word recognition, and even in the recognition of newly-acquired words.

The best framework to interpret our results is thus a hybrid model with both episodic and abstractionist components (Cutler *et al*. 2010; Goldinger, 2007). In such a model, an episodic memory system would store the idiosyncrasies of specific speech episodes. This system could then interact with both prelexical and lexical abstract representations and could be involved in the consolidation of new traces into abstract forms (Goldinger, 2007). On this view, then, word recognition is based on phonological abstraction, but that process is supported by an episodic memory system. The evidence that memories of episodic detail can influence word recognition (*e.g.*, Creel *et al.*, 2008; Goldinger, 1998; Nygaard *et al.*, 1994) arises in this account not because those details are stored in the mental lexicon, but rather because they are stored elsewhere, in a manner that they can nevertheless influence word recognition.

We suggest that this episodic influence has its effect at the prelexical level. Previous research has suggested that, with respect to segmental information, abstraction is a prelexical process, such that abstract representations of speech sounds mediate between the speech signal and the mental lexicon (Cutler *et al.*, 2010; McQueen *et al.*, 2006; Sjerps & McQueen, 2010). These representations are flexible, allowing listeners to learn about idiosyncratic pronunciations (through exposure to talker-specific details; Eisner & McQueen, 2005). These representations may thus be based on episodic memories, and may be modulated by experience with specific talkers. Importantly, because they are prelexical and abstract, they support generalization of speech learning over the lexicon. Abstract knowledge about the component segments of words can thus benefit word recognition.

The current findings, along with those on the use of prior knowledge about the durational properties of prosodic words (Shatzman & McQueen, 2006), suggest that similar prelexical abstraction processes apply to the suprasegmental properties of the speech signal. Our findings show that listeners have abstract knowledge not only about the form of prosodic words (the relative durations of syllables in monosyllabic versus polysyllabic words; Shatzman & McQueen, 2006), but also about other prosodic properties (lexical stress patterns). Because Italians have acquired abstract knowledge about stress in Italian – the penultimate-stress bias in trisyllabic words, and the acoustic properties associated with antepenultimate stress – they can bring that knowledge to bear when recognizing newly-learnt words. As we suggested earlier, these cues could be extracted from the speech signal by a Prosody Analyzer (Cho *et al.*, 2007), working in parallel with the prelexical segmental abstraction process.

We draw three related conclusions. First, Italian listeners have abstract knowledge about lexical stress. They know that a distributional bias in trisyllabic words exists which favors penultimate stress, and they use that knowledge during the recognition of well-known and newly-learnt spoken words. Moreover, they know that the uncommon pattern (antepenultimate stress) is revealed by specific sources of acoustic-phonetic information in the speech signal, and again they use that knowledge in the recognition of known and new words. Second, it appears that listeners extract and compute prosodic information at the same time as they compute segmental information. These two processes seem to occur in parallel, as the speech signal unfolds over time. This means, as we have shown, that stress information can sometimes be used to disambiguate Italian words before segmental disambiguation is available. Third, prosodic knowledge about lexical structure appears to be phonologically abstract rather than word-specific, suggesting in turn that lexical representations are abstract rather than episodic in nature. Prelexical processing may thus involve abstraction processes not only for segmental material (McQueen *et al*., 2006) but also for suprasegmental material.

APPENDIX

Target stress pairs used in Experiment 1

	Frequency	No. letters		Frequency	No. letters
Abaco	0	5	aBAte	1.1	5
abacus			*abbot*		
Acaro	0	5	aCAcia	0	6
mite			*acacia*		
Acero	1	5	aCEto	1.3	5
maple			*vinegar*		
ALluce	0.7	6	alLUme	0.3	6
big toe			*alum*		
Asino	1	5	aSIlo	1.7	5
donkey			*kindergarten*		
ATtico	0.7	6	atTIguo	0.4	7
penthouse			*adjacent*		
COLlera	1.4	7	colLEga	2.2	7
anger			*colleague*		
CAlamo	0	6	caLAta	1.3	6
quill			*invasion*		
CAnapa	0	6	caNAle	2.1	6
hemp			*channel*		
CANdido	1.3	7	canDIto	0.4	7
candid			*candied*		
CElebre	2.1	7	ceLEste	1.3	7
famous			*pale blue*		
COdice	2.3	6	coDIno	1.1	6
code			*ponytail*		
COmico	1.7	6	coMIzio	1.3	7

Funny			*meeting*		
Eremo *hermitage*	0.9	5	eREde *heir*	1.8	5
Estero *Foreign*	2.3	5	eSTEta *aesthete*	0	5
FEdera *pillow case*	0	6	feDEle *faithful*	2	6
FORbice *Scissor*	0.6	7	forBIto *polished*	0.4	7
FRAgola *strawberry*	0.7	7	fraGOre *uproar*	0.9	7
IMpeto *Impetus*	1.2	7	imPEro *empire*	1.9	7
LATtice *Latex*	0.6	7	latTIna *can*	0.8	7
LOculo *burial niche*	0	6	loCUsta *locust*	0	7
MAcabro *gruesome*	1.1	7	maCAco *macaque*	0	6
MAstice *Putty*	0.6	7	maSTIno *mastif*	0.3	7
MISsile *Missile*	1.3	7	misSIva *missive*	0.9	7
Monito *Warning*	1.2	6	moNIle *jewel*	0	6
Panico *Panic*	1.8	6	paNIno *sandwich*	1.1	6
PROtesi *prothesis*	1.6	7	proTEsta *complaint*	2.2	7
Remora	0.3	6	reMOto	1.4	6

Hesitation			*remote*		
SAlice	0.8	6	saLIva	1.4	6
willow			*spittle*		
SEnape	1.2	6	seNAto	2.3	6
mustard			*senate*		
TOnaca	0.9	6	toNAle	0.9	6
habit			*tonal*		
ZIgomo	0	6	ziGote	0	6
cheekbone			*zigote*		

Note. The stressed syllables are in capital letters. Frequency is log transformed.

Interim

In the first two chapters, we dealt with lexical stress processing in spoken word recognition. As highlighted in Chapter 1, lexical stress is driven by three main acoustic correlates – pitch, duration, and amplitude – and the importance of each correlate varies among languages. For example, in English stress is signaled mainly by pitch (Fry, 1958), while in Dutch it is mainly signaled by duration (Reinisch *et al.*, 2010). In Chapter 2 we showed that Italian listeners use mainly amplitude to detect stress in normal conditions, but they can also use a combination of correlates – amplitude and duration – in case they have to recognize words in more adverse conditions. Listeners use stress information to perform lexical selection (see, *e.g.*, Cutler & Donselaar, 2011; Tagliapietra & Tabossi, 2005), especially in those languages in which unstressed syllables do not contain schwa or any other significant segmental modification. Moreover, when we considered at which processing stage stress information intervenes in word recognition, we found that listeners use stress as soon as it becomes available (see also Reinisch *et al.*, 2010). Finally, we considered stress also in relation with the domain of word prosody. In Chapter 1 we described evidence for Dutch listeners who have stored information about syllable duration and use such prosodic information when recognizing new words (Shatzam & McQueen, 2006). Our investigation went further in this direction. We found that Italian listeners have abstract information not only about the acoustic cues that drive stress, but also about the distribution of the different stress patterns within the lexicon.

The experimental results in Chapter 2 suggested that prosodic knowledge about lexical stress (and about phonemes too) is phonologically abstract: This mean that such prosodic knowledge is generalized in the lexicon and available to label novel words with the same prosodic pattern. Listeners use a pre-lexical level of abstraction to map the acoustic input onto the lexicon: At this level, acoustic information is categorized in abstract suprasegmental (and segmental) units which get in contact with the lexical knowledge (Cutler, 2010; McQueen *et al.*, 2006). Moreover, it has been proposed that the pre-lexical abstraction process may work through two different mechanisms that compute phonemes and word prosody, respectively (Cho *et al.*, 2007).

Our data shed new light on the nature of lexical stress, showing that it may be viewed as abstract knowledge that listeners have about word prosody. These conclusions point to understanding the nature of lexical stress in spoken word recognition. However, we are interested in investigating the representation of lexical stress within different linguistic processes, in order to get an all-accomplished idea of lexical stress and to draw overall conclusions on the nature of lexical stress. To move on in this direction, we studied lexical stress within a different linguistic process and in a task that implies the production of words. By using the reading aloud task, we aimed at assessing how lexical stress is processed in polysyllabic word reading, and whether any similarities exist between the computation of stress in word recognition and in word naming. We addressed the following issues: Do readers compute word prosody apart from phonemes? At which time in processing does stress computation take place? Do readers use the statistical information on the stress patterns' distribution?

Since we assume that the representation of lexical stress in spoken word recognition and reading aloud is quite similar, it is reasonable to assume that: Segmental and suprasegmental information are part of two partially autonomous domains; the segmental and the suprasegmental domains interact at different levels of word processing. These assumptions are not process-specific, but they may be applied to both word recognition and word production. The nature of suprasegmental as well as segmental information might partially overlap in spoken word recognition and reading aloud. If this is the case, it does not mean that the two processes work exactly in the same way, but only that their nature is quite similar.

Further support for our main assumption comes from the literature on the self-monitoring system (see, *e.g.*, Roelofs, 2003; Roelofs, Özdemir, & Levelt, 2007). The monitoring system is a device that allows us to check our speech production. The monitoring system works at two level: An internal monitoring device checks whether the phonological plan has been correctly encoded (for a somewhat different view, see Oppenheim & Dell, 2010); then, external monitoring checks whether the utterance has been correctly articulated. Research on this issue has concluded that the monitoring of an utterance takes place through a procedure that involves two separate but closely linked systems, one related to speech planning and the other one to speech perception. Monitoring would occur through the matching of the perceived word-form into the encoded word-form. This means that the information about the word form which is activated in the speech recognition system is similar in nature and compatible with the information activated in the speech production system. Therefore, the considerations on the monitoring systems are in line with the view that suprasegmental information may be encoded and represented similarly in both the word comprehension and word production systems.

To conclude, in order to advance in understanding the representation of lexical stress, it is useful to investigate its nature in both word perception and production. While in the first two chapters we have dealt with lexical stress in spoken word recognition, in the next two chapters we will investigate stress assignment in reading aloud. We will address the issue of whether readers compute word stress apart from phonemes, at which processing stage readers compute stress, and whether readers use information on the stress patterns' distribution during lexical processing. We will compare how stress is represented in the speech comprehension system and in the speech production system and whether the two representations have a similar nature or not. What we expect is a comparable pattern of results for the two processes in terms of abstractness and autonomy of lexical stress. We do not exclude that specific results may occur depending on the process. However, we assume that the suprasegmental information, similarly to segmental information, is represented in a very similar form in the different linguistic processes.

Lexical stress in reading aloud

Chapter 3

3.1 Introduction

In free stress languages, polysyllabic words can bear stress on different syllables. In such languages stress position may be more or less predictable, depending on which factors contribute to determine stress position. In order to assign stress to a word, readers can use information derived from different sources, such as distributional properties of language, explicit rules, and lexically stored knowledge. Moreover, in dealing with written language, in some cases a graphic mark may signal the word stress position (*e.g.*, *coliBRÌ*, ‘hummingbird’). The role that each factor has in determining stress assignment may vary across languages. In dealing with stress we have thus to take into account the cross-linguistic differences: Stress assignment is at most a language-specific process that needs specific investigations in each polysyllabic language. When considering the previous issues altogether, stress assignment appears as a quite complex process that needs to be studied from different points of view to rightly understand how people read words.

Let us consider the case in which an Italian reader has to read out the polysyllabic word *TAvolo* (table). The reader has to convert the printed string into its corresponding phonetic realization and during this process he/she must place stress, which can occupy one of the three syllables. Here and in the next sections we deal with how the process of stress assignment may take place and which information contributes in determining stress position when participants read a word. Before reporting our investigation on whether stress and suprasegmental information can be computed autonomously from segmental material and how such process may take place, in the present chapter we revise three issues that are fundamental to understand how stress computation works in reading. First, we briefly revise the lexical and sub-lexical computations of stress in reading and which kind of relation exists between segmental and suprasegmental information. Second, we discuss the role of distributional information and, specifically, the role of orthographical information in driving stress assignment. Previous research has shown that distributional information can drive readers in stress assignment to words and pseudowords. Third, we revise how computational models have implemented stress computation in polysyllabic word reading and whether such models are able to explain the human behavior in stress assignment and its interaction with the other reading components.

Following the present introductory chapter, we report two experimental studies investigating the relationship between segmental and suprasegmental information in reading polysyllabic Italian words. In Chapter 4 we investigate whether stress information is computed apart from segmental information and whether it can affect reading performance autonomously from the computation of phonemes. In Chapter 5, we investigate at what level stress computation may affect the word reading and whether this process may interact with the phonetic realization of the words.

To conclude, our investigations aim to shed new light on how Italian people compute stress when reading polysyllabic words. Investigating such issue is important to understand how prosodic knowledge works during reading, and how it is represented in the mind. Before reporting our results, we will review those studies relevant for stress assignment and the models that have tried to implement stress computation in polysyllabic word reading.

3.2 Multiple sources for stress assignment

In polysyllabic languages, the information about word stress is needed to read out a word. Studies on reading aloud have argued that stress information can be retrieved or computed in different ways. In a language like Italian, where stress may occupy one of the last three syllables and its position is often unpredictable, stress information may have a lexical source. Consider the word *CAnapa* (hemp) as an example: To assign the correct stress pattern, a reader has to know the word and retrieve its relative stress pattern, that is the antepenultimate stress. Thus, as Colombo argued (1992; see also Colombo, 1991; Colombo & Tabossi, 1992), lexical knowledge is an important source to assign stress to polysyllabic words: Stress is stored with the other lexical information and readers retrieve the stored phonological representation of the target word to address the correct pronunciation of the stimulus.

In a cross-linguistic perspective, the assumption of a lexical source for word stress may be extended to those languages where stress position is not fixed or strictly governed by rules. When participants read aloud a word with an unpredictable stress pattern, they can retrieve the stress information that is stored in the lexicon. Results in line with this view have been obtained not only in Italian (Colombo, 1992), but also in other languages with no fixed stress position. Studies in English (Rastle & Coltheart, 2000) and in Greek (Protopapas, Gerakaki, & Alexandri, 2006; 2007) have shown that the lexicon constitutes an important source for stress assignment and, at least in Greek, such source appears available since the early development of reading ability (Protopapas *et al.*, 2006).

However, lexical information is not the only source available to the reader for assigning stress to polysyllabic words. Readers may also use the sub-lexical route to drive stress assignment. Initially, studies on different languages have interpreted the sub-lexical source as a mechanism that assigns stress through the computation of rules. Research in Italian (Colombo, 1992) and English (Rastle & Coltheart, 2000) argued that the most common stress pattern in a language could be considered as the regular stress and that the sub-lexical route assigns such pattern by default during word reading. Evidence for a tendency to assign the regular stress was also found in the Italian literature on pseudoword naming, with readers assigning more often penultimate stress to pseudoword stimuli (Colombo, 1992; Colombo *et al.*, 2000; Colombo, Fonti, & Cappa, 2004). Studies in Italian (Miceli & Caramazza, 1993) and Spanish (Gutiérrez Palma & Palma Reyes, 2004) have argued that the sub-lexical procedure is able to assign stress following the word's syllabic structure and its phonological related rules. For example, in Italian if the penultimate syllable of a word ends with consonant (*e.g.*, *conCERto,* ‘concert’), then it has to be stressed (Krämer, 2009), although some exceptions exist. In such a view, readers assign stress by exploiting the stress information coming from both lexical and sub-lexical activation.

More recently, research in reading has started to study whether other types of sub-lexical information are able to drive stress assignment to polysyllabic words. It is assumed that readers can select the word’s stress pattern driven by some sub-parts of the word – in particular the word ending – that work as orthographic cues for stress. In line with a growing interest for how orthography allows readers to establish the word’s stress pattern, the sub-lexical computation of stress has been interpreted more in a connectionist-distributed way: In such a view, stress would be assigned sub-lexically not applying any explicit rule, but following distributional tendencies that allow readers to associate some recurrent orthographic sequences to a certain stress pattern. Considering the growing number of studies that investigate whether and how orthographic patterns affect stress assignment, this topic will be briefly reviewed in the next paragraph, although it will not be investigated in the following chapters.

The role of orthographic information in driving stress selection is not the only issue of interest when we deal with the sub-lexical computation of stress. In a recent investigation, Colombo and Zevin (2009) proposed that, when reading a word sub-lexically, Italian participants may compute stress information apart from segmental information, since the word prosodic patterns are represented separately from lexical and segmental information. The view that stress information may have an autonomous status finds support in the word production literature that assumes that speakers compute segmental information apart from suprasegmental information and the two processes take place in parallel (Levelt, Roelofs, & Meyer, 1999; Roelofs & Meyer, 1998). Such parallel computation of segmental and suprasegmental material may occur also when reading a word aloud. Consequently, we addressed the hypothesis that in reading, participants are able to retrieve and compute stress information apart from segmental material and this may occur in case of both lexical and sublexical computation. This issue will be addressed in the experiments reported in Chapters 4 and 5. The experiments reported in Chapter 5 also suggest that the assignment of word's stress involves the phonological output buffer, the place where lexical and sub-lexical information are merged together.

3.3 The distributional information and its role in stress assignment

When reading a word aloud, people may have access to stress information from different sources. In the last years, a growing number of studies is showing that different types of distributional knowledge may be one of those sources that address stress assignment in polysyllabic word and pseudoword reading. Two types of distributional knowledge with different origins can be considered. The first type of knowledge concerns the distribution of the different stress patterns in a language: Readers know how the stress patterns of their language are distributed in the lexicon. This means that readers implicitly know whether a certain stress pattern is widely represented in their lexicon and thus there is a high probability to read a word with that stress pattern. The second type of knowledge concerns the relationship between orthographic information and stress patterns. Many studies have shown that some orthographic sequences – that is, the initial and mainly the final part of a word – are mostly associated with a certain stress pattern. These sequences can work as cues able to address readers in stress assignment. However, as we will see below, the importance of the distributional knowledge is not universally determined, but it varies among languages. Distributional information may play a role mainly in those languages with a free stress system and its importance may vary according to the strictness of rules that govern stress assignment.

Investigating word reading in English, Rastle and Colthert (2000) assessed whether readers assign stress following a default mechanism. Since 83% of English bisyllables bear initial stress, they assumed that such pattern is assigned by default during bisyllabic word reading. On the basis of this distributional asymmetry, Rastle and Coltheart defined the initial stress as the regular one and the final stress as the irregular one. However, their first experiment did not reveal any difference between reading regular and irregular stress words. Because of this result, the authors re-defined the notion of regularity, in terms of a morpheme-based system. They assumed that stress is assigned according to word morphology. Some morphemes are mostly associated to a certain stress pattern; Thus, through the sub-lexical route, readers apply an algorithmic procedure to identify such morphological units in the orthographic input and determine the associated stress pattern. With this morpheme-based approach, Rastle and Coltheart found that regular stress words were read faster and more accurately than irregular stress words, but only in the case of low frequency words. However, Chateu and Jared (2003) showed that the regularity effect found by Rastle and Coltheart could be interpreted in a different way. They analyzed the Rastle and Coltheart's stimuli and they found that not stress regularity, but the stimulus' orthographic consistency could account for the difference between regular and irregular stress words: Regular stress words were more consistent in their spelling-to-sound relation and thus easier to read than irregular words.

Within an approach not based on rules, but on distributional tendencies, Kelly and colleagues (Kelly, Morris, & Verrekia, 1998) investigated stress assignment in English word reading. They investigated whether a relation exists between stress patterns and word's orthographic sequences. In two reading aloud experiments (Experiments 1 and 3), using bisyllabic words, they showed that a strong distributional relation exists between some word endings and a given stress pattern and this relation may affect participants in lexical decision and reading aloud tasks. Consider the word ending *et* and *ette* as examples. While English words that end with *et* bear mostly initial stress (91% of bisyllables ending with *et* bear initial stress, *e.g., COmet, SONnet*), words ending with *ette* bear mostly final stress (96% of bisyllables ending with *ette* bears final stress *diNETTE*). Kelly *et al.* (1998) showed that such distribution affected the reading performance, with participants being more accurate when reading words with a stress pattern congruent with their ending sequence. Again in English, Arciuli and Cupples (2006) investigated how readers use distributional knowledge about stress in relation to word's grammatical information. In their reading experiments, using bisyllabic nouns and verbs, Arciuli and Cupples found that orthographic information activated by word ending not only cued lexical stress, but also grammatical class of the word. In English while nouns have mostly initial stress, verbs have mostly final stress and the word ending can cue the word's grammatical information as well as the word's stress pattern (see also Arciuli & Cupples, 2003). Some studies also showed that not only word ending, but also word beginning might work as an orthographic cue for stress assignment to English words. Throughout a corpus analysis, Kelly (2004) showed that a correlation exists between the onset complexity of the first syllable and the word stress pattern. The author highlighted that the occurrence of initial stress increased with the increasing number of consonants in the word onset: A complex word onset enhanced the probability of receiving initial stress.

To summarize, previous studies have shown that, when reading a word, English participants may use the orthographic information as a cue to assign stress, with a prominent role of the word ending and a weak contribute of the distributional knowledge about how stress patterns are represented in the language (*i.e.*, initial stress is the most common). Arciuli and colleagues (Arciuli, Monaghan, & Seva, 2010) recently showed that a similar picture may also be found developmentally. While younger children (5-6 years old) assign stress to pseudowords relying on the information driven by both word beginning and word ending, older children (7-8 years old) rely mainly on the word ending. Furthermore, the distributional bias toward initial stress (the most common in English) is shown only by the younger children and it becomes weaker with the development of the reading system.

Differently from English, studies conducted in Greek have shown that, in such language the orthographic information driven by word ending and word beginning plays a less crucial role than other types of information. As shown in different studies involving pseudoword processing (Protopapas *et al.*, 2006; 2007), Greek readers assign stress to polysyllabic stimuli mainly relying on two sources of information: They retrieve stress from the lexicon, or they assign stress in accordance with the graphic mark (the diacritic) that signals the place of stress. When these two mechanisms do not allow readers to select a stress pattern, a default mechanism assigns stress to the penultimate syllable. Protopapas and colleagues (2007) suggest that one possible source for the default mechanism can be found in the lexical stress neighborhood, defined as the number of words sharing their ending and the stress pattern (see below, Burani & Arduino, 2004). Therefore, when we consider Greek, we note that the role of distributional information in stress assignment is quite marginal and subordinate to lexical and graphic information.

In Italian, a seminal study was conducted by Lucia Colombo (1992) who in a series of experiments investigated the multiple sources of stress assignment in polysyllabic word reading. She found that, other than the lexical information, Italian readers may use two types of distributional information to assign stress: The knowledge of the distribution of penultimate and antepenultimate stress patterns in Italian, and the orthographic information driven by word ending. Because the most part of polysyllabic words bear penultimate stress, the tendency to assign it as a default could be expected. Investigating the effect of stress assignment in reading three-syllabic words aloud, Colombo found that the default stress affected the reading process of low-frequency stimuli: Penultimate stress words (*piSTOla,* ‘gun’) were read faster and better than antepenultimate stress words (*TAvolo,* ‘table’). Moreover, Colombo found that also stress neighborhood – defined as the number of words sharing the same ending and the same stress pattern (*e.g., PENtola,TOMbola,BAMbola,* ‘pot, bingo, doll’) – facilitated the computation of the target word, but stress neighborhood affected only antepenultimate stress words and not penultimate stress words.

Subsequent studies challenged the idea of a default mechanism that follows the dominance of penultimate stress in the Italian lexicon. Burani and Arduino (2004) showed that stress neighborhood facilitates both low-frequency words with antepenultimate and penultimate stress in the same way. Furthermore, their experiments did not show any evidence in favor of a default mechanism assigning penultimate stress. Similar results were obtained in another study with skilled and dyslexic children. Paizi and colleagues (Paizi, Zoccolotti, & Burani, 2011) found that a consistent stress neighborhood facilitates children’s reading performance in the case of both penultimate and antepenultimate stress words. Paizi *et al.* (2011) found only a weak evidence for a default mechanism, which would be exploited only by dyslexic children. In summary, for Italian readers, the word ending can be considered as the main distributional information able to address stress to polysyllabic words. Although readers may be sensitive to the distributional asymmetry between penultimate and antepenultimate stress, such knowledge would weakly affect word reading.

To conclude, the review of the reading studies conducted in different languages has highlighted that readers can exploit different types of distributional information when assigning stress to polysyllabic words. However, the information driven by orthography, with a prominent role of word ending, seems to be the most crucial source of stress assignment. Differently, the overall distribution of stress patterns appears less useful to establish word stress, with readers being only slightly affected by the asymmetrical distribution that the stress patterns have in the languages. At this regard, some of the results that will be reported in the present studies (see Chapters 4 and 5) support the view that the distributional difference between penultimate and antepenultimate stress does not produce any visible effect in word reading, especially when participants read words through the lexical route.

3.4 How do computational models account for word stress?

The investigation of reading has initially focused on understanding how people read monosyllabic words. Dealing with monosyllabic units the first computational models of reading (see, for example, Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Seidenberg & McClelland, 1989) did not address the issue of how stress is assigned when a word has to be read out. The first effort to incorporate stress processing within a computational model was done by Rastle and Coltheart (2000). They assumed that word stress is lexically stored. However, Rastle and Coltheart also assumed that, in order to deal with stress assignment to pseudowords, readers must be able to compute stress also through a sub-lexical mechanism.

Thus, they implemented a sub-lexical algorithm that applies stress according to the morphological structure of the stimulus. Rastle and Coltheart (2000) observed that some morphemes are associated to a certain stress pattern: for example, the suffix –eer receives stress, being associated to final stress. Their sub-lexical algorithm was able to discover the presence of such morphemes within a stimulus and to assign stress according to them. However, this pioneer study on modeling stress assignment raises some problematic issues: First, the proposed algorithm did not consider any kind of orthographic pattern that has no morphological status; second, the authors did not explain how the sub-lexical mechanism can merge the segmental and the suprasegmental information together.

Recently, some authors have implemented connectionist models of reading able to assign stress to polysyllabic words. For English, Arciuli and colleagues (Arciuli *et al*, 2010) have proposed a connectionist model able to assign stress from orthography (see also Ševa, Monaghan, & Arciuli, 2009). The model is a simple network that learns to map the distributional information driven by orthography into stress positions. However, while showing that the model is able to assign stress correctly, the authors do not implement any phonological representation of stress, neither they implement how polysyllabic words are phonetically executed. The lack of the implementation of phonological and phonetic components makes the model partially unable to account for some stress phenomena. Another connectionist model of polysyllabic word reading has been recently proposed by Pagliuca and Monaghan (2010). The aim of this model is to simulate word naming in a transparent orthography as Italian. Also in this case, the model is able to assign the word's stress on the basis of orthographic information, and it can read quite accurately three-syllabic stimuli. However, the model failed to simulate behavioral effects related to stress computation in reading Italian, as for example the stress neighborhood effect (Burani & Arduino, 2004). The authors argued that a possible reason for this failure may be found in the way stress is represented: In the model, stress was encoded as a segmental feature – that is, connected to the phonemes (for example, there are two different phonemes for unstressed /a/ and stressed /a/) – and not as a suprasegmental feature able to affect the word's syllabic realization.

Another account for stress computation in reading has been recently offered by a connectionist dual process model, the CDP++ (Perry, Ziegler, & Zorzi, 2010). Differently from the previous reading models, the general architecture of the CDP++ assumes the existence of a specific component for stress computation that, according to the authors, would be located in the phonological output buffer. Thus, the stress output nodes would receive activation of the word's stress pattern from both the lexical and the sub-lexical routes. The stress system would work in parallel to the phoneme output nodes that again receive activation of phonemes from both routes. Finally, the authors assume the existence of a stress parameter, the stress naming criterion: Words are read out only when the word's stress, reaching the activation threshold, has been established. However, until now, there is no empirical evidence in reading studies that justifies the assumption of an autonomous level for stress computation in reading. In contrast, the idea of an autonomous level for stress computation is widely accepted in the speech production literature. As noted above, Roelofs and Meyer

(1998) showed that metrical information is computed apart from segmental information, and the WEAVER++ model of word production (Levelt *et al.*, 1999; Roelofs, 2000) assumes two different components to compute stress and phonemes. Accordingly, it may be assumed that a similar process is at work also in word reading, as argued by Perry and colleagues (2010) in their CDP++ model. At this regard, the studies that will be presented in Chapters 4 and 5 investigate the hypothesis that metrical information is computed apart from segmental information. The results obtained in a series of naming experiments are in line with a theory that assumes the existence of separate components for the computation of metrical and segmental materials: Word stress may be computed apart from the word's phonemes and these two kinds of information may be assembled together within the phonological buffer.

To conclude, in a language with no fixed stress position as Italian, understanding how people compute stress when reading a word aloud is a fundamental issue that needs to be further investigated to fully understand how reading aloud works. When the most part of the lexicon is composed of polysyllabic words and stress may appear in different positions, then the reading process cannot disregard how readers assign stress. As argued in this chapter, studies conducted in different languages mainly focused on two aspects. First, they identified the possible sources of stress, suggesting that stress can be assigned both through the lexical and the sub-lexical routes. Second, previous studies focused on understanding how the sub-lexical route is able to assign stress. At this regard, most of the authors agree that readers are driven by some orthographic patterns – as the word ending

– that work as cues for stress assignment. However, some important issues are still open and need to be investigated. In the following chapters we address three related questions that will be investigated in Italian. First, do readers retrieve and compute lexical stress apart from segmental material? Second, does the distributional difference between the two main Italian stress patterns – 80% of words with penultimate stress vs. 18% of words with antepenultimate stress – affect the way in which the lexical route assigns stress in word reading? Third, does computation of stress affect the latest stages of word reading, *i.e.*, stages where the phonological word is assembled and converted into a phonetic realization? These three issues will be addressed in Chapters 4 and 5. Using a priming paradigm, we ran a series of reading aloud experiments to explore how readers compute word stress and whether such computation takes place apart from the computation of segmental material.

Priming lexical stress in reading Italian Aloud

Chapter 4

Sulpizio, S., Job, R., & Burani, C. (in press). Priming lexical stress in reading Italian aloud. *Language & Cognitive Processes.*

Two experiments using a lexical priming paradigm investigated how stress information is processed in reading Italian words. In both experiments, prime and target words either shared the stress pattern or they had different stress patterns. We expected that lexical activation of the prime would favor the assignment of congruent stress to the target. Results showed that participants were faster in naming target words that had the same stress pattern as the prime. Similar effects were found on target words that were included in lists in which all prime and target stimuli had the same stress type (Experiment 1) and in lists with mixed stress type and congruency between primes and targets (Experiment 2). Results indicate that, in single word reading, metrical information about stress position is activated in the lexicon, independent of segmental information.

Chapter 4

Introduction

In reading Italian aloud, the assignment of stress to three- and more syllable words is the only process that cannot be accomplished by applying rules, but rather requires accessing lexical entries. In a transparent orthography like Italian, simple rules are sufficient to obtain the correct print-to-sound mapping of all the words at the segmental level (Burani, Barca, & Ellis, 2006); but in contrast with this high regularity, there are no rules dictating that a three-syllabic word like “matita” [pencil] bears stress on the penultimate syllable (maTIta), whereas the word “bibita” [drink] bears stress on the antepenultimate syllable (BIbita)[4]. A reader of Italian must learn the correct stress for these words by rote.

It is thus conceivable that, in the lexicon of Italian readers, the phonological representation of a polysyllabic word includes the representation of its metrical structure. However, considering the asymmetrical distribution of the two main Italian stress patterns – about 80% of three-syllables bear stress on penultimate syllable, and 18% bear stress on the antepenultimate syllable[5] (Thornton, Iacobini, & Burani, 1997) – it may also be assumed that only antepenultimate stress is included in the lexical representation, whereas the penultimate stress is the default pattern, consistent with the statistical properties of stress distribution (Colombo, 1992).

Whether the word’s metrical structure may be represented independently of the representation of its phonemic segments – as an

---

[4] Penultimate stress is assigned by rule only in the case that a word has a heavy penultimate syllable (e.g., *bisonte*).

[5] The remaining 2% of three-syllabic words bear stress on the final syllable, and in this case stress is graphically marked (e.g., *colibrì*).

autonomous level of representation – is still an open issue. For word production, it has been proposed that the metrical structure is computed separately from segmental information, and can be autonomously involved in preparing an utterance. Roelofs and Meyer (1998), for instance, found that the production of a Dutch response word was facilitated when participants knew in advance both the number of syllables and the stress location of the word.

Evidence of the latter type is lacking for reading aloud. If the stress of an Italian three-syllabic word is represented in the lexicon as a part of its metrical structure, autonomously from its segmental representation, then in reading it should be possible to prime the production of the stress pattern of a target word by accessing a prime word that has the same stress pattern as the target. Two main predictions can be conceived.

The first prediction follows from positing that the penultimate stress is applied sub-lexically by default (Colombo, 1992), whereas only the antepenultimate stress is lexically represented. Within this view, a low-frequency target word, prone to be read by means of sublexical print-to-sound conversion (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), should be read faster and more accurately when it bears penultimate stress than when it bears antepenultimate stress. Accordingly, two results are expected: a main effect of stress type, with penultimate stress words read faster than antepenultimate stress targets (Colombo, 1992), and a larger stress priming facilitation on antepenultimate stress targets than on penultimate stress targets.

Within the contrasting view which posits that stress is autonomously represented for both antepenultimate and penultimate stress words, and no default mechanism is at work, no difference in processing penultimate and antepenultimate stress words is expected, and stress priming should occur for both word targets with penultimate and antepenultimate stress pattern.

In reading aloud, the locus of the stress priming effect may also be at a non-lexical level, if readers rely on some sort of rhythmic pattern. Colombo and Zevin (2009) investigated stress and metrical computation in reading aloud. By using a “pathway priming” methodology (Zevin & Balota, 2000), in which a list of five primes preceded a target and all stimuli, both primes and targets, were read aloud, Colombo and Zevin (2009) tested stress computation within a lexical (word primes) or a sub-lexical (nonword primes) context. They found that stress can be represented separately from lexical and segmental information: A stress representation could be primed in reading aloud, but only when a sub-lexical mechanism was involved and there was a homogeneous stress context (primes and targets sharing the stress pattern). On the basis of these results the authors concluded that stress priming can be induced as a consequence of sub-lexical rhythmic processing.

In the present study, we adopted a priming paradigm in which a prime word is presented briefly before a target word, and only the target is read aloud. In contrast to the “pathway priming” paradigm, our paradigm allows us to investigate lexical priming in the absence of an overt prosodic/rhythmic context induced by reading primes aloud. To further ascertain that stress priming requires lexical retrieval and does not result from rhythmic priming, we manipulated the list context in which the prime-target pairs were presented. In experiment 1, prime-target pairs with only congruent (or incongruent) stress patterns were presented in the same list; in experiment 2, the two stress types and the two congruency conditions (prime and target with the same or different stress type) were mixed. This mixed list condition aimed at ruling out the possibility that stress could be assigned sub-lexically as a predictable prosodic pattern. If stress priming requires lexical retrieval and is not a consequence of sub-lexical rhythmic processing as argued by Colombo and Zevin (2009), then we can expect the same pattern of results in the two experiments, that is stress priming should occur for both word targets with penultimate and antepenultimate stress, with no difference for the two stress patterns, irrespective of list context.

Experiment 1

In experiment 1, we investigated lexical phonological priming (Ferrand & Grainger, 1993), with primes and targets sharing the stress pattern (congruent condition), or having different stress patterns (incongruent condition). If access to the prime word activates its metrical representation in the phonological lexicon as autonomous information, then targets in the congruent condition (in which the prime stress matches the target stress) should be named faster than targets in the incongruent condition, in line with Roelofs and Meyer's findings (1998). If both stress patterns are lexically represented, the congruent stress prime condition should facilitate the production of both antepenultimate and penultimate stress words.

Method

Participants

Thirty-two students of the University of Trento, all native Italian speakers.

Materials and Design

Two sets of 32 low-frequency three-syllabic words, selected from the CoLFIS database (Bertinetto *et al.*, 2005), were used as targets. One set included penultimate stress words and the other antepenultimate stress words. Stimuli were matched on familiarity, length in letters, orthographic neighborhood size, orthographic neighbors' summed frequency, bigram frequency, orthographic complexity, number of embedded words, embedded words' summed frequency, and two initial phonemes (Table1). Both targets with penultimate and antepenultimate stress had a stress neighborhood composed mainly of stress friends, *i.e.*, their orthographic ending was shared by a majority of words with either penultimate or antepenultimate stress, respectively (Burani & Arduino, 2004; Colombo, 1992). Accordingly, there was no bias toward assigning the penultimate stress on the basis of orthographic/phonological cues of the word ending (Arciuli, Monaghan, & Ševa, 2010; Ševa, Monaghan, & Arciuli, 2009).

	Stress type	
Item variables	Penultimate	Antepenultimate
Word frequency	1.39 (0.82)	1.46 (0.98)
Length in letters	7.13 (0.61)	6.84 (0.63)
Bigram frequency	11.18 (0.37)	11.16 (0.46)

N of orthographic neighbors	0.75 (0.98)	0.81 (0.9)
Neighbors' frequency	6.71 (11.29)	4.96 (8.98)
Familiarity	5.37 (1.1)	5.32 (1.13)
Contextual rules	0.43 (0.61)	0.65 (0.70)
N of embedded words	0.18 (0.39)	0.15 (0.36)
Embedded word frequency	0.34 (1.12)	0.65 (2.02)

Table 1. Summary statistics: Means (and standard deviations) for the three-syllabic target words in Experiments 1 and 2. *Note*: Word frequency measures are calculated out of 1 million occurrences (Bertinetto *et al.*, 2005); bigram frequency is log transformed on the basis of the natural logarithm; number of contextual rules is a measure of orthographic complexity (see Burani *et al.* 2006); familiarity was measured on a 1-7 rating scale (1 = low familiarity, 7 = high familiarity.

Two sets of 32 medium-high frequency three-syllabic words were used as primes. One set included penultimate stress words and the other antepenultimate stress words, all selected from CoLFIS (Bertinetto *et al.*, 2005). They were matched on the same variables as targets (Table 2)[6]. The two sets of 32 primes were paired with two sets of 32 target words, with no semantic relation between prime and target. Targets were divided between the two stress conditions (congruent and incongruent), matching initial phonemes and word length within each subgroup.

Item variables	Stress type	
	Penultimate	Antepenultimate
Word frequency	18.56 (8.26)	19.06 (6.02)
Length in letters	7.06 (0.66)	6.81 (0.59)
Bigram frequency	11.38 (0.33)	11.23 (0.53)
N of orthographic neighbors	1.37 (1.73)	1.31 (1.33)
Neighbors’ frequency	13.53 (25.39)	13.25 (18.13)
Contextual rules	0.65 (0.70)	0.59 (0.61)
N of embedded words	0.12 (0.33)	0.18 (0.39)
Embedded word frequency	3.15 (16.62)	3.09 (10.43)

---

[6] Because of their medium-high frequency, the familiarity ratings were not collected for the prime stimuli.

Table 2. Summary statistics: Means (and standard deviations) for the three-syllabic prime words used in Experiments 1 and 2. *Note*: Word frequency measures are calculated out of 1 million occurrences (Bertinetto *et al.*, 2005); bigram frequency is log transformed on the basis of the natural logarithm; number of contextual rules is a measure of orthographic complexity (see Burani *et al.*, 2006).

The experiment had a 2 (congruent-incongruent stress pattern) x 2 (penultimate-antepenultimate stress) design, with both factors within participants. Four pure blocks were created: each block included stimuli from only one condition (penultimate-stress prime & penultimate-stress target; antepenultimate-stress prime & penultimate-stress target; antepenultimate-stress prime & antepenultimate-stress target; penultimate-stress prime & antepenultimate-stress target). To avoid facilitating effects due to sharing initial phoneme (Malouf & Kinoshita, 2007), primes and targets differed on initial phoneme.

Apparatus and procedure

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

Each trial started with a fixation cross, centered on the screen, for 400 ms. The prime was then presented for 86 ms (Ferrand & Grainger, 1993) in lower-case letters in the center of the screen, followed by the target word displayed in the same position as the prime, in upper-case letters. The target remained on the screen until the participant began to read it aloud or for a maximum of 1500 ms. The interstimulus interval was 1500 ms. A practice preceded the experiment. Naming times were recorded by means of E-Prime software.

Each participant received 64 trials, presented in four blocks. Primes and targets were paired in such a way that for half of the participants a target was preceded by a penultimate stress word, and for the other half the same target was preceded by an antepenultimate one.

The order of prime-target pairs was randomized within blocks and block order was counterbalanced between participants. The experimenter noted the naming errors.

Results

Responses shorter than 250 ms or longer than 1500 ms (1.6% of all data points) were excluded from the analyses.

Results are reported in Table 3. A 2x2 analysis of variance was conducted on RTs as the dependent variable, with condition (congruent-incongruent) and stress type (penultimate-antepenultimate) as within-participant factors (in the analysis by items, the factors were between participants).

There was a main effect of condition (F_1 (1,31) = 9.54, MSE = 2595, $p < .005$; F_2 (1,124) = 12.89, MSE = 2223, $p < .001$), with congruent target words read faster than incongruent targets. There was no effect of stress type (F_1 (1,31) = 3.27, MSE = 795; $F_2 < 1$) and no interaction between the two factors (F_1 (1,31) = 1.38, MSE = 1091; F_2 (1,124) = 1.58, MSE = 2223).

Naming errors, including both phonemic and stress errors, were also submitted to a 2x2 ANOVA with error percentages as dependent variable and condition and stress type as within-participant (or, in the analysis by items, as between participants) factors. No factor reached significance (all Fs < 1).

	Prime-Target stress congruency			
	Congruent		Incongruent	
Target Stress	Mean RT	*%E*	Mean RT	*%E*
Penultimate	579	4.4	598	3.7
	(80)	(5)	(70)	(5.6)
Antepenultimate	577	4.1	614	4.1
	(88)	(5.4)	(72)	(4.5)

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

Discussion

Word targets preceded by stress-congruent primes were named faster than targets preceded by stress-incongruent primes. No main effect of stress type was found, and prime congruency similarly affected the reading of penultimate and antepenultimate words. The priming effect found in the congruent stress condition can be associated to the pre-activation of stress information during processing of the prime. When the stress pattern activated by the prime matches the target stress pattern, then facilitation in reading the target aloud is obtained.

The next experiment investigated the presence of this effect in a mixed list condition, to rule out the possibility that stress was assigned as a predictable prosodic pattern.

Experiment 2

In experiment 2, we mixed the two stress types and the two congruency conditions (primes and targets with the same or different stress type), in order to check whether the homogeneous stress of the targets in each list adopted in experiment 1 affected metrical processing, and assignment of stress specifically, in target words. As Colombo and Zevin (2009) showed, there may be a tendency to homogenize the stress pattern assigned to a word with the stress pattern of its list context; thus, the effects obtained in experiment 1 might be inflated because of this context effect. Since target's stress position in experiment 1 was predictable, readers might have assigned stress in an automatic rhythmic way, homogenizing the stress pattern to the metrical information activated on earlier trials. In this sense, the stress congruency effect could be strategic, depending on context and not on the task (Rastle, Kinoshita, Lupker, & Coltheart, 2003).

In order to test whether the stress congruency effect obtained in experiment 1 was strategic in nature, experiment 2 manipulated list context so that stress could not be assigned sub-lexically as a predictable prosodic pattern. If stress priming depends on lexical retrieval, then we expect the same pattern of results obtained in the first experiment.

Method

Participants

Thirty-two students of the University of Trento, all native Italian speakers.

Materials and Design

The same materials as in experiment 1 were adopted. Four mixed blocks were created: each block was composed of 16 stimuli, four from each experimental condition (penultimate-stress prime & penultimate-stress target; antepenultimate-stress prime & penultimate-stress target;

antepenultimate-stress prime & antepenultimate-stress target; penultimate-stress prime & antepenultimate-stress target).

Apparatus and procedure

Procedure was the same as in experiment 1. Each participant read the 64 target stimuli in four different mixed blocks. Thirty-two target stimuli were assigned to each condition (congruent and incongruent stress pattern), counterbalanced across two lists.

Results

Responses shorter than 250 ms or longer than 1500 ms (1.1% of all data points) were excluded from the analyses. Results are reported in Table 4. A 2x2 analysis of variance was conducted on RTs as the dependent variable, with condition (congruent-incongruent) and stress type (penultimate-antepenultimate) as within-participant factors (in the analysis by items, the factors were between participants).

	Prime-Target stress congruency			
	Congruent		Incongruent	
Target Stress	Mean RT	*%E*	Mean RT	*%E*
Penultimate	598	3.5	609	2.14
	(97)	(3.1)	(106)	(3)
Antepenultimate	600	2.05	614	2.8
	(102)	(2.2)	(106)	(2.9)

Table 4. Mean latencies for correct responses and percentage of errors by condition (with standard deviations), in Experiment 2.

The effect of condition was marginally significant by participants (F_1 (1,31) = 3.77, *MSE* = 1267, *p* = .06), and significant by items (F_2 (1,124) = 4.98, *MSE* = 1214,56, $p < .05$), showing that words in the congruent condition were read faster than words in the incongruent condition. There was no effect of stress type (*Fs* < 1), and no stress type x congruency interaction (*Fs* < 1).

Naming errors, including both phonemic and stress errors, were submitted to a 2x2 ANOVA with error percentages as dependent variable and condition and stress type as within-participant (or, in the analysis by items, as between participants) factors. No factor reached significance (condition and stress type: both *Fs* < 1; interaction: F_1 (1,31) = 3.17, *MSE* = 22.153; F_2 < 1).

Joint analysis for Experiments 1 & 2

To compare results from the two experiments, an analysis of variance was conducted with condition (congruent-incongruent), stress type (penultimate-antepenultimate) and experiment/context (blocked-mixed) as factors. Condition and stress type were within-participant measures in the analysis by participants, and between-participants measures in the analysis by items. Experiment/context was a between-participants factor. There was a significant effect of stress congruency again (F_1 (1,62) =13.28, *MSE* = 1931, $p < .01$; F_2 (1,248) = 17.20, *MSE* = 1740, $p < .01$). RTs in experiment 2 were slower than in experiment 1, resulting in a main effect of experiment in the analysis by items (F_1 < 1; F_2 (1, 248) = 5.97, *MSE* = 1740.46, $p < .05$); but, importantly, experiment/context did not interact with any other factor (stress type-experiment/context interaction: *Fs* < 1; stress congruency-experiment/context interaction: F_1 (1,62) = 1.372, *MSE* = 863; F_2 < 1).

Error percentages were submitted to analysis, with condition and stress type as within-participants measures in the analysis by participants, and between-participants measures in the analysis by items. Experiment/context was a between participants factor. There was a main effect of experiment, in the analysis by participants (F_1 (1,62) = 6.68, *MSE* = 1273,193, $p<.05$; F_2(1,248) = 2.89, *MSE* = 44,327), with more errors in experiment1. Experiment/context did not interact with any factor (all *Fs* < 1).

Discussion

The results of experiment 2 are consistent with those of experiment 1, and show that the stress congruency effect was present even when participants could not rely on rhythmic strategies to assign stress. The effect was not modulated by list context, as shown by the absence of any interaction between stress congruency and list composition. This pattern of results rules out the use of task-specific strategies.

General Discussion

In two experiments, stress information coming from prime activation affected the processing of a target word. Readers were faster to read a word when it was preceded by another word with the same stress pattern than when it was preceded by a word with a different stress.

The stress congruency effect was present on both antepenultimate and penultimate stress word targets. The absence of a main difference in latencies to penultimate and antepenultimate stress words confirms that no default mechanism is at work in stress assignment for words with a neighborhood composed mostly of stress friends (Burani & Arduino, 2004). The similarity in stress priming effects for penultimate and antepenultimate stress targets suggests that both stress patterns are represented in the phonological lexicon and can be activated as a consequence of prime processing.

This stress priming effect may have a lexical source, because it was found in a list where target stress was unpredictable and thus it was not possible to apply any rhythmic cue. The presence of the stress priming effect under such conditions indicates that, in lexical access, the retrieval of stress information is partially autonomous with respect to the phonemic segmental material: When processing a prime, readers retrieve the metrical structure of the word, containing stress position, which then exerts an influence on target word reading.

Our results contrast with those reported by Schiller, Fikkert and Levelt (2004) for Dutch picture naming. In that study, participants named pictures corresponding to bisyllabic words stressed on the first or second syllable. Target pictures were preceded by the auditory presentation of another bisyllabic word with same or different stress. Unlike the present study, Schiller *et al*. (2004) did not find a stress priming effect in Dutch. However, it may be observed that all the Dutch word targets had a predictable stress. Stress was predictable both for words with initial stress, which is by far the dominant stress (or default pattern) in Dutch, as well as for words with final stress (all had a “super-heavy final syllable” to which “metrically regular stress” is applied). Thus, stress could be assigned through a non-lexical mechanism (see Miceli & Caramazza, 1993). The absence of lexical stress retrieval might be the main source for the absence of stress priming in the Dutch study. In contrast, the Italian words used in our study had a stress not predictable on the basis of metrical characteristics, with subsequent retrieval of stress from the lexicon and lexical priming.

The view that metrical information is stored in the lexicon apart from segmental information has been developed with reference to speech production (Levelt, Roelofs, & Meyer, 1999; Roelofs & Meyer, 1998). However, Roelofs (2004) argued that speech production and reading aloud may share the last stages of processing, *i.e.*, phonological and phonetic encoding of the word. According to Roelofs (2004), a model of speech production as the WEAVER++ and the DRC model of reading (Coltheart *et al*., 2001) could be merged at the level of segmental spellout, which precedes the prosodification process.

In a dual route framework of reading (Coltheart *et al*., 2001; Perry, Ziegler, & Zorzi, 2010), the activation of a polysyllabic word in the phonological lexicon may entail its prosodification, which involves syllabification of the word and stress retrieval. The pre-activation of metrical information – stored separately from the segmental material – in the lexicon caused by a prime word would affect some component of the phonological output buffer that keeps trace of stress information during processing. In the CDP++ model (Perry *et al*. 2010) the planning of a target's articulation would be affected by the pre-activation of a congruent metrical structure in the Stress Output Nodes contained in the phonological output buffer. There, the information concerning prime stress may affect the reading of a target word at the level of its phonological encoding, which is also considered the locus of lexical stress encoding in naming (Schiller, 2006). Thus, the prime metrical structure can be exploited during prosodification of the target.

Single route connectionist models of stress assignment (Arciuli *et al*., 2010; Ševa *et al*., 2009) may also be able to account for the present set of results by positing that the pattern of activation characterizing the prime stress may affect the stress unit processing the target. Assuming that stress is part of an output representation, stress priming might affect the resting level of this output representation (Colombo & Zevin, 2009). However, the existing models are still underspecified regarding this issue, so they do not allow us to make more specific predictions at this stage.

In conclusion, metrical information can play an autonomous role in priming the assignment of the correct phonology to a word. Further investigations are needed to understand how metrical information may interact with other orthographic and/or phonological cues that speakers rely on when reading words aloud.

Appendix

Congruent stress pairs – Penultimate stress target: licenza-BADESSA, cuscino-BUDELLO, profeta-CAMBUSA, medaglia-COMPASSO, stivale-CONCIME, cotone-FRAGORE, cancello-GAVETTA, sirena-IBISCO, bagaglio-INGHIPPO, polmone-INTRUSO, sicario-LASAGNA, tempesta-LOMBRICO, commedia-MANGUSTA, patata-MEDUSA, prigione-MIRTILLO, candela-NASELLO, allievo-NIRVANA, ribelle-PERNICE, cugino-POLENTA, nipote-POMATA, vigilia-RAMARRO, miscela-RUGIADA, coltello-STARNUTO, maiale-SUSINA, indizio-TARALLO, cravatta-TIMBALLO, rancore-TOPAZIO, tabella-VANGELO, furgone-VESSILLO, metallo-VILUPPO, padella-ZAVORRA, stupore-ZITELLA.

Congruent stress pairs – Antepenultimate stress target: panico-BALSAMO, tessera-BUFALA, modulo-CALAMO, vicolo-COLICA, sintomo-COTTIMO, preside-FREGOLA, raffica-GANGHERO, comico-IBRIDO, cupola-INDACO, povero-INDOLE, complice-LASTRICO, canone-LOCULO, crimine-MAMMOLA, parroco-MESCITA, stomaco-MICROBO, protesi-NACCHERA, replica-NINNOLO, bambola-PERTICA, incubo-POLIPO, cellula-PORPORA, margine-RANTOLO, stimolo-RUGGINE, arbitro-STIPITE, calibro-SUGHERO, fulmine-TARTARO, coniuge-TIMPANO, sintesi-TOSSICO, tattica-VANDALO, maschera-VERTEBRA, missile-VIRGOLA, liquido-ZAZZERA, nuvola-ZIGOMO,

Incongruent stress pairs – Penultimate stress: tessera-BADESSA, panico-BUDELLO, modulo-CAMBUSA, vicolo-COMPASSO, sintomo-CONCIME, preside-FRAGORE, raffica-GAVETTA, comico-IBISCO, cupola-INGHIPPO, calibro-INTRUSO, complice-

LASAGNA, canone-LOMBRICO, crimine-MANGUSTA, parroco-MEDUSA, stomaco-MIRTILLO, replica-NASELLO, protesi-NIRVANA, bambola-PERNICE, cellula-POLENTA, incubo-POMATA, margine-RAMARRO, stimolo-RUGIADA, arbitro-STARNUTO, povero-SUSINA, fulmine-TARALLO, coniuge-TIMBALLO, sintesi-TOPAZIO, tattica-VANGELO, maschera-VESSILLO, missile-VILUPPO, liquido-ZAVORRA, nuvola-ZITELLA.

Incongruent stress pairs – Antepenultimate stress: cuscino-BALSAMO, cugino-BUFALA, patata-CALAMO, medaglia-COLICA, stivale-COTTIMO, cotone-FREGOLE, cancello-GANGHERO, sirena-IBRIDO, bagaglio-INDACO, polmone-INDOLE, sicario-LASTRICO, tempesta-LOCULO, commedia-MAMMOLA, profeta-MESCITA, prigione-MICROBO, allievo-NACCHERA, candela-NINNOLO, ribelle-PERTICA, nipote-POLIPO, licenza-PORPORA, vigilia-RANTOLO, miscela-RUGGINE, coltello-STIPITE, maiale-SUGHERO, indizio-TARTARO, furgone-TIMPANO, rancore-TOSSICO, tabella-VANDALO, cravatta-VERTEBRA, metallo-VIRGOLA, padella-ZAZZERA, stupore-ZIGOMO.

Syllable Frequency and Stress Priming Interact in reading Aloud

---

Chapter 5

Sulpizio, S. & Job, R. (in preparation). Syllable Frequency and Stress Priming Interact in Reading Aloud.

In current models of reading aloud the structure and operations of the phonological buffer are quite underspecified. We investigated this issue by asking participants to read aloud Italian three-syllabic words in a priming condition: target words varying for the frequency of their initial syllable were preceded by words congruent or incongruent for the stress pattern. The results showed an interaction between syllable frequency and stress prime, with longer reading times for target words with an initial low-frequency syllable in the incongruent stress condition. This pattern does not support a strictly sequential or threshold processing of metrical and segmental information and suggests that both stress assignment and syllable computation affect reading at the level of the phonological-to-phonetic interface.

Chapter 5

Introduction

Reading aloud requires the execution of multiple operations, *e.g.*, perceiving the stimulus, converting the printed information in a speech signal, and articulating the word's sounds, taking into account both segmental (*e.g.*, sounds) and suprasegmental (*e.g.*, stress) information. While many studies have investigated the operations involved in word recognition, the phonological encoding of a word and its phonetic realization have received less attention. The same happens with computational models of reading aloud: They usually implement in a detailed way the procedures readers use to recognize words, but they are less specific about those phenomena related to the production stages (see, *e.g.*, Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), and the very few that have attempted to implement procedures for stress assignment differ in the solutions they propose (see, *e.g.*, Perry, Ziegler, & Zorzi, 2010; Rastle & Coltheart, 2000; Sibley, Kello, & Seidenberg, 2010).

The speech production literature may be helpful in investigating this aspect of reading aloud, as it has been argued that speech production and reading aloud may share the last stages of processing, that is the phonological and phonetic encoding of the word (Roelofs, 2004). In the model developed by Levelt and colleagues (Levelt, Roelofs, & Meyer, 1999) it is assumed that during phonological encoding speakers retrieve in parallel the segmental material and the metrical structure – number of syllables and word's stress pattern – and combine them into the phonological word (see also Roelofs & Meyer, 1998). At this point, the phonological word is phonetically encoded and it is then translated into its phonetic realization. In their connectionist Dual Process model of reading Perry at al. (2010) implement an analogous double process. The model presents stress-output nodes, i.e. nodes specifying the position of the stress within the lexical string, for both the sublexical network and the lexical route. Such nodes are activated autonomously from the segmental information, although full processing of the latter is conditional upon the former: Articulation of the word phonemes cannot be initiated until the word stress has been fully determined.

Two reading studies, both run in Italian (Colombo & Zevin, 2009; Sulpizio, Burani, & Job, in press), support the view that metrical and segmental information are autonomously involved in planning and assembling an utterance, both when stress is sub-lexically computed (Colombo & Zevin, 2009) or lexically retrieved (Sulpizio *et al.*, in press). In particular, the latter studies showed an effect of stress position priming for segmentally different prime-target pairs. Specifically, readers are faster in reading a word when it is preceded by a word with the same stress, *e.g.*, TESsera (card) – BUfala (hoax), than when in is preceded by a word with a different stress, *e.g,.* cuGIno (cousin) – BUfala[7](hoax). The patter was interpreted as showing that stress priming affects the stage of phonological word encoding in the phonological buffer.

An effect that has also been ascribed to the later stages of reading aloud is that of syllable frequency. Researches in different languages have shown that participants are faster in producing a word that starts with a high-frequency syllable than one with a low-frequency

---

[7] Capital letters indicate stressed syllable.

syllable (see, among others, for Dutch: Cholin, Levelt, & Schiller, 2006; English: Cholin, Dell, & Levelt, 2011; French: Laganaro & Alario, 2006; German: Aichert & Ziegler, 2004; Italian: Sulpizio & Job, 2010; Spanish: Carreiras & Perea, 2004) and there is consensus on the claim that such effect is be attributed to the phonetic encoding, when readers convert the abstract phonological word into abstract motor programs.

Jointly considering the effects of stress assignment and of syllable frequency in reading aloud may allow us to better articulate the operations involved in the phonological-to-phonetics interface, the rather neglected and oversimplified component of reading models. Both stress priming and syllable frequency are assumed to affect the latest stages of reading process, when readers (a) spell out segmental and metrical information and (b) articulate the word, with syllable frequency affecting the word's phonetic encoding (Carreiras, Mechelli, & Price, 2006; Laganaro & Alario, 2006).

Thus, an additive pattern of syllable frequency and stress priming would be consistent with models that postulate two separate consecutive stages for the two effects, (*e.g.*, Levelt *et al.*, 1999), or with models that postulate a threshold of activation for one component (*e.g*,. stress assignment) before the other may start its computations (*e.g.*, word articulation) (Perry *et al.*, 2010). Differently, according to Sternberg (1969) logic, an interaction between syllable frequency and stress priming would suggest a common locus for syllable computation and stress assignment.

Experiment

Three-syllabic Italian words were used as stimuli as stress position for these words is not always predictable. Indeed, Italian three-syllabic words have two[8] main stress patterns (Thornton, Iacobini, & Burani, 1997): Antepenultimate stress (*i.e.*, the first syllable bears stress, *e.g.*, TAvolo ‘table’), and penultimate stress (*i.e.*, the second syllable bears stress, *e.g.*, coLOre ‘color’). Although their distribution differs – 80% of three-syllable words bear penultimate stress and 18% bear antepenultimate stress – the two patterns are lexically stored within the phonological lexicon and the asymmetry does not affect lexical reading (Paizi, Zoccolotti, & Burani, 2011).

By jointly manipulating stress priming and syllable frequency we aimed at investigating the operations involved in the phonological-to-phonetics interface that take place during the later stages of word reading. Specifically, if stress priming and syllable frequency originate at different stages of processing then the stress priming effect should be of similar size for both words starting with a high- and words staring with a low-frequency syllable. Differently, if stress priming and syllable frequency affect word reading at the same level, an interaction between the two effects should be expected.

Method

Participants

Twenty-four students (14 male, mean age: 24, sd: 3.8) of the University of Trento. They were all Italian native speakers and they had normal or corrected-to-normal vision. They received credit course for their participation.

Materials and Design

---

[8] The remaining 2% of three-syllabic words bears stress on the final syllable, and in this case stress it is graphically marked (e.g., *colibrì*).

Four sets of three-syllabic words were used as targets. The sets were selected by combining two variables: Frequency of the first syllable (high or low) and stress pattern (penultimate or antepenultimate). Each set was composed of 22 low-frequency words selected from the CoLFIS database (Bertinetto *et al.*, 2005). Stimuli were matched on length in letters, orthographic neighborhood size, orthographic neighbors' summed frequency, frequency of the second and third syllable, mean bigram frequency, orthographic complexity, initial phoneme (Table 1), and had a stress neighborhood composed mainly of stress friends (Burani & Arduino, 2004). Targets were pre-tested to ensure that none of the initial syllables were a probabilistic orthographic cue for stress (Arciuli, Monaghan, & Ševa, 2010). Thus, syllable frequency was not expected to interact with word's stress pattern. To further rule out such possibility, we ran a pilot experiment asking 18 university students to read aloud all targets. Stimuli appeared in capital letters in the center of the screen, after a fixation cross displayed for 400 ms. Each stimulus remained on the screen until the participant began to read or for a maximum of 1500 ms. The presentation order was randomized between participants. Mean RTs for correct responses were submitted to a 2 (high- *vs*. low-frequency syllable) x 2 (penultimate *vs*. antepenultimate stress) ANOVA. The analysis showed an effect of syllable frequency (F_1 (1,17) = 22.196, MSE = 1246, p < .01; F_2 (1,84) = 17.29, MSE = 2033, p < .01), with faster reaction time for words with a high-frequency syllable. Neither stress type (F_1 (1,17) = 1.6, MSE = 246; F_2 < 1) nor the interaction were significant (F_1 (1,17) = 3.6, MSE = 217; F_2 < 1). No effect was significant in the analysis of errors (4.8%).

	First Syllable Frequency			
	High		Low	
	Penultimate Stress	Antepenultimate Stress	Penultimate Stress	Antepenultimate Stress
First Syllable Frequency	690.78 (561.47)	720.36 (505.64)	27.87 (25.61)	41.51 (30.44)
Second+Third Syllable Frequency	1588.93 (847.88)	1711.78 (809.56)	2088.09 (919.7)	2228 (769.35)
Word frequency	4.55 (4.91)	6.54 (11.24)	7.1 (12)	6.05 (7.37)
Length in letters	7 (0.69)	6.87 (0.46)	7.14 (0.35)	7.09 (0.29)
Mean Bigram frequency	11.6 (0.28)	11.57 (0.25)	11.47 (0.27)	11.51 (0.42)
N of orthographic neighbors	1.05 (1.21)	1.09 (1.10)	1.14 (0.94)	1.05 (1.01)
Neighbors' summed frequency	4.74 (9.92)	8.09 (22.9)	2.18 (2.83)	6.04 (14.96)
Contextual rules	0.68 (0.83)	0.82 (0.73)	0.55 (0.8)	0.73 (0.76)

Table 1. Summary statistics: Means (and standard deviations) for the three-syllabic target words. *Note*: Syllable frequency measures are calculated out of 1,000,000 occurrences (Stella & Job, 2001); word

frequency measures are calculated out of 1 million occurrences (Bertinetto *et al.*, 2005); mean bigram frequency is log transformed on the basis of the natural logarithm; number of contextual rules is a measure of orthographic complexity (see Burani *et al*., 2006).

Two sets of 44 high frequency three-syllabic words were used as primes. One set included penultimate stress words and the other antepenultimate stress words, all selected from CoLFIS (Bertinetto *et al.*, 2005). The two sets were matched on: Length in letters, orthographic neighborhood size, orthographic neighbors' summed frequency, mean bigram frequency, orthographic complexity (Burani, Barca, & Ellis, 2006), and initial phoneme (Table 2). Primes were paired with target words in such a way that neither semantic relation nor orthographic overlapping existed between prime and target. Targets were divided between the two prime stress conditions (congruent and incongruent).

	Stress Type	
	Penultimate	Antepenultimate
Word frequency	216 (118)	228 (127)
Length in letters	6.95 (0.78)	6.73 (0.76)
Mean Bigram frequency	11.59 (0.4)	11.47 (0.39)
N of orthographic neighbours	1.95 (1.73)	1.84 (1.44)
Neighbors' summed frequency	51.52 (68.77)	52.65 (65.07)

Contextual rules	0.41 (0.58)	0.77 (0.77)

Table 2. Summary statistics: Means (and standard deviations) for the three-syllabic prime words. *Note*: Syllable Word frequency measures are calculated out of 1 million occurrences (Bertinetto *et al.*, 2005); mean bigram frequency is log transformed on the basis of the natural logarithm; number of contextual rules is a measure of orthographic complexity (see Burani *et* al., , 2006).

The Experiment had a 2 (congruent *vs*. incongruent stress pattern) x 2 (high- *vs*. low-syllable frequency) design. Prime-target pairs were divided in 4 blocks. In each block, prime and target shared the stress pattern & target had a high-frequency initial syllable; prime and target shared the stress pattern & target had a low-frequency initial syllable; prime and target with different stress patterns & target had high-frequency initial syllable; prime and target with different stress patterns & target had a low-frequency initial syllable. Furthermore, in each block, half of the targets had penultimate stress and half had antepenultimate stress, and in no case prime and target shared the initial phoneme. Primes and targets were paired in such a way that for half of the participants a target was in a congruent stress condition (prime and target having same stress), and for the other half the same target was presented in the incongruent stress position (prime and target having different stress). The order of prime-target pairs was randomized within blocks and block order was counterbalanced among participants.

Apparatus and procedure

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

Each trial started with a fixation cross, centered on the screen, for 400 ms. The prime was then presented in lower-case letters just above the center of the screen for 86 ms and it was followed by a 86 ms blank; then, the target stimulus was displayed in upper-case letters just below the center of the screen. The target remained on the screen until the participant began to read it or for a maximum of 1500 ms. The inter-stimulus interval was 1500 ms. A practice session with 8 trials preceded the experiment. Naming times were recorded by means of E-Prime software. The experimenter noted the naming errors.

Results

Responses shorter than 250 ms or longer than 1500 ms (2.4% of all data points) were excluded from the analyses. Naming errors, including both phonemic and stress errors, summed to 2.7% of all data points and were not analyzed. Results are reported in Figure 1.

A 2x2 analysis of variance with syllable frequency (high- vs. low-frequency syllable) and condition (congruent vs. incongruent stress) was conducted on the reaction times (RTs) of correct responses. The factors were within participants in the analysis by participants and between participants in the analysis by items. Analyses treating both partcipants (F_1) and items (F_2) as random variables are reported. The main effect of condition was significant, with congruent target words read faster than incongruent target words (F_1 (1,23) = 10.49, MSE = 3771, $p < .01$, $\eta^2 = .27$; F_2 (1,176) = 51.49, MSE = 1558, $p < .001$, $\eta^2 = .23$). The main effect of syllable frequency was also significant, showing that targets with an initial high-frequency syllable were read faster than targets with a low-frequency syllable (F_1 (1,23) = 8.73, *MSE* = 995, $p < .01$, $\eta^2 = .31$; F_2 (1,176) = 10.24, *MSE* = 1558, $p < .01$, $\eta^2 = .15$). Finally, there was a significant interaction between congruency condition and syllable frequency, (F_1 (1,23) = 4.39, *MSE* = 675, $p < .05$, $\eta^2 = .16$; F_2 (1,176) = 4.26, *MSE* = 1558, $p < .05$, $\eta^2 = .12$): LSD post-hoc comparisons showed that the 55 ms stress priming effect ($p < .005$, $\eta^2 = .31$) for targets with a low-frequency initial syllable was significantly different from the 31 ms effect ($p < .05$, $\eta^2 = .23$) for the targets with a high-frequency initial syllable. Interestingly, the size of syllable frequency effect (8 ms) in the congruent prime condition is close to the effect found in other languages in a no prime condition (see, *e.g.*, Cholin *et al.*, 2011), while in the incongruent prime condition the size of syllable frequency effect is large (32 ms), suggesting that stress priming effect may be inhibitory, affecting target reading when preceded by an incongruent stress prime.

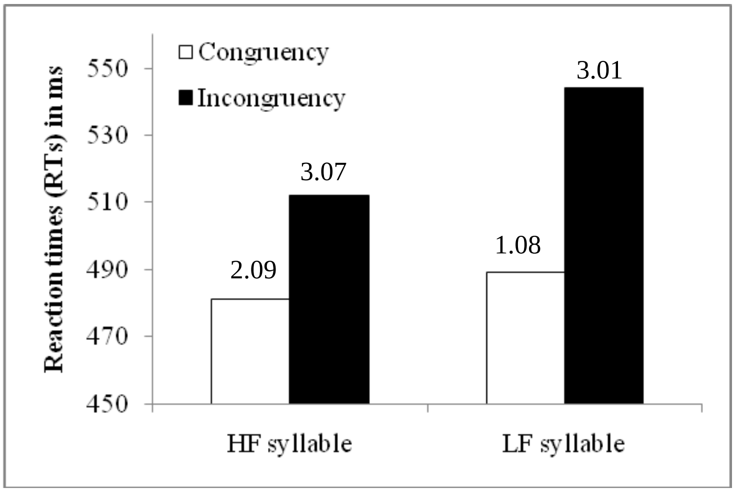

Figure 1. Reaction times and percentage of errors by condition

The results of the present experiment are clear. Word targets preceded by stress-congruent primes were read faster than targets preceded by stress-incongruent primes. Moreover, words with a high-frequency first syllable were read faster than words with a low-frequency first syllable. Finally, the priming effect was larger for targets with a low-frequency first syllable.

General Discussion

The main finding of our study is that syllable frequency and stress priming interact. This result constrains the functional architecture of the reading system not only by suggesting a common locus for the two effects but also by indicating the relative timing of the operations underlying stress retrieval and word articulation in reading aloud.

The effect of syllable frequency has been generally ascribed to the phonetic encoding level, assuming that readers are facilitated in articulating those syllables they produce frequently. Specifically, Levelt *et al.* (1999) argue that high-frequency syllables can be retrieved from a mental syllabary, while low-frequency syllables are assembled using the phonological word as input. The effect of stress priming has been ascribed to mechanisms operating at the level of the phonological buffer (Sulpizio *et al.*, in press) for assigning stress during the word phonological encoding.

The interaction suggests that the common locus for both syllable frequency and stress assignment is the phonological output buffer, where the phonological word is realized. One might argue – contra Levelt *at al.* (1999) – that syllable frequency may affect reading during the orthography-to-phonology conversion. If that were the case, the syllable frequency effect would have emerged only in the congruent stress condition, while in the incongruent stress condition the time needed to solve the stress mismatch would have delayed the assembling of segmental and metrical information, with the consequence of allowing enough time for fully computing low-frequency syllables. In this case, the syllable frequency effect would be greatly reduced or even annulled. This is not the case, and our results support Levelt *et al.*'s (1999) proposal that syllable frequency effect arises at the phonetic encoding.

While both Levelt *et al.*'s (1999) and Perry *et al.*'s (2010) models, the two most explicit models on this issue, would agree that both effects originate at quite late stages of word production, the functional architectures they propose are not fully compatible with our results, though for different reasons. In Levelt *et al.*'s model (1999) it is assumed that the phonological word encoding and the phonetic encoding are sequentially ordered and, in addition, there is no feed-back between the two levels. In Perry *et al.*'s (2010) model the timing of the operations in the phonological output buffer is such that only after the relevant stress pattern has been activated phonemes, structured in their syllabic constituents, can be overtly articulated. In both cases, additive effects are expected between stress assignment and the subsequent syllabification. Although our data support the view that stress assignment is essential for articulation to take place, they also suggest an interactive process at the level of phonology-to-phonetic interface.

The difference in speed of processing between high- and low-frequency first syllables seems to be the critical factor in the pattern we obtained and the interaction would suggest that at the level of phonology-to-phonetic interface words with a high-frequency initial syllable are less prone to interference from the stress mismatch. One way to account for this pattern would be to postulate that the speed with which high-frequency syllable are processed allow for information from the following syllable to be rapidly available, and to be less interfered with by incongruent stress information. For low-frequency syllable, information about the following syllable will be available later in time, and the incongruent stress information may be particularly damaging this situation. Alternatively, it may be assumed that the possible role of speed asymmetry between high- and low-frequency syllables might be linked to the different procedures for syllabification of high- and low-frequency syllables. In the framework of Levelt *et al.*'s (1999) model the former are retrieved from the repertoire of syllables while for the latter a composition from their constituent phonemes is postulated. It may be further assumed that for high-frequency syllables both the stressed and the unstressed forms are stored in the repertoire. Upon reading, both forms are available and as soon as stress information becomes available the correct form is selected. Low-frequency syllables, instead, would be computed from their constituent phonemes, and in the incongruent stress condition the initial computed form must be discarded and the correct form re-computed.

To conclude, our findings show that words with an initial low-frequency syllable are more strongly affected by manipulation of incongruent stress priming than words with a high-frequency initial syllable. This is consistent with the view that the phonological buffer acts as the locus of phonological-to-phonetics interface, where the abstract phonological word is converted into its phonetic representation, and where stress and syllable information may interact.

Appendix

TARGET STIMULI

High frequency syllable words: binario, bisturi, canapa, candela, candido, canguro, concavo, congiura, demone, diploma, folata, fosforo, lacrima, lapide, lasagna, livore, muflone, muscolo, pergola, pernice, persiana, persico, pertica, pertugio, precario, presbite, procione, prologo, rachide, rasoio, sovrano, sughero, tegame, tenaglia, tenebra, trafila, tragico, trapano, veleno, veranda, vescica, vescovo, vipera, viscido.

Low frequency fyllable words: berbero, bottone, cappero, cappone, circolo, cirrosi, cucchiaio, cucciolo, delfino, donnola, fulgido, fulgore, lattice, lattuga, longevo, luppolo, missile, missiva, pallido, palpebra, palpito, pattume, plenario, pollice, polmone, pompelmo, pulcino, pulpito, roncola, ruggito, succube, sultano, taccola, tappeto, tessera, tombino, tombola, tossina, vassoio, velluto, vincolo, virgola, vortice, vulcano.

PRIME STIMULI

Penultimate stress words: affare, appello, azienda, campagna, catena, cavallo, coltello, commedia, confine, contatto, convegno, cultura, destino, destino, esame, figura, finanza, finestra, fortuna, incendio, lezione, locale, malato, maniera, medaglia, memoria, missione, modello, nipote, palazzo, pianeta, pistola, poltrona, profumo, regime, salute, segreto, senato, settore, tesoro, timore, tragedia, vestito, vettura.

Antepenultimate stress words: albero, angolo, attimo, camera, carcere, carica, clinica, codice, compito, coniuge, credito, cronaca, debito, diavolo, epoca, fabbrica, fascino, favola, formula, incubo, lettera, limite, macchina, margine, maschera, medico, metodo, missile, nobile, pagina, plastica, polvere, popolo, principe, reddito, scandalo, secolo, simbolo, spirito, tavolo, termine, traffico, vertice, vittima.

Summary & Conclusions

Chapter 6

The aim of this thesis was to study the representation of lexical stress in Italian. In polysyllabic languages with free-stress position, the investigation of lexical stress is a powerful tool to determine how people understand and produce words: If I am not able to recognize stress position, then I am not able to recognize the word correctly. To illustrate, let us consider the minimal pair ANcora ‘anchor’ *vs*. anCOra ‘again’, in which the two words differ only in the stress pattern. Similarly, if I am not able to assign stress, then I am not able to produce a word correctly.

Previous research on lexical stress investigated the topic considering mainly one linguistic domain at a time: Some studies focused on the role of stress in word comprehension, while other studies focused on word production, and still other on reading aloud. The novel approach taken in this thesis is to investigate across tasks how lexical stress is represented and whether its underlying processes differ during word production and word comprehension. Thus, focusing on the representation of lexical stress in Italian, We addressed two related questions: First, does lexical stress have an abstract and autonomous representation with respect to the segmental material? Second, is the postulated abstract representation involved in both language comprehension and the language production and reading systems?

The experiments reported in Chapter 2 investigated how Italian listeners use lexical stress in spoken-word recognition and whether they use abstract knowledge about lexical stress when recognizing spoken words. In two experiments, using the printed-word eye-tracking paradigm (Huettig & McQueen, 2007; McQueen & Viebahn, 2007) and the artificial-lexicon eye-tracking paradigm (Creel *et al.*, 2006; Magnuson *et al.*, 2003; Shatzman & McQueen, 2006) respectively, we tested how Italians use lexical stress when recognizing known and newly-learnt words. The results of these experiments showed three related findings.

First, Italian listeners have abstract knowledge about lexical stress. They are sensitive to the distributional bias for penultimate stress in three-syllabic words and are able to exploit such bias when recognizing known and newly-learnt words. Moreover, listeners are aware that the less common pattern (antepenultimate stress) is revealed by specific acoustic information in the speech signal, and again they use that knowledge in the recognition of known and newly-learnt words.

Second, listeners extract and compute prosodic information at the same time as they compute segmental information. These two processes seem to occur in parallel, as the speech signal unfolds over time. This means that stress information can sometimes be used to disambiguate Italian words before segmental disambiguation is available, optimizing the word-recognition process.

Third, prosodic knowledge about lexical structure is phonologically abstract rather than word-specific, suggesting that lexical representations are abstract – that is, generalized overt the lexicon and available to label words as having a certain prosodic pattern – and not episodic in nature. Thus, listeners use a pre-lexical level of abstraction to map the sounds into the lexicon: At this pre-lexical level, acoustic information is categorized in abstract segmental (*e.g.*, phonemes) and suprasegmental (*e.g.*, lexical stress) units, which are used to make contact with the lexical knowledge (Cutler, 2010; McQueen *et al.*, 2006).

Chapters 4 and 5 investigated the representation of lexical stress in reading polysyllabic words aloud. In Chapter 4, we ran two experiments using a lexical priming paradigm. Participants had to read aloud a target word (*e.g.*, camBUsa, 'storeroom') that was preceded by a prime word either sharing (*e.g.*, proFEta 'prophet') or not sharing the stress pattern (*e.g.*, MOdulo 'form'). The results showed three main findings.

First, lexical stress can be primed, suggesting that the metrical information of a word is stored in the lexicon and is autonomous from segmental information. This finding is in line with the view developed in the speech production literature: In planning the production of a word, speakers retrieve and compute the metrical structure separately from the phonemes, and they can autonomously involve the metrical structure when preparing an utterance (Levelt *et al.*, 1999; Roelofs & Meyer, 1998). Second, stress priming affected both antepenultimate and penultimate stress word targets, showing that lexical reading is not modulated by the distributional asymmetry between penultimate stress (the most common pattern) and antepenultimate stress. This result suggests that both stress patterns are equally represented in the phonological lexicon and can be activated as a consequence of prime processing.

Third, the fact that metrical and segmental information are stored separately in the lexicon suggests that the computation of lexical stress and the computation of phonemes are performed by different underlying mechanisms. It further suggests the need of postulating a mechanism for integrating the two kinds of information during production.

In Chapter 5 we report one experiment run to further test (a) the activation of lexical stress in reading a word aloud and (b) at what processing level stress priming may affect word reading. We again used a lexical priming paradigm with primes and targets sharing or not their stress pattern. In addition, we manipulated the frequency of the initial syllable of the target words. The results of the experiment showed an interesting interaction between stress priming and syllable frequency: The effect of stress priming was larger for words with a low-frequency initial syllable than for words with a high-frequency initial syllable. This finding suggests two conclusions.

First, the effects of stress priming and syllable frequency occur at a common locus, that might be identified in the phonological output buffer. At this stage of processing, the interaction between the metrical and the segmental information might take place at the level of phonology-to-phonetic interface, *i.e.* when the assembled phonological unit has to be converted into its phonetic realization.

Second, the interaction we found allows us to constraints the functional architecture of any psychologically viable model of word production as it suggests that, at least in reading, the interface between the phonological word and its phonetic realization works interactively. The analysis of the two most explicit models on this issue, namely the Levelt *et al.*’s model (1999) of word production and the Perry *et al.*’s model (2010) of visual word recognition and reading aloud show that, although for different reasons, both these models predict additive effects of stress assignment and the subsequent syllabification.

The studies reported here allow us to draw conclusions on two different issues: First, a methodological issue, namely the use of different paradigms when investigating a given topic; second, and more important, the theoretical issue of the nature of lexical stress. Let us consider the methodological issue and its implications first. In our experiments, we employed different methods and techniques to investigate lexical stress: Eye-tracking to test the role of lexical stress in spoken-word recognition (Allopena, Magnuson, & Tanenhaus, 1998; McQueen & Viebahn, 2007) and lexical-phonological priming to test the role of lexical stress in reading aloud (Ferrand & Grainger, 1993). The use of different methods rests on the assumptions that (a) some tasks and techniques are best suited for some research questions than others, and (b) converging evidence from different techniques may rule out method specific effects. Eye-tracking allows investigating the time course of spoken-word recognition. As Allopena *et al.* (1998) argue, the probability to fixate a target is a function of the activation level of the target: This link between fixations and lexical competition makes the technique very useful to test explicit predictions on how spoken-word recognition unfolds over time. Moreover, the results obtained with the printed-word (with words displayed on the screen) and the visual-world (with objects displayed on the screen) eye-tracking paradigms are consistent with other results in the literature obtained with different tasks (McQueen & Viebahn, 2007): This convergence suggests that our results and, more in general, results obtained with the eye-tracking paradigms, are robust and reflect the spoken-word recognition process, not any task-specific processes. However, in using an eye-tracking paradigm there are two main concerns that must be kept in mind. First, the visual display contains few stimuli, and thus the number of candidates that compete during word-recognition under these conditions is small. This situation may differ from every-day life where a target word has to be selected among many possible alternatives. Second, language processing and eye movements are strongly related, but their link is indirect. In this regard, it is important to define what kind of representation of visual and auditory inputs listeners activate, and at which level and how the visual information affects the processing of auditory information (Huettig, Rommers, & Meyer, 2011; Salverda & Tanenhaus, 2010).

In the reading aloud experiments, we employed a priming paradigm to test the activation of lexical stress in word reading. Thanks to its versatility, the priming paradigm has been used thousands of times to investigate different aspects of visual word recognition and reading aloud (see, *e.g.*, Kinoshita & Lupker, 2003). Several critical parameters must be considered when using a priming methodology, including the level of representation (*e.g.*, orthographic, phonological, or semantic), and the temporal dynamics of the prime-target interaction. In our experiments we wanted to verify whether lexical stress – a phonological feature of the word – is a suprasegmental feature, independent of the segmental level. Consequently, we employed a phonological priming paradigm (Ferrand & Grainger, 1993), manipulating the stress pattern relationship between prime and target.

Let us now consider the nature of lexical stress and its representation within different linguistic domains, as it emerges from the results reported above.

The investigations here reported explicitly assume that language production, language comprehension, and reading aloud share a subset of processes with respect to lexical stress. As for the overlap of word production and reading aloud, it has been postulated (Indefrey & Levelt, 2004; Roelofs, 2004; for evidence for shared cortical structures see, *e.g.*, Price, McCrory, Noppeney, Mechelli, Moore, Biggio, & Devlind, 2006) that they share mechanisms that allow both word retrieval and utterance articulation during language production. Thus, both speaking and reading require the construction of the word's phonological form as well as the phonological planning of the word (Roelofs, 2004). Our results are in line with this view, and suggest that the word's metrical information is computed apart from segmental information. Furthermore, they suggest that the computation of the word's metrical information takes place at the level of the phonological buffer.

But what about lexical stress in spoken word recognition? Some evidence for the overlap in the domains of word production and word comprehension comes from different sources. It has been argued that self-monitoring during the processing of an utterance takes place through a procedure that involves two separate but closely linked systems, *i.e.*, through the matching of the perceived word-form into the encoded word-form (Oppenheim & Dell, 2010; Roelofs, 2003; Roelofs *et al.*, 2007). Thus, the word-form information activated in the speech recognition system has to be similar in nature and, above all, compatible with the information activated in the speech production system. If this is true, prosodic information must be encoded and represented quite similarly in both the word comprehension and the word production systems. Further evidence for a common representation of lexical information in the production and comprehension systems comes from studies on the cortical organization of speech. It has been proposed that some brain regions are involved in both the comprehension and the production processes (Hickok & Poeppel, 2007; Menenti, Gierhan, Segaert, & Hagoort, in press). Considering the lexical level, Menenti and colleagues (in press) argue that, in both hemispheres, the posterior and anterior middle temporal gyrus, the left inferior and middle frontal gyrus are involved in lexical comprehension and production. Although having common regions for two processes does not mean using the same kind of information to perform them, the large overlap between brain regions activated during word comprehension and word production suggests that these two processes may relay on the same kind of information.

The opportunity to compare the different linguistic processes using a common perspective allows us to jointly consider how suprasegmental information is represented in language comprehension, language production, and reading aloud. In such a view, we can think about the nature of lexical stress in a unified manner, taking also into account to what extent the production and the comprehension systems differ from each other in the computation of word’s prosody.

Taken together, the experimental results reported in this thesis highlight two main issues on the nature of lexical stress: First, the autonomy of suprasegmental information from segmental information; second, the possibility of a unique mechanism that processes suprasegmental information in word comprehension, word production and word reading. Both issues, that were addressed separately, refer to how suprasegmental information is computed within different linguistic processes.

Consider the autonomy of suprasegmental information first. The independence of suprasegmental information from segmental information is a widely accepted assumption in linguistic theories (see, *e.g.*, Nespor & Vogel, 1986). In psycholinguistics, the first evidence in favor of such distinction came from studies on word production: In their theory of lexical access, following the findings of Roelofs and Meyer (1998), Levelt and colleagues (1999) postulated two different mechanisms to retrieve and compute suprasegmental and segmental information. As suggested above, the results of our experiments go in a similar direction: Suprasegmental information is stored apart from segmental information and two different mechanisms can be postulated for the computation of word prosody and word phonemes respectively. In such a view, when people listen or plan to produce a phonological unit, they process it by means of two mechanisms, each of them responsible to compute word prosody and the phonemes constituting the word, respectively. Similar proposals have been recently made by different authors. In the language comprehension literature, Cho and colleagues (2007) proposed a double mechanism for the computation of suprasegmental and segmental information in the spoken-word recognition system. They assumed that a “Prosody Analyzer” is at work at a pre-lexical level, operating in parallel with the segmental analysis performed by the recognition system. Similarly, in the reading aloud literature, several/some recent computational models have also implemented two distinct mechanisms for the computation of stress and the computation of phonemes (see, *e.g.*, the CDP++ model (Perry *et al*., 2010), or the model proposed by Sibley, Kello, & Seidenberg (2010)). In these models, suprasegmental and segmental information are computed from different components and they are assembled together later on for the production of the word. To sum up, in line with some proposals recently advanced in literature, our findings suggest that lexical stress is computed by a specific/dedicated mechanism that works pre-lexically in the spoken-word recognition system and post-lexically in the word production and reading aloud system.

As for the second issue – *i.e.*, the possibility of a unique mechanism for processing suprasegmental information *vs*. domain specific mechanisms involved in word production and comprehension – studies on the self-monitoring and on the cortical organization of speech showed that word comprehension and word production may share some mechanisms (see, *e.g*., Menenti *et al*., in press; Roelofs *et al*., 2007). As our data suggest, this might be the case for the processing of lexical stress. Although our results in spoken-word recognition (Chapter 2 in this thesis) and in reading aloud (Chapter 4 and 5 in this thesis) are not directly comparable, all together our findings may suggest the existence of a unique mechanism to compute lexical stress and, more in general, suprasegmental information. This mechanism interacts with the other linguistic mechanisms involved in word comprehension and production and it may work in parallel and contemporarily to the mechanism responsible for the computation of segmental information.

Having argued for the autonomy of suprasegmental information from segmental information, we may consider which kind of information is associated with the prosodic knowledge of a word. We have no direct data on this, but we would like to suggest that the stress pattern of a word is one of the prosodic features stored in the lexicon. Other types of prosodic information concern the acoustic-phonetic information signaling the different stress patterns, and the statistical information about the stress patterns' distribution in the language. Considering the stored acoustic-phonetic information signaling stress, in Chapter 2 we showed that listeners not only know which cues allow them to recognize whether a word bears antepenultimate or penultimate stress (by exploiting the relative difference in amplitude and duration of the vowels), but that they also possess abstract knowledge concerning the acoustic-phonetic cues related to stress. This means that Italian listeners appear to know that words with a particular stress pattern tend to have particular acoustic properties. This knowledge may be stored as part of the prosodic knowledge that people have about words and this knowledge comes into play in hearing someone speaking as well as in speaking or reading: In the former case, we need to recognize in the acoustic signal the presence of a certain stress patter to correctly understand the words; differently, in the latter case, we need to place our oral tract in order to produce the acoustic signal corresponding to the right stress pattern.

Finally, let us consider the asymmetrical distribution of the stress patterns in the language. As described above, in Italian the two main stress patterns have a strong asymmetrical distribution: 80% of words bears penultimate stress and only 18% of words bears antepenultimate stress. The asymmetry in the distribution of the two stress patterns might be part of the stored prosodic knowledge. The information that penultimate stress is the most common stress pattern in Italian might be encoded as a distributional bias. However, as shown in Chapter 2, 4, and 5, the asymmetry between penultimate and antepenultimate stress – and the distributional bias – seems to affect spoken-word recognition, but not reading aloud, especially when reading takes place through the lexical route (see also Colombo & Zevin, 2009). But if the distributional bias is part of the stress prosodic knowledge, why does it affect only spoken-word recognition?

A possible interpretation for this difference might be found in the different role that the statistical bias can have in optimizing each of the two processes. During spoken-word recognition, listeners have to identify spoken words within a continuous speech stream. To recognize words, listeners use both segmental and suprasegmental information as soon as it becomes available (Norris & McQueen, 2008; Reinisch *et al*, 2010; Tanenhaus *et al*. 1995; Warren & Marslen-Wilson, 1987). In such incremental process, the assumption of the penultimate stress as a default pattern might help Italian listeners to optimize the word-recognition process. By assigning penultimate stress by default, the initial disadvantage for those possible competitors with antepenultimate stress will be limited: Checking the acoustic-phonetic information in the speech signal will allow to detect the presence of the antepenultimate stress, and listeners can revise their original assumption. In this way, 80% of the times listeners will have computed the right stress pattern just following the default, with a great benefit for the recognition process in terms of time and efficiency. However, the situation might be different for reading aloud. When people read a word aloud, with a strong involvement of the lexical route, they activate the word lexical representation, thus making available both segmental and suprasegmental information of the target word. Since, by hypothesis, a word stress is represented irrespective of the fact that it is a penultimate stress or an antepenultimate stress, no difference is expected for the two types of stress in the retrieval and computational processes during word reading. Therefore, in such a view, the automatic assignment of a default stress would not favour the process in the case of penultimate stress words.

The fact that we exclude any effect of the statistical bias during lexical reading does not exclude that the default bias may play a role during sub-lexical reading (Colombo & Zevin, 2009): When people read a pseudoword, they cannot retrieve any stress pattern from the lexicon and they might use distributional information to assign stress to the stimulus. In this case, the default bias may be helpful to establish the pseudoword stress pattern and to optimize pseudoword reading, but only when readers cannot assign stress basing on any other distributional information driven by orthography (cf. Sulpizio, Arduino, Paizi, & Burani, *under review*).

To sum up, we propose that, in Italian, lexical stress is a stored information pertaining to the prosodic knowledge about a word. In accordance with other linguistic and psycholinguistic theories (Cho *et al.*, 2007; Levelt, 1999; Nespor & Vogel, 1986; Perry *et al.*, 2010), we have argued that suprasegmental and segmental information are partially independent from each other and that they are computed by two different mechanisms. Moreover, when considering lexical stress, we may assert that the word's prosodic information is a knowledge that specifies the set of prosodic properties of the lexical entries. People have at least three types of stored information about lexical stress: First, the word's stress pattern, which is a feature stored in the lexicon; second, the acoustic cues that define each stress pattern at the phonetic level, which people have to know both to correctly recognize and produce words; third, the asymmetrical distribution between penultimate and antepenultimate stress, which is reflected into a distributional bias toward the penultimate stress: People can exploit it to optimize spoken-word recognition and, in some cases, reading aloud.

Finally, it is worth noting that the role of lexical stress in both word comprehension and word production may vary across languages (see, *e.g.*, the difference between fixed- and free-stress languages), as well as the types of stored knowledge that people may have about it. Consider the acoustic-phonetic cues of stress as an example. As described in Chapter 1, listeners of different languages use different acoustic cues to detect word stress: While Italian listeners use mainly amplitude as stress detector, Dutch listeners use duration, and English listeners use pitch. These differences at the perceptual level will necessarily be reflected at the cognitive level, that is in the types of prosodic knowledge people have about lexical stress. Therefore, although we suggest, together with other studies in the literature, that lexical stress is part of a more general prosodic knowledge concerning the words, we have to keep in mind that specifying what types of information is contained in this knowledge remains a language-specific task.

References

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language, 88,* 148-159.

Albano Leoni, F., & Maturi, P. (1998). *Manuale di fonetica*. Roma, Italy: Carocci.

Alfano, I. (2006). La percezione dell’accento lessicale: un test sull’italiano a confronto con lo spagnolo. In Savy, R., & Crocco, C. (Eds.), *Atti del II AISV, Associazione Italiana di Scienze della Voce* (pp. 632-656). Padova, Italy: EDK Editore

Alfano, I., Savy, R., & Llisterri, J. (2009). Sulla realtà acustica dell’accento lessicale in italiano ed in spagnolo: la durata vocalica in produzione e percezione. In Romito, L, Galatà, V., & Lio (Eds.), *Atti del IV AISV, Associazione Italiana di Scienze della Voce* (pp. 22-39). Torriana, Italy: EDK Editore.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419-439.

Arciuli, J., & Cupples, L. (2003). Effects of stress typicality during speeded grammatical classification. *Language and Speech, 46*, 353–374.

References

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

Arciuli, J., Monaghan, P., & Ševa, N. (2010). Learning to assign lexical stress during reading aloud: Corpus, behavioral, and computational, investigation. *Journal of Memory & Language, 63,* 180-196.

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

Barr, D. J. (2008). Analyzing 'visual world' eye tracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457-474.

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]. www.istc.cnr.it/material/database/.

Boersma, P., & Weenink, D. (2007). PRAAT, a system for doing phonetics by computer (version 5.1.32) [computer program]. Retrieved from http://www.praat.org

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

References

Burani, C., Barca, L., & Ellis, A. W. (2006). Orthographic complexity and word naming in Italian: Some words are more transparent than others. *Psychonomic Bullettin & Review, 13,* 346-352.

Carreiras, M., Mechelli, A., & Price, C. J. (2006). Effect of Word and Syllable Frequency on Activation During Lexical Decision and Reading Aloud. *Human Brain Mapping, 27,* 963-972.

Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain and Language, 90*, 393-400.

Chateau, D., & Jared, D. (2003). Spelling-sound consistency effects in disyllabic word naming. *Journal of Memory and Language, 48,* 255-280.

Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically-driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics, 35*, 210-243.

Cholin, J., Dell, G. S., & Levelt, W. J. M. (2011). Planning and articulation in incremental word production: Syllable-frequency effects in English. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37,* 109-122.

Cholin, J., Levelt, W. J. M., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition, 99,* 205–235.

Colombo, L. (1991). The role of lexical stress in word recognition and pronunciation. *Psychological Research, 53,* 71-79.

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.

References

Colombo, L., Brivio, C., Benaglio, I., Siri, S., & Cappa, S. F. (2000). Alzheimer patients' ability to read words with irregular stress. *Cortex, 36,* 703–714.

Colombo, L., Fonti, C., & Cappa, S. (2004). The impact of lexical-semantic impairment and of executive dysfunction on the word reading performance of patients with probable Alzheimer dementia. *Neuropsychologia, 42,* 1192–1202.

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.

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

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

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech, 45,* 207-228.

Creel, S., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, *106,* 633-664.

References

Creel, S. C., Tanenhaus, M. K., & Aslin, R. N. (2006). Consequences of lexical stress on learning an artificial lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 15-32.

Cutler, A. (2005). Lexical stress. In D. B. Pisoni & R. E. Remez (eds.), *The Handbook of Speech Perception* (pp. 264-289). Oxford: Blackwell.

Cutler, A. (2008). The 34th Sir Frederick Bartlett Lecture The abstract representations in speech processing. *The Quarterly Journal of Experimental Psychology, 61*, 1601-1619.

Cutler, A. (2010). Abstraction-based efficiency in the lexicon. *Laboratory Phonology, 1*, 301-318.

Cutler, A., Eisner, F., McQueen, J. M., & Norris, D. (2010). How abstract phonemic categories are necessary for coping with speaker-related variation. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory phonology 10* (pp. 91-111). Berlin, Germany: de Gruyte.

Cutler, A., Norris, D., & Sebastián-Gallés, N. (2004). Phonemic repertoire and similarity within the vocabulary. In S. Kin & M. J. Bae (Eds.), *Proceedings of the 8th International Conference on Spoken Language Processing* (Interspeech) (pp. 65-68). Seoul: Sunjijn Printing Co.

Cutler, A., & Pasveer, D. (2006). Explaining cross-linguistic differences in effects of lexical stress on spoken-word recognition. In R. Hoffmann & H. Mixdorff (Eds.), *Speech Prosody*. Dresden: TUD press.

References

Cutler, A., Wales, R., Cooper, N., & Janssen, J. (2007). Dutch listeners' use of suprasegmental cues to English stress. In J. Trouvain & W. J. Barry (Eds.). *Proceedings of the XVI International Congress of Phonetic Sciences* (pp. 1913-1916). Dudweiler, Germany: Pirrot.

Cutler, A., & Van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech*, 44, 171–195.

Dupoux, E., Pallier, C., Sebastiàn-Gallfo, N., & Mehler, J. (1997). A destressing deafness in French. *Journal of Memory and Language, 36,* 399-421.

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience.* 21, 803–820.

Dumay, N., & Gaskell, G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18,* 35-39.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics, 67*, 224-238.

Fear, B. D., Cutler, A., & Butterfield, S. (1995). The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America, 97,* 1893-1904.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society*, *31*, 119-122.

Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of words and picture naming: The role of syllabic units. *Journal of Memory & Language, 35,* 708–723.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27,* 765-768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1, 126-52.

Garde, P. (1972). *Introduzione ad una teoria dell'accento*. Roma, Di Officina Edizioni.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12,* 613-656.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1166-1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105,* 251-279.

Goldinger, S. D. (2007). A complementary-systems approach to abstract and episodic speech perception. *Proceedings of XVI International Congress on Phonetic Sciences* (pp. 49-54). Dudweiler, Germany: Pirrot.

Gutiérrez Palma, N., & Palma Reyes, A. (2004). Lexical stress and reading: a study with children. *Electronic Journal of Research in Educational Psychology, 2*, 143–160.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8,* 393-402.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57,* 460-482.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*, 151-171.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92,* 101-144.

Kelly, M. H., Morris, J., & Verrekia, L. (1998). Orthographic cues to lexical stress: Effects on naming and lexical decision. *Memory and Cognition*, *26*, 822-32.

Kelly, M. H. (2004). Word onset patterns and lexical stress in English. *Journal of Memory and Language, 50,* 231–244.

Kinoshita, S., & Lupker, S.J.(Eds.) (2003). *Masked priming: The state of the art*. New York: Psychology Press.

Kouider, S., & Dupoux, E. (2005). Subliminal speech priming. *Psychological Science, 16,* 617-625.

Krämer, M. (2009). *The Phonology of Italian*. New York: Oxford University Press.

Ladefoged, P., Draper, M. H., & Whitteridge, D. (1958). Syllables and stress. In *Miscellanea Phonetica III* (pp. 1-14). London: International Phonetic Association, University College.

References

Laganaro, M., & Alario, F. X. (2006). On the locus of the syllable frequency effect in speech production. *Journal of Memory and Language*, *55*, 178-196.

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

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America,* 33, 451-454.

Llisterri, J., Machuca, M., de la Mota, C., Riera, M., & Rios, A. (2003). The perception of lexical stress in Spanish. *Proceedings of the 15th International Congresso f Phonetic Sciences*, Barcelona, 2023-26.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, *132,* 202-227.

Malouf, T., & Kinoshita, S. (2007). Masked onset priming effect for high-frequency words: Further support for the speech-planning account. *The Quarterly Journal of Experimental Psychology*, *60,* 1155-67.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics, 53,* 372-380.

McClelland, J. L., & Elman, J. L. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology, 18,* 1-86.

McQueen, J. M. (2007). Eight questions about spoken word recognition. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 37-53). Oxford: Oxford University Press

McQueen, J. M., & Cutler, A. (2010). Cognitive processes in speech perception. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 489-520). Oxford: Blackwell

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological Abstraction in the Mental Lexicon. *Cognitive Science, 30*, 1113-1126.

McQueen, J. M., & Viebahn, M. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology, 60,* 661-671.

Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort P. (in press). Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science*.

Miceli, G., & Caramazza, A. (1993). The assignment of word stress in oral reading: Evidence from a case of acquired dyslexia. *Cognitive Neuropsychology*, *10*, 273-296.

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology.* Dordrecht: Foris Publications.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52,* 189-234.

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian Model of Continuous Speech Recognition. *Psychological Review, 115,* 357-395.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47,* 204–238.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech-perception as a talker-contingent process. *Psychological Science, 5,* 42-46.

Obleser, J., & Eisner, F. (2009). Pre-lexical abstraction of speech in the auditory cortex. *Trends in Cognitive Sciences, 13,* 14-19.

Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: The flexible abstractness of inner speech. *Memory & Cognition, 38,* 1147-1160.

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 & Writing: An Interdisciplinary Journal, 24,* 443-461.

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.

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology VII* (pp.101-140). Berlin, Germany: Mouton de Gruyter.

Potisuk, S., Gandour, J., & Harper, M. P. (1996). Acoustic correlates of stress in Thai. *Phonetica,* 53, 200-20.

References

Price, C. J., McCrory, E., Noppeney, U., Mechelli, A., Moore, C. J., Biggio, N., & Devlin, J. T. (2006). How reading differs from object naming at the neuronal level. *NeuroImage, 29*, 643-648.

Prince, A. (1990). Quantitative Consequences of Rhythmic Organization. In K. Deaton, M. Noske, & M. Ziolkowski (Eds.), *CSL 26-II: Papers from the Parasession on the Syllable in Phonetics and Phonology* (pp. 355-398). Chicago: Chicago Linguistics Society.

Protopapas, A., Gerakaki, S., & Alexandri, S. (2006). Lexical and default stress assignment in reading Greek. *Journal of Research in Reading, 29,* 418–432.

Protopapas, A., Gerakaki, S., & Alexandri, S. (2007). Sources of information for stress assignment in reading Greek. *Applied Psycholinguistics, 28*, 695-720.

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.

Rastle, K., Kinoshita, S., Lupker, S. J., & Coltheart, M. (2003). Cross-task strategic effects. *Memory and Cognition, 31*, 867-876.

Reinisch, E., Jesse, A., & McQueen, J. M. (2010). Early use of phonetic information in spoken word recognition: lexical stress drives eye movements immediately. *The Quarterly Journal of Experimental Psychology, 63*, 1-12.

Roelofs, A. (2000). WEAVER++ and other computational models of lemma retrieval and word-form encoding. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 71-114). Sussex, UK: Psychology Press.

References

Roelofs, A. (2003). Modeling the relation between the production and recognition of spoken word forms. In N. O. Schiller & A. S. Meyer (Eds.). *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 115-158). Berlin, Germany: Mouton de Gruyter.

Roelofs, A. (2004). Seriality of phonological encoding in naming objects and reading their names. *Memory & Cognition, 32,* 212-222.

Roelofs, A., & Meyer, A. S. (1998). Metrical Structure in Planning the Production of Spoken Words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24,* 922-939.

Roelofs, A., Özdemir, R., & Levelt, W. J. M. (2007). Influences of Spoken Word Planning on Speech Recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 33,* 900-913.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition, 90,* 51-89.

Salverda, A. P., & Tanenhaus, M. K. (2010). Tracking the Time Course of Orthographic Information in Spoken-Word Recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 36*, 1108-1117.

Schiller, N. O. (2006). Lexical stress encoding in single word production estimated by event-related brain potentials. *Brain Research,1112*, 201-212.

Schiller, N. O., Fikkert, P., & Levelt, C. C. (2004). Stress priming in picture naming: An SOA study. *Brain & Language, 90,* 231-240.

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

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

Shatzman, K. B., & McQueen, J. M. (2006). Prosodic Knowledge Affects the Recognition of Newly Acquired Words. *Psychological Science, 17*, 372-377.

Sibley, D. E., Kello, C. T., & Seidenberg, M. S. (2010). Learning orthographic and phonological representations in models of monosyllabic and bisyllabic naming. *Journal of Cognitive Psychology, 22*, 650-668.

Sjerps, M. J., & McQueen, J. M. (2010). The bounds on flexibility in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 195-211.

Sluijter A. M. C., & van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100, 2471-85.

Soto-Faraco, S., Sebastian-Galles, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language, 45*, 412-432.

Stella, V., & Job, R. (2001). Le sillabe PD/DPSS. Una base di dati sulla frequenza sillabica dell’italiano scritto [The syllables PD/DPSS. A database for syllable frequency of written Italian]. *Giornale Italiano di Psicologia, 3,* 633-642, http://dpss.psy.unipd.it/files/strumenti.php.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica,* 30, 267–315.

Sulpizio, S., Arduino, L. S., Paizi, D., & Burani, C. (*under review*). Stress assignment in reading Italian polysyllabic pseudowords. *Journal of Experimental Psychology: Learning, Memory & Cognition.*

Sulpizio, S., & Job, R. (2010). L'intervento delle sillabe nei processi di produzione e riconoscimento visivo delle parole [The role of syllable in word production and visual word recognition]. *Giornale Italiano di Psicologia, 3,* 731-738.

Sulpizio, S., Job, R., & Burani, C. (in press). Priming lexical stress in reading Italian aloud. *Language & Cognitive Processes*.

Sulpizio, S., & McQueen, J. M. (2011). When two newly-acquired words are one: New words differing in stress alone are not automatically represented differently. *Proceedings of Interspeech 2011*, Florence, Italy (pp. 1385-1388).

Tagliapietra, L., & McQueen, J. M. (2010). What and where in speech recognition: Geminates and singletons in spoken Italian. *Journal of Memory and Language, 63,* 306-323.

Tagliapietra, L., & Tabossi, P. (2005). Lexical Stress Effects in Italian Spoken Word Recognition. *Proceeding in XXVII Conference of the Cognitive Science Society* (pp. 2140-2144).

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632-1634.

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.

van Donselaar, W., Koster, M., & Cutler, M. (2005). Exploring the role of lexical stress in lexical recognition. *Quarterly Journal of Experimental Psychology, 58,* 251-273.

Warren, P., & Marslen-Wilson, W. D. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception and Psychophysics, 4,* 262-275.

Zevin, J.D., & Balota, D.A. (2000). Priming and attentional control of lexical and sublexical pathways during naming. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26,* 121-135.