



# Syllable effects in a fragment-detection task in Italian listeners

Caroline Floccia<sup>1\*</sup>, Jeremy Goslin<sup>1</sup>, José Junça De Moraes<sup>2</sup> and Régine Kolinsky<sup>2,3</sup>

<sup>1</sup> School of Psychology, University of Plymouth, Plymouth, UK

<sup>2</sup> Research Unit in Cognitive Neuroscience, Université Libre de Bruxelles, Brussels, Belgium

<sup>3</sup> Fonds de la Recherche Scientifique, FRS-FNRS, Brussels, Belgium

## Edited by:

Carlo Semenza, Università degli Studi di Padova, Italy

## Reviewed by:

Cristina Romani, Aston University, UK  
Lisa D. Sanders, University of Massachusetts Amherst, USA

## \*Correspondence:

Caroline Floccia, School of Psychology, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK.  
e-mail: caroline.floccia@plymouth.ac.uk

In the line of the monitoring studies initiated by Mehler et al. (1981), a group of Italian listeners were asked to detect auditory CV and CVC targets in carrier words beginning with a CV, a CVC, or a CVG (G = geminate) syllable with variable initial syllable stress. By slowing participants reaction times (RTs), using both catch and foil trials, a syllable effect was found, partially modulated by participants' speed and stress location. When catch trials were removed in a second experiment the syllable effect was not observed, even if RTs were similar to that of the first experiment. We discuss these data in relation to the language transparency hypothesis, the nature of the pivotal consonant, and the resonance-based ART model for speech perception (Grossberg, 2003).

**Keywords:** prelexical processing, syllable, lexical access, Italian, spoken word recognition, syllable-timed languages

## INTRODUCTION

Prelexical processing refers to an intermediate stage during which the acoustic speech stream is parsed into phonological units used to contact the lexicon and access word meaning (see Pallier et al., 2001; McQueen et al., 2006). Most models of spoken word processing assume the existence of such a prelexical level of processing and representation, whose aim is to achieve perceptual constancy through the retrieval of a limited set of intermediate units (Mehler et al., 1981; McClelland and Elman, 1986; Norris, 1994; Gaskell and Marslen-Wilson, 1997; Cutler et al., 2010). Unfortunately, the variability that has been encountered as a function of the task (e.g., Tagliapietra et al., 2009), the listener's language (e.g., Cutler et al., 1986), the choice of the stimuli (e.g., Content et al., 2001) or even the instructions given to the participants, or the experimenter (Goldinger and Azuma, 2003), has prevented a consensual conclusion on the identity of these perceptual building blocks even after four decades of research. However, converging evidence for the existence of prelexical representations and their necessity in mediating lexical access (e.g., McQueen et al., 2006; see Oleser and Eisner, 2009, for neural correlates) continues to drive the search for potential candidates and identify the particular role they play interacting between levels of representation (Jusczyk, 1986) and top-down and bottom-up processes (Goldinger and Azuma, 2003). In this study we revisit a well-known paradigm in the history of the quest for a prelexical unit, the fragment-detection paradigm, and examine how new data can be accommodated by older (Sebastián-Gallés et al., 1992; Dupoux, 1993) and more recent accounts (Goldinger and Azuma, 2003) of prelexical processing.

Mehler et al. (1981) published a seminal paper indicating the use of syllabic prelexical units in spoken word processing. Their main experiment used a fragment-detection task, in which French speaking subjects had to detect a pre-specified target like/ba/in

a list of carrier words like *ba.lance*<sup>1</sup> (/balās/) or *bal.con* (/balkō/). Because targets were detected faster when they precisely matched the first syllable of the carrier (e.g., *ba* in *ba.lance*, and *bal* in *bal.con*) than when they did not (such as *ba* in *bal.con* or *bal* in *ba.lance*), an advantage known as the syllable effect, it was assumed that syllables played a prelexical role in spoken language processing. Since the original study, Mehler et al.'s (1981) paradigm has been used many times in different languages (e.g., Cutler et al., 1986; Sebastián-Gallés et al., 1992; Zwitserlood et al., 1993; Tabossi et al., 2000; Aquil, 2011) with mixed results, indicating a link between the nature of the prelexical segmentation unit and the language phonology. In particular it would appear that the syllable has greater perceptual relevance in syllable-timed languages such as Romance languages than in stress-timed languages like English, Dutch, or Arabic (for a summary see Mattys and Melhorn, 2005). The syllable may not make for a stable perceptual unit in stress-timed languages as syllable boundaries are often obscured by germination and ambisyllabicity (e.g., Treiman and Danis, 1988), whilst unstressed syllables are extremely dependent upon contextual variables (e.g., Brown, 1977). Instead it has been proposed that the regularity of stress placement in stress-timed languages allow listeners to use a metrically based segmentation strategy (the MSS, proposed by Cutler and Norris, 1988; Vroomen and de Gelder, 1995).

However, several researchers argued that the syllable effect observed in fragment detection may not necessarily reflect initial access to prelexical units (e.g., Kolinsky, 1998). Indeed, the occurrence of a syllable effect would appear to depend upon the nature of the task. For example, no syllabic effects were found when using the fragment-detection task with Spanish and Italian

<sup>1</sup>The syllable boundary is indicated with a full stop.

speakers (Sebastià-Gallés et al., 1992; Tabossi et al., 2000; but see Bradley et al., 1993 for a syllable effect in Spanish), whilst these effects were evidenced in Catalan (Sebastià-Gallés et al., 1992), or using the attention paradigm (Pallier et al., 1993; Tabossi et al., 2000) and cross-modal priming (Tabossi et al., 2000; Tagliapietra et al., 2009). More recently, Mattys and Melhorn (2005) showed that even English speakers, who consistently fail to show syllable effects in fragment detection (Cutler et al., 1986; but see Zwitserlood et al., 1993, in Dutch), display syllable-based perceptual illusions (i.e., migration errors, cf. Kolinsky et al., 1995) indicating that syllables have a greater contribution to the perception of spoken English than previously assumed. As such, when considering syllabic-effect discrepancies across tasks and languages, Kolinsky (1998) concluded that the evidence is best explained by a stage-processing approach to language differences, in which syllable effects arise as a function of the level tapped by the task. Even within the fragment-detection paradigm inter-language differences indicate that the task may tap into a variety of response codes.

Here we examine Italian listeners' behavior in a fragment-detection task by manipulating factors that have been shown to color the observation of syllabic effects in Italian and other syllable-timed languages, namely the stress status of the to-be-identified fragment, participants' reaction times (RTs), and the presence or absence of catch trials in the filler list (pairs like *bal* – *bacare* in which the target and the carrier share one or two phonemes, but should be ignored).

Syllabic effects are not consistently observed amongst syllable-timed languages in fragment-detection tasks. The concept of language *acoustic transparency*, first introduced by Sebastià-Gallés et al. (1992), accounts for the lack of such an effect in Spanish (Sebastià-Gallés et al., 1992) and Italian (Tabossi et al., 2000) by positing an increased likelihood of access to non-syllabic phonemic prelexical units, prior to the syllable, in transparent languages or segments. An acoustically transparent language (or stimulus) would be one in which the acoustic correlates of linguistic information are sufficiently specified to allow the retrieval of this information with a minimal involvement of top-down information. Transparency would depend upon the number of vowels, the presence of vocalic reduction and ambisyllabicity, and the predictability of a potential lexical accent in a given language (it might be therefore more accurate to call it “phonological transparency”). The more acoustic, phonetic, or phonological ambiguity there is (as would be found with a high number of vowels, vocalic reduction, and the presence of ambisyllabicity) the less transparent a language would be. Using these criteria, French is less transparent than Spanish, and possibly even Catalan, explaining the presence of a syllable effect in French. Italian would stand between Spanish and Catalan, due to its intermediate number of vowels (five in Spanish, seven in Italian, eight in Catalan) and the lack of reduced vowels (see Tabossi et al., 2000). In addition, within a specific language, acoustic transparency may vary as a function of lexical stress, with stressed syllables having greater transparency than unstressed ones. This would account for the lack of syllable effect with initial stressed syllables in Catalan (Sebastià-Gallés et al., 1992). In short, with sufficient acoustic transparency listeners may be able to bypass the syllable, giving rapid responses on the

basis of non-syllabic phonemic prelexical units<sup>2</sup>. More generally it has been found that rapidity of response (promoted by acoustic transparency) seems to make the syllable effect disappear. Re-analyzing the French data obtained by Mehler et al. (1981) as well as other data obtained in phoneme-monitoring studies, Dupoux (1993) observed that the original syllable effect was mainly evident in the slowest participants. Faster participants tended to respond on the basis of information from the first half of the syllable, what Dupoux refers to as a *truncation effect*<sup>3</sup> (see also Content et al., 2001). The idea is that when responses are fast enough, participants can respond at a sub-stage of prelexical processing at which the open part of the syllable (onset plus nucleus) is already available, but not the complete syllable. For later responses the full syllabic representation is calculated, and the listener goes through a mandatory syllabic stage, which is reflected in a syllable-based response. Based upon these observations, Dupoux (1993) predicted that truncation should be observed when segments are acoustically transparent, either due to the inherent characteristics of the language, or specific syllabic stress. The final case can be accounted for by interpreting stress as a length effect, as stressed syllables are usually longer than unstressed syllables, resulting in an increased truncation effect.

Thus far, truncation effects linked to acoustic transparency have only been explored in Spanish (Sebastià-Gallés et al., 1992) and Italian (Tabossi et al., 2000), each having a broadly similar high transparent rating along the scale proposed by Sebastià-Gallés et al. (1992). As in the original study of Spanish, the Italian study by Tabossi et al. (2000) failed to elicit the syllable effect in fast participants, in support of Dupoux's (1993) predictions. However, general agreement with Sebastià-Gallés et al.'s (1992) original claim for Italian would also require the elicitation of a syllable effect with slower participants. In Tabossi et al. (2000) study there was no comparison between fast and slow participants, and no conditions where the RTs of participants were slowed.

Therefore, the first aim of the present study was to establish if Italian participants exhibit the syllable effect in fragment detection when their reactions are slowed. To implement this slowing in our first experiment we included not only the usual foil trials used, e.g., by Tabossi et al. (2000), in which there is no overlap between target and carrier (e.g., *ba* followed by *tento*), but also catch trials previously used by Content et al. (2001) and Norris and Cutler (1988), with an overlap in the first consonant (e.g., *ba* – *bolo*) or vowel (*ba* – *salare*). In addition, the response pattern of the participants was analyzed as a function of RTs, as previous studies

<sup>2</sup>It is important to note that the acoustic transparency effect would predict strong syllabic effects in English, as this language has a large number of vowels, reduced vowels, ambisyllabicity, and variable lexical stress. However syllabic effects have not been reported in English using the fragment-detection task, which could be explained by the fact that English listeners use different segmentation unit to process their language (as proposed by Cutler et al., 1986). It is also possible that English listeners accommodate the high ambisyllabicity of the language by using multiple units to access the lexicon (Mattys and Melhorn, 2005).

<sup>3</sup>Theoretically, truncation effects are to be distinguished from phonemic effects, because the former refer to the use of semi-syllables, whereas the latter to the use of phonemic units. However, in fragment-detection tasks, they are very hard to distinguish, because they both lead to an advantage of CV target detection over CVC target detection, and can equally account for a disappearance of a carrier type effect for CV and CVC targets.

suggest that syllable effects are more prevalent in slower RTs (e.g., Dupoux, 1993; Content et al., 2001). We also investigated whether the syllable effect could be modulated by changes in syllabic stress, as the presence of such an effect in Italian has yet to be established. Given that Italian stands between Spanish and Catalan in terms of acoustic transparency, syllable effects should emerge irrespective of stress position, as in Spanish. However, if syllabic effects were modulated by stress they would be more likely to be observed on the less transparent segments, namely words starting with an initial unstressed syllable, as in Catalan. In Experiment 2, we focused our investigation on the role of the catch trials, removing them from the procedure of the previous experiment whilst adapting participant's instructions to keep RTs similar across experiments.

## EXPERIMENT 1

### PARTICIPANTS

Twenty-four native monolingual Italian-speaking participants from the University of Bologna were tested (aged 25.6, from 19 to 39, including three males). The data of three additional participants were rejected due to slow average RTs (above 1100 ms). Another participant was excluded because over 15% of his responses were missing.

### STIMULI

A total of 36 Italian carrier words were selected, evenly split between those with initial CV (such as *calore* and *calo*), CVC (*calvizie* and *caldo*), and CVG (*calloso* and *callo*) syllables, each having a liquid pivotal consonant (that is, a consonant which can be either the coda of a syllable as the /l/ in *cal.do* or the onset of a syllable as in *ca.lo*). Geminate consonants in Italian are considered heterosyllabic (Wiltshire and Maranzana, 1998), so that a word like *callo* should be divided as *cal.lo*. This condition provides an interesting test of the robustness of the syllable effect when the phoneme(s) constituting the pivotal cluster is the same in CV or CVC carriers. These categories of stimuli were also evenly split between those with first syllable stress (S1 words, such as *calo*, *caldo*, or *callo*), and those with second syllable stress (S2 words, such as *calvizie*, *calore*, or *calloso*). In order to respect the most common accent pattern of Italian, S1 words were disyllabic and S2 words trisyllabic. The lexical frequency of these stimuli (see Appendix) averaged 32.5 (SD = 67.2, COLFIS database for written Italian, by Laudanna et al., 1995). Both CV and CVC targets were generated for each of the carrier words, resulting in two target-carrier pairs for each carrier. Each resulting target-carrier pair was presented twice to each participant. There was no effect of stress pattern (S1 or S2) or carrier type (CV, CVC, or CVG) on lexical frequency [respectively  $F(1, 5) = 1.29$ ;  $F(2, 10) < 1$ ] and no interaction between stress pattern and lexical frequency [ $F(2, 10) < 1$ ].

These stimuli were augmented with 72 foil carriers, 36 of which had no overlap with the target (*ba* followed by *tento*), with the remaining 36 forming the basis of the catch trials. Of these 11 only shared the first consonant with their CV target (*ba* – *bolo*), 11 shared the first consonant and the first vowel with their CVC target (*bal* – *bacare* or *cal* – *cactus*) and 14 only shared the first vowel with their CV or CVC target (*ba* – *salare* or *car* – *tabella*). All target-foil pairs were also presented twice to each participant,

resulting in a total of 144 target-foil presentations. Test and foil pairs were randomized in two blocks of 144 target-stimulus pairs, whose presentation order was counterbalanced. An additional set of 10 training pairs was also selected, that did not contain any test stimulus.

To maximize the acoustic distance between the pairs the targets were recorded by a male speaker, carriers by a female speaker, both of whom were native speakers of Italian originating from Bologna. Stimuli and targets were recorded in isolation. Stimulus presentation and response feedback was mediated through the software EXPE (Pallier et al., 1997).

The mean duration of the target syllables was 385 ms (std 49 ms), and 904 ms (std 143 ms) for the stimuli. Disyllabic words (accented on the first syllable) were shorter than trisyllabic words [accented on the second syllable; 803 vs. 1004 ms;  $F(1, 35) = 80.3$ ,  $p < 0.001$ ], and overall CV words were shorter than CVC and CVG words [CV: 800 ms; CVC: 967 ms; CVG: 944 ms;  $F(2, 35) = 21.8$ ,  $p < 0.001$ ]. These differences were found across the two stress patterns [ $F(2, 35) = 2.25$ ,  $p = 0.12$ ]. The initial consonant-vowel portion of the stimuli (such as *ba* in *balia* or *baldo*) was shorter in trisyllabic words than disyllabic words [195 vs. 338 ms;  $F(1, 35) = 86.0$ ,  $p < 0.001$ ]. However its duration was identical across all stimuli [CV stimuli: 289 ms; CVC: 250 ms; CVG: 260 ms;  $F(2, 35) = 2.22$ ;  $p = 0.12$ ], and this was found whatever the stress pattern of the words [interaction between stress type and stimuli:  $F(2, 35) < 1$ ]. Regarding the initial consonant-vowel-consonant portion (such as *bal* in *balia* and *baldo*), again it was shorter in trisyllabic words than in disyllabic words [281 vs. 457 ms;  $F(1, 35) = 150.2$ ,  $p < 0.001$ ], but it was similar across all stimuli [CV stimuli: 378 ms; CVC stimuli: 346 ms; CVG stimuli: 382 ms;  $F(2, 35) = 2.50$ ,  $p = 0.10$ ] and this was found across the two stress patterns [ $F(2, 35) = 1.67$ ].

### Procedure

The standard fragment-detection task was used in this experiment, in which each auditory target was presented before a paired carrier word. The delay between target and carrier word was 1000 ms, with pairs presented every 3000 ms. Participants were asked to make a speeded decision as to whether the target was found at the beginning of the carrier by pressing a response button with their favored hand when there was a match.

### RESULTS

Of the 3456 no-go trials, 86 false alarms were reported, an average of 2.5% of responses with error rates for participants ranging between 0 and 9.0%. For the 3456 expected go responses 1.6% were missed. For further analyses an initial rejection of the slowest 5% of responses cumulated across participants was applied (1175 ms), followed by a rejection of responses above or below 2.5 standard deviations of each subject average, consisting of 1.2% of responses.

Three within-participant factors were considered in analyses of variance (ANOVA) performed by participants ( $F_1$ ) and by items ( $F_2$ ): stress position (S1: carriers stressed on the first syllable, or on the second, S2), target type (CV or CVC), and carrier type (CV, CVC, or CVG).

A main effect of stress position was observed,  $F_1(1, 23) = 9.45$ ,  $p = 0.005$ ,  $F_2(1, 5) = 7.81$ ,  $p = 0.038$ , showing that words stressed

on the first syllable were processed slower than those with second syllable stress (655 vs. 629 ms). There was also a main effect of target type,  $F_1(1, 23) = 12.3$ ,  $p = 0.02$ ,  $F_2(1, 5) = 14.5$ ,  $p = 0.012$ , revealing that CV targets were detected faster than CVC ones (625 vs. 659 ms). No main effect of carrier was observed,  $F_1(2, 46) < 1$ ,  $F_2(2, 10) < 1$  (CV words averaging 642, 644 ms for CVC words, and 640 ms for CVG words).

The interaction between stress position and target type was significant,  $F_1(1, 23) = 45.7$ ,  $p < 0.0001$ ,  $F_2(1, 5) = 10.38$ ,  $p = 0.023$ , showing an advantage of CV targets over CVC targets in S1 words [622 vs. 688 ms,  $F_1(1, 23) = 34.3$ ,  $p < 0.001$ ,  $F_2(1, 5) = 83.7$ ,  $p < 0.001$ ], but not in S2 words [629 vs. 630 ms,  $F_1(1, 23) < 1$ ,  $F_2(1, 5) < 1$ ].

The interaction between stress position and carrier type was significant by participant,  $F_1(2, 46) = 7.97$ ,  $p = 0.001$ ,  $F_2(2, 10) < 1$ , showing that amongst S2 words, CV carriers were processed faster than CVC and CVG carriers [614 vs. 638 ms and 636 ms respectively,  $F_1(2, 46) = 2.75$ ,  $p = 0.074$ ,  $F_2(2, 10) < 1$ ]. Amongst S1 words, CV carriers were processed more slowly than CVC and CVG carriers [669, 651, and 645 ms respectively,  $F_1(2, 46) = 3.61$ ,  $p = 0.035$ ,  $F_2(2, 10) < 1$ ].

The interaction between carrier and target was significant by participant and marginal by item,  $F_1(2, 46) = 3.33$ ,  $p = 0.045$ ,  $F_2(2, 10) = 3.69$ ,  $p = 0.063$ . As can be seen in **Figure 1**, it appeared that CV targets were detected faster in CV carriers than in CVC carriers or CVG carriers (617, 628, and 631 ms respectively) but this was not significant.

CVC targets were detected equally fast in CV than in CVC carriers [667 vs. 660 ms,  $F_1(1, 23) < 1$ ,  $F_2(1, 5) = 1.58$ ], or in CVC carriers as compared to CVG carriers [660 vs. 650 ms,  $F_1(1, 23) = 2.70$ ,  $p = 0.11$ ,  $F_2(1, 5) = 4.06$ ,  $p = 0.10$ ], but tended to be detected slower in CV carriers than in CVG carriers [667 vs. 650 ms,  $F_1(1, 23) = 3.75$ ,  $p = 0.065$ ,  $F_2(1, 5) = 8.10$ ,  $p = 0.036$ ].

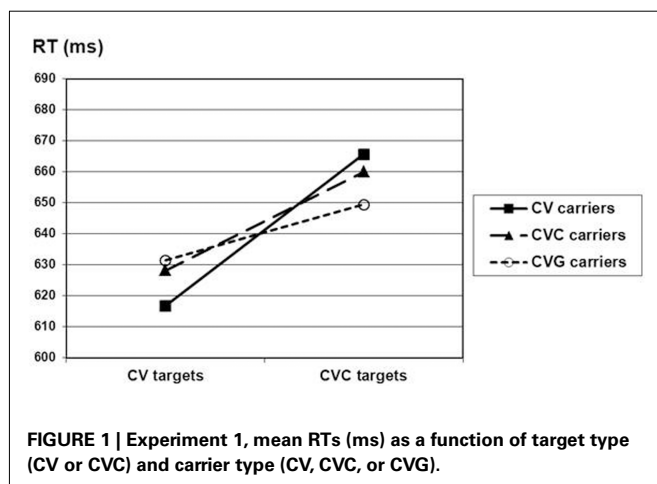
The triple interaction between stress position, carrier, and target was significant by participant,  $F_1(2, 46) = 4.11$ ,  $p = 0.023$ ,  $F_2(2, 10) = 1.44$ . The pattern of results for stressed and unstressed segments displayed in **Figure 2** suggests a syllable effect in unstressed segments [interaction between targets and carriers in unstressed segments:  $F_1(2, 46) = 7.08$ ,  $p = 0.002$ ,  $F_2(2, 10) = 3.55$ ,

$p = 0.068$ ] while there is no such effect for the stressed segments,  $F_1(2, 46) < 1$ ,  $F_2(2, 10) < 1$ . This pattern of results strongly resembles what was found by Sebastià-Gallés et al. (1992) in Catalan, although these authors did not provide the value of the triple interaction between stress location, target, and carrier type, and displayed the results for each stress location separately.

To shed further light upon the different behavior observed for stressed and unstressed segments we also conducted an examination of the distribution of miss rates and false alarms. No significant effect was found on miss rates, but false alarms in the no-go catch trials revealed very low error rates for pairs of target-carrier sharing the first vowel (such as *ca* – *rana*; mean false alarm rate: 0.15%), whereas more false alarms were found for pairs sharing the first consonant (such as *ba* – *bolo* or *bar* – *busta*; mean false alarm rate: 2.5%;  $t(23) = 2.88$ ,  $p = 0.009$ ) and many more for pairs sharing the first two consonants (such as *bal* – *banca* or *bal* – *bacare*; mean false alarm rate: 16.7%;  $t(23) = 4.92$ ,  $p < 0.001$ ).

The simple comparison of slow and fast subjects assumes that each participant's responses are closely distributed around its mean, which is rarely the case. Therefore, following Dupoux's (1993) procedure, for each participant and target type we based our analyses upon the differences in target detection time between different carriers (e.g., RT for detecting *bar* in *barile* minus *bar* in *barcone*). This allowed us to compute the effect of the carrier upon both CV and CVC target detection latencies for each participant (see also Content et al., 2001). The size of the carrier effect for each of the target types was then analyzed in relation to the average RTs for each target type and carrier pair. For example, the average RT of 450 ms for the target *ba* across CV and CVG carriers *barile* and *barrito* for a participant was compared to his/her 34 ms advantage for *ba* detection in *barile* as compared to *barrito*. These analyses revealed a significant positive correlation between the size of the CV target-carrier effect and average RT [CV vs. CVC carriers:  $r(286) = 0.18$ ,  $p = 0.002$ ; CV vs. CVG carriers:  $r(286) = 0.21$ ,  $p < 0.001$ ]. In other words, it is more likely to observe that *ba* is detected faster in *baleno* than in *balcone* or *balletto* in slow responses. Contrastively, for CVC targets the only significant correlation between the size of the carrier effect and average RT was negative (CV vs. CVC carriers:  $r(286) = -0.13$ ,  $p = 0.032$ ; CV vs. CVG carriers:  $r(286) = -0.03$ ). In other words, it is more likely to observe that *bal* is detected faster in *balcone* or *balletto* than in *baleno* in fast responses. In addition, as would be predicted by the similar pattern of results for both CVC and CVG carriers, time to detect either CV or CVC targets in CVC vs. CVG carriers was independent of average RT,  $r(286) = 0.01$  and  $=0.10$  respectively. These analyses suggest that the two components of the syllabic interaction, that is, the advantage of CV detection in CV carriers over CVC or CVG carriers on one hand, and the advantage of CVC detection in CVC or CVG carriers over CV carriers on the other hand, do not emerge simultaneously in the time course of word processing (see also Dupoux, 1993; Content et al., 2001).

It was also found that carriers stressed on the first syllable were processed significantly slower than carriers stressed on the second syllable (655 vs. 629 ms, respectively). This is at odds with other results reporting an advantage of stressed over unstressed segments. Such an advantage was observed in fast Spanish listeners



(369 ms for S1 words vs. 383 ms for S2 words, significant by participant only), while slow Spanish participants or Catalan participants did not show any effect at all (Sebastià-Gallés et al., 1992). For Italian listeners, Tabossi et al. (2000) reported in their first and second experiments that S1 words were processed faster than S2 words (a significant difference of 14 ms in Experiment 1, and 21 ms in Experiment 2). Usually, the advantage of stressed segments over unstressed segments is interpreted as an acoustic transparency effect, stressed segments providing more straightforward phonetic information than unstressed ones. However, RTs in fragment-detection tasks are not only related to acoustic transparency, but also to fragment duration within the carrier words: longer fragments naturally delay the availability of the critical information. Because stressed fragments in Italian are by definition longer than unstressed fragments, RTs in fragment-detection

tasks can be seen as the result of the combination of at least two components: the advantage of stressed segments over unstressed ones due to enhanced acoustic transparency, and the disadvantage of stressed segments over unstressed ones due to their longer duration.

Indeed it appears that the difference in duration between stressed and unstressed fragments was much more important in our study than in the Tabossi et al. (2000) study. Specifically, the mean CV duration was 338 ms in stressed fragments, and 195 ms in unstressed fragments, and the mean CVC duration was 457 ms in stressed fragments and 281 ms in unstressed fragments (see Figure 3). In the Tabossi et al.'s study (Experiment 2) the duration of stressed and unstressed portions was not very different (mean CV duration: 197 ms in S1 words vs. 120 ms in S2 words; mean CVC duration: 271 ms in S1 words vs. 192 ms in S2 words).

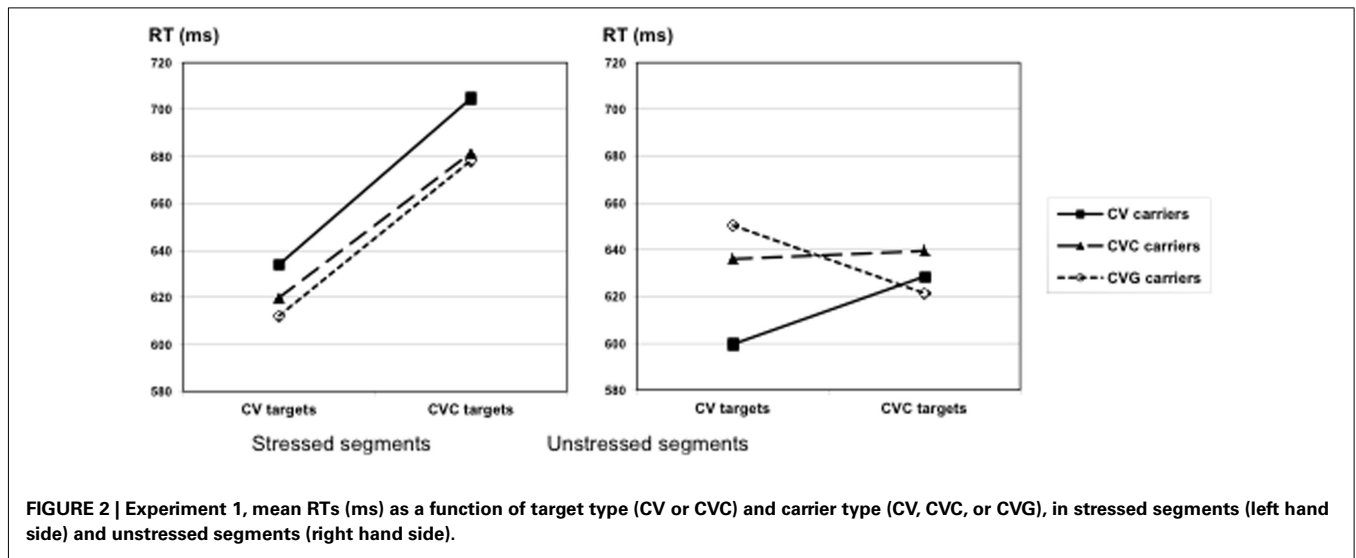


FIGURE 2 | Experiment 1, mean RTs (ms) as a function of target type (CV or CVC) and carrier type (CV, CVC, or CVG), in stressed segments (left hand side) and unstressed segments (right hand side).

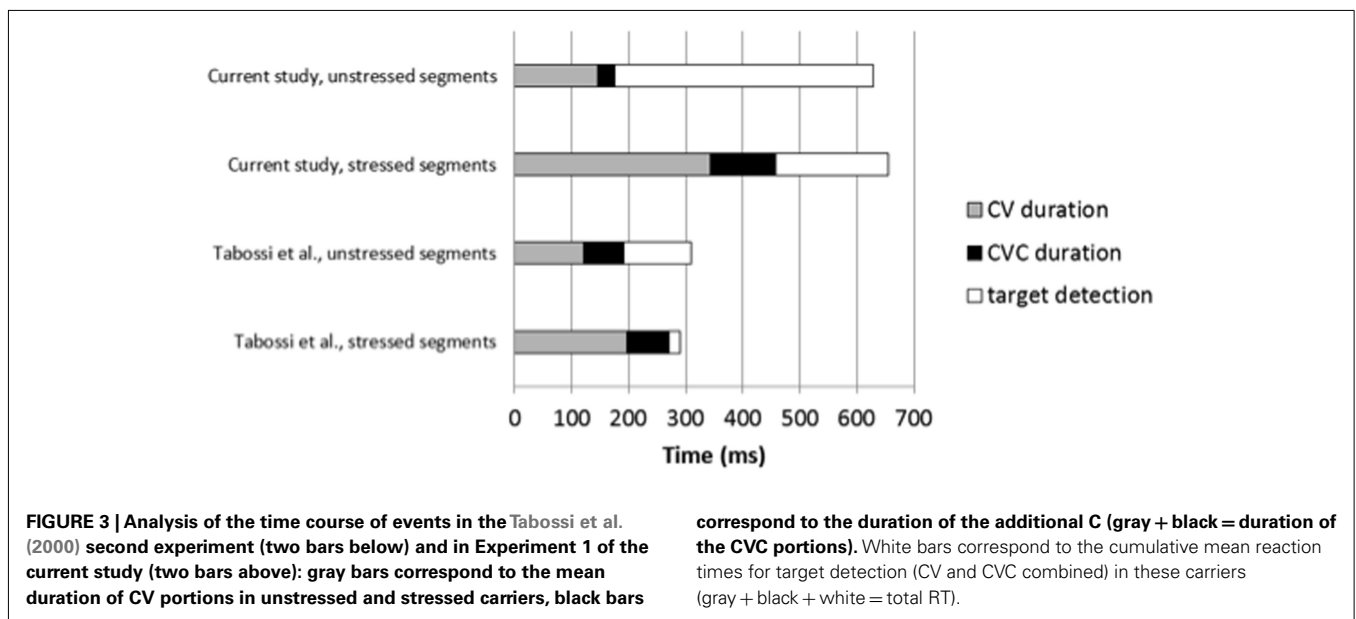


FIGURE 3 | Analysis of the time course of events in the Tabossi et al. (2000) second experiment (two bars below) and in Experiment 1 of the current study (two bars above): gray bars correspond to the mean duration of CV portions in unstressed and stressed carriers, black bars correspond to the duration of the additional C (gray + black = duration of the CVC portions). White bars correspond to the cumulative mean reaction times for target detection (CV and CVC combined) in these carriers (gray + black + white = total RT).

This was very likely due to the use of disyllabic and trisyllabic items in our study, as opposed to only trisyllabic words in the Tabossi et al.'s study. Therefore, in our study the disadvantage of stressed segments over unstressed segments due to duration difference may have overwhelmed the advantage of stressed segments due to acoustic transparency, resulting in a slight advantage to the processing of unstressed carriers.

## DISCUSSION OF EXPERIMENT 1

This first experiment shows that a syllable effect can be obtained with Italian participants when they are slowed down by the adjunction of catch trials. This complements the Tabossi et al. (2000) study, where there were no catch trials and participants responded faster, and which did not find a syllable effect. In addition, our analyses show that the size of the syllable effect is colored by participants' RTs, and by the stress status of the to-be-retrieved syllable. They also show that the syllabic effect is stronger in CV–CVG carrier pairs like *baleno* – *balletto* than in typical CV–CVC pairs like *baleno* – *balcone*, as revealed by the stronger advantage of CVC target detection in CVG carriers over CVC carriers (as compared to CV carriers). When the pivotal phoneme from the CV carrier (*baleno*) is the same as the pivotal cluster of phonemes from the CVG carrier (*balletto*), the syllabic-effect emerges more robustly, perhaps because the preservation of the phonemes' identity at the syllable junction helps reducing the between-item variability.

However, before discussing the implications of these findings for models of prelexical processing, we first needed to determine whether we could exhibit a disappearance of the syllable effect, as Tabossi et al. (2000), using the same procedure and stimuli, but different conditions. Indeed several methodological differences could explain the discrepancies between this study and ours, such as the modality (we used an auditory presentation of the primes whereas they used a visual presentation), the use of disyllabic items here as words stressed on the first syllable instead of trisyllabic ones in Tabossi et al., or the nature of the pivotal consonants (liquids here, vs. liquids and nasals in Tabossi et al.).

Therefore in Experiment 2, we tested a new group of Italian participants with the same test material, but removed all catch trials (such as *bal* – *bacare*). In these conditions, we predicted an overall advantage of CV target detection over CVC target detection, as was reported by Tabossi et al. (2000) and by Sebastián-Gallés et al. (1992) with fast Spanish participants. However, we took this opportunity to examine also the role of catch trials as opposed to absolute RT values in modulating the observation of a syllabic effect.

## EXPERIMENT 2

Catch trials are an important methodological feature of the fragment-detection task, because they are meant to slow participants down, as speed is perhaps the critical dimension in the observation of syllabic vs. truncation effects (see Dupoux, 1993). The use of catch trials varies from one study to the other, and there does not seem to be a clear link between the presence of these foils and the observation of a syllabic effect in syllable-timed languages. Mehler et al. (1981), Sebastián-Gallés et al. (1992), Tabossi et al. (2000), and Content et al., 2001, Experiment 1a) did not use catch trials, but the first two of these studies reported a syllabic effect

while the last two did not. Bradley et al. (1993) and Content et al., 2001, Experiment 1b and 2) did use catch trials, but the former reported a syllabic effect while the latter only found it in restricted conditions (blocked vowels or blocked pivotal consonant, and only with liquid pivotal consonants). Only the study by Content et al. (2001) has systematically examined the impact of catch trials upon the observation of a syllabic effect with the same stimuli and population (French). All the other previously mentioned studies have tested one condition only (catch trials or no catch trials), or used other techniques to slow participants down, such as asking them to remember some words in parallel with the task (as in Sebastián-Gallés et al., 1992). Yet, beyond their direct influence on RT values, the presence or absence of catch trials could be partially responsible for the observation of truncation effects. Without catch trials, participants could answer on the basis of a partial match between the prime and the carrier, and could even answer using the first phoneme of the carrier. With catch trials, this strategy is not viable since it would lead to many false alarms, behavior observed by Content et al., 2001, Experiment 1) in addition to slower RTs. However, since the usage of catch trials and the associated changes in RTs have been confounded in Content et al. (2001), it is difficult to determine whether the truncation vs. syllabic effects are due to one or both of these factors. In our second experiment we examine the role of catch trials in the observation of syllabic effects by removing catch trials but keeping RTs at approximately the same level as in the previous experiment by adapting the instructions given to the participants (ascertained in a pilot experiment). If truncation effects are only obtained when participants answer rapidly then they should not be observed in Experiment 2, and a syllabic effect should still be found. On the other hand if truncation effects are partially the result of participants' strategy, then the absence of catch trials should result in truncation effects and therefore no syllabic effect should be obtained.

## PARTICIPANTS

Twenty-four native monolingual Italian-speaking participants from the University of Pavia were tested (aged 24.6 years from 22 to 34 years, including five males). The data of five additional participants were rejected due to slow average RTs (above 1100 ms). Another participant was excluded because over 15% of his responses were missing.

## STIMULI AND PROCEDURE

As in Experiment 1, apart from the foil list whose set was reduced to 35 pairs with no overlap with the target (*ba* – *tento*). In addition, the instructions given to the participants did not emphasize the requirement for a fast answer. This contrasts with the previous experiment in which the experimenter made sure participants were aware of the necessity to speed up during the task. Pilot experiments ensured us that this simple change resulted in RTs at the same level as in the previous experiment, despite the removal of all catch trials.

## RESULTS

Of the 1680 no-go trials, 13 false alarms were reported, an average of 0.8% of responses with error rates for participants differing between 0 and 5.7%. For the 3456 expected go responses 0.5%



were missed. An initial rejection of the slowest 5% of responses cumulated across participants was applied (1143 ms), followed by a rejection of 1.2% of responses above or below 2.5 standard deviations of each subject average.

Three within-participant factors were considered in ANOVA performed by participants ( $F_1$ ) and by items ( $F_2$ ): stress position (S1: carriers stressed on the first syllable, or on the second, S2), target type (CV or CVC), and carrier type (CV, CVC, or CVG).

The effect of stress position was not significant,  $F_1(1, 23) = 1.57$ ,  $F_2(1, 5) < 1$  but, as in Experiment 1, did show a slight tendency for slower processing on words with stressed rather than unstressed initial syllables (662 vs. 656 ms). There was a main effect of target type,  $F_1(1, 21) = 11.11$ ,  $p < 0.05$ ,  $F_2(1, 5) = 34.04$ ,  $p = 0.002$ , revealing that CV targets were detected faster than CVC ones (646 vs. 672 ms). As can be seen in **Figure 4**, CV targets were detected faster than CVC targets in all three types of carriers, in CVC carriers [637 vs. 668 ms,  $F_1(1, 23) = 6.60$ ,  $p = 0.017$ ,  $F_2(1, 5) = 45.60$ ,  $p = 0.001$ ] and in CVG carriers [645 vs. 675 ms,  $F_1(1, 23) = 10.20$ ,  $p = 0.004$ ,  $F_2(1, 5) = 14.45$ ,  $p = 0.013$ ], but it was not significant in CV carriers [656 vs. 672 ms,  $F_1(1, 23) = 2.22$ ,  $p = 0.15$ ,  $F_2(1, 5) = 2.92$ ,  $p = 0.15$ ].

No main effect of carrier was observed,  $F_1(2, 46) = 1.57$ ,  $F_2(2, 10) < 1$  (CV words averaging 664 ms, 653 ms for CVC words, and 660 ms for CVG words), nor were there any significant interactions [stress position and target type,  $F_1(1, 23) = 2.72$ ,  $p = 0.11$ ,  $F_2(1, 5) = 1.34$ ; stress position and carrier type,  $F_1(2, 46) < 1$ ,  $F_2(2, 10) < 1$ ; carrier and target type,  $F_1(2, 46) < 1$ ,  $F_2(2, 10) < 1$ ].

The triple interaction between stress position, carrier, and target was not significant either,  $F_1(2, 46) < 1$ ,  $F_2(2, 10) < 1$  (see **Figure 5**), and visual inspection reveals no trace of a tendency toward a syllabic interaction.

As in Experiment 1, the size of the carrier effect for each target type was analyzed in relation to the average RTs for each target type and carrier pair. No correlation was found between the size of the CV target effect and average RT. In other words, time to detect a CV target in a CV carrier or a CVC carrier was similar no matter the average RTs ( $r(288) = 0.009$ ). The same was found for all other correlations between the size of the CV or CVC target effect and the average RT for all pairs of carriers (CV vs. CVC, CV vs. CVG, and CVC vs. CVG). Therefore the same patterns of results were found in slow and fast responses.

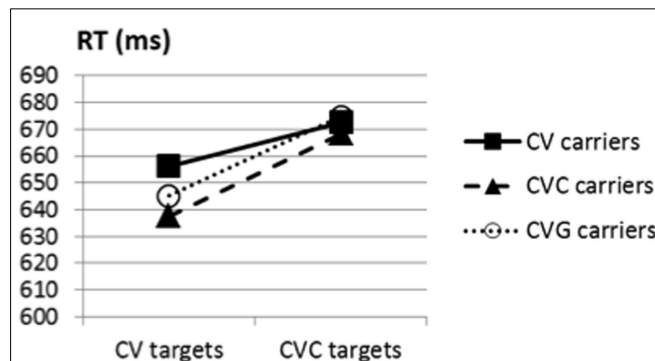
### COMPARISON OF THE TWO EXPERIMENTS

To test the significance of the difference between experiments we conducted an additional ANOVA analysis with a between-participant factor of experiment and within-participant factors (stress position, target type, and carrier type) by participants ( $F_1$ ) and by items ( $F_2$ ). First, the main effect of experiment was not significant, showing that overall RTs were at the same level in both experiments [642 ms in Experiment 1 vs. 659 ms in Experiment 2,  $F_1(1, 46) < 1$ ;  $F_2(1, 10) = 1.62$ ]. The triple interaction between stress location, target type, and experiment was significant by participant [ $F_1(1, 46) = 16.90$ ,  $p < 0.001$ ;  $F_2(1, 10) = 4.0$ ,  $p = 0.073$ ], due to the advantage of CV over CVC target detection found in all conditions except with unstressed syllables in Experiment 1. The triple interaction between stress, carrier type, and experiment was also significant by participant [ $F_1(2, 92) = 4.57$ ,  $p = 0.013$ ,  $F_2(2,$

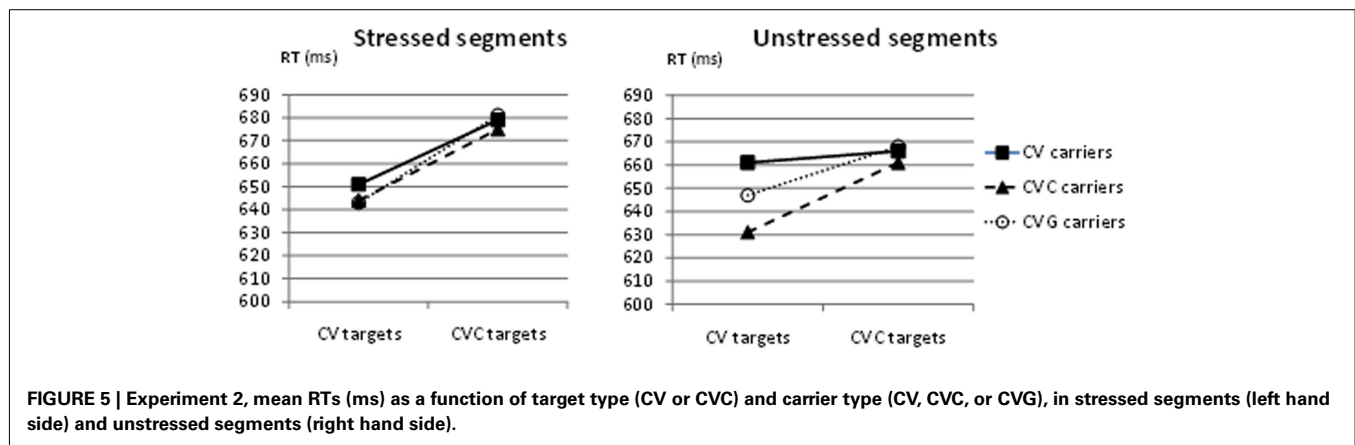
20)  $< 1$ ]. This was caused by CV carriers being processed faster than CVC and CVG carriers with unstressed syllables in Experiment 1, whereas for all other conditions the three types of carriers were processed equally fast. Finally, a significant by participant interaction between carrier type, target type, and experiment was found [ $F_1(2, 92) = 3.25$ ,  $p = 0.043$ ,  $F_2(2, 20) = 2.18$ ,  $p = 0.14$ ]. No other interaction was found to be significant. Given that the results in both experiments seem to diverge mainly for words with an initial unstressed syllable (S2), we examined the effect of experiment on these words only. When considering only CV and CVC carriers, the interaction between carrier type, target type, and experiment was significant by participant [ $F_1(1, 46) = 4.21$ ,  $p = 0.046$ ,  $F_2(1, 10) < 1$ ]. When considering only CV and CVG carriers, again the interaction between carrier type, target type, and experiment was also significant by participant [ $F_1(1, 46) = 5.59$ ,  $p = 0.022$ ,  $F_2(1, 10) = 2.69$ ,  $p = 0.13$ ]. Therefore, it would appear that the target effects observed for CV vs. CVC carriers and CV vs. CVG carriers were different in both experiments: a syllabic effect was obtained in Experiment 1, whereas an advantage of CV target detection was observed in Experiment 2.

### GENERAL DISCUSSION

Sebastià-Gallés et al. (1992) hypothesized that the likelihood to observe a syllable effect in a fragment-detection task is related to the transparency of the language mastered by the listeners. On the scale of acoustic transparency proposed by these authors Spanish is more transparent than Catalan, with Italian standing possibly between them. Therefore, syllable effects should be observed in Italian when participants are relatively slow, but not in faster participants where responses are liable to truncation (Dupoux, 1993). Support for this hypothesis in Italian was provided in a study by Tabossi et al. (2000), which failed to elicit a significant syllable effect in fast participants. In the present study we provide complementary evidence to support the original hypothesis by eliciting a syllable effect in Italian participants that have been slowed down by the use of catch trials. Thus, in-line with the predictions raised by Dupoux and by Sebastià-Gallés et al., Italian participants would appear to behave similarly to Spanish listeners, with the syllable effect observed at relatively slow RTs (average of 642 ms in our study) but not when participants respond faster (around 350 ms



**FIGURE 4 | Experiment 2, mean RTs (ms) as a function of target type (CV or CVC) and carrier type (CV, CVC, or CVG).**



**FIGURE 5 | Experiment 2, mean RTs (ms) as a function of target type (CV or CVC) and carrier type (CV, CVC, or CVG), in stressed segments (left hand side) and unstressed segments (right hand side).**

in Tabossi et al., 2000). We also found support for another of Sebastián-Galles et al.'s findings, in that the syllable effect was more likely to be observed on less acoustically transparent segments, that is, for words with an initial unstressed initial syllable, and not when the first syllable was stressed.

Further analyses of slow and fast responses revealed a different pattern of correlation between RTs and effect sizes depending on whether the target was CV or CVC. Both Dupoux (1993) and Content et al. (2001) have discussed the possibility that the two components of the syllabic effect, namely the CV target advantage in CV carriers, and the CVC advantage in CVC carriers, might capture different detection or perceptual mechanisms. Content et al. (2001) conducted regression analyses between measures of phoneme durations and response latencies, and concluded that the carrier effects for CV and CVC targets resulted from different underlying processes. For CVC targets an 85% prediction model was based upon perceptual estimates of phonemic durations. That is, the time to detect CVC targets was mostly related to the duration of initial phonemes in the carriers (time to vowel onset plus  $\frac{1}{2}$  the time from vowel onset to consonant onset), meaning that the speed of the response was related the time it takes for the phoneme to be produced. Contrastively, the carrier effect for CV targets was more likely to be explained through a model based on the syllabic structure of the carrier together with perceptual and acoustic estimates of phoneme duration, although the influence of the latter component was far less salient than in CVC target detection. If the key component of the syllable effect is the advantage of CV target detection in CV carriers over CVC carriers, then it should be particularly sensitive to speed of response (as the syllable effect as a whole is meant to be modulated by speed of response). This was supported by our observations in the first experiment, where CVC target detection did not differ significantly over time between the carrier types, whereas C target detection was more likely to be found on slow RTs. This is similar to that reported by Dupoux (1993) in French, who claimed that carrier effects are mostly attributed to CV rather than CVC targets.

However, results of Experiment 2 suggest a slightly more complex picture. In this second experiment we removed all the catch trials, as was done in Tabossi et al.'s, 2000; see also Mehler et al., 1981; Sebastián-Gallés et al., 1992), but modified the instructions to obtain RTs at the same level as in Experiment 1. This

was implemented to determine whether truncation effects are only found when participants speed up, or whether they are partially due to the strategy that can be used during the task. Indeed the absence of catch trials should allow participants to respond on the sole basis of the first phoneme(s), therefore encouraging truncation effects. It was found that although main RTs in both experiments did not differ significantly, the pattern of results with unstressed segments was different. The syllabic effect observed in Experiment 1 with unstressed syllables disappeared in Experiment 2. Instead we observed a general advantage of CV target over CVC target detection. This pattern of results would suggest that it is the temporal relation between the availability of critical information at the moment of decision that modulates the RTs leading to the syllable effect, more than the absolute RT (fast or slow). In short, the presence of the syllable effect would appear to be related to participants' focus of attention and/or strategy.

As none of the "no-go" pairs (as in *bal – tento*) in Experiment 2 have any overlap, whilst all "go" pairs share the first phoneme (as in *bal – baleno*), decisions could theoretically be made on the very first phoneme of the carriers. However, our results do not fit that pattern, as in that case RTs should be identical across all conditions as it should take just approximately the same time to detect/b/ from *ba* or *bal* in *baleno* or *balcone*. Another possibility would be to ignore the third phoneme in CVC targets, and only retrieve the first two phonemes, which are always shared by the target and the carrier (this strategy was also suggested by Dupoux, 1993). In this case, one would also expect identical RTs for CV and CVC detection. Again, this is not what we observed. Rather, participants use the entire phonemic string provided in the target, as shown by the general advantage of CV targets over CVC targets. Most importantly, the time course of responses in relation to phoneme durations in Experiment 1, as shown in Figure 3, shows that on average decisions are made long after the entire first syllable has been heard, that is, after the critical information is available.

In Experiment 1 the presence of catch trials negates the use of a strategy based upon the simple matching of the initial two or three phonemes of the target and carriers. In this situation, participants fell back on the syllable as a more reliable source of information, especially with words with an unstressed first syllable. This was also visible in the analyses of the distribution of false alarms in Experiment 1, which showed fewer errors in unstressed segments



than stressed segments, reinforcing the idea that the syllabic strategy would lead to a more reliable outcome. However, the fact that the CV detection advantage in CV carriers over CVC or CVG carriers is mainly found in slow RTs (Experiment 1) suggests that this syllabic strategy is the outcome of a late perceptual construct, also proposed by Dupoux (1993).

Therefore, we can rephrase our initial conclusion as follows: our results are consistent with the concept that in Italian the key component of the syllable interaction (the carrier effect for CV detection) is more likely to emerge at slower response times, provided that the task structure prevents participants from using a phonemic strategy. This finding does not necessarily indicate mandatory syllable processing in Italian, but rather that the structure of the task allows participants to tap on different levels of representation or processing (e.g., Kolinsky, 1998).

Another issue that should be noted in the design of stimuli used in this experiment is that all pivotal consonants consisted of liquid clusters, as in the seminal study of Mehler et al. (1981). However, in a wider study using non-words, Content et al. (2001) reported that robust syllable effects were only evident in carriers with a liquid pivotal consonant, but not if it was a fricative or an obstruent. In the study of Tabossi et al. (2000), 4 nasals and 1 liquid formed the pivotal consonants in their first experiment, with a mix of 12 nasals and 4 liquids in their second. Would this imply that the lack of syllable effect in the previous studies of Italian could be due to the use of non-liquid pivotal consonants?

In Italian and French, syllabification of CC consonantal clusters follows broadly similar rules (e.g., McCrary, 2002; Goslin and Floccia, 2007). If one distinguishes clusters starting with liquid (L-clusters), fricative (F-clusters), plosive (P-clusters), or nasal (N-clusters) consonants, it turns out that only L and N-clusters are systematically heterosyllabic, whereas F-clusters and P-clusters can be tauto or heterosyllabic, depending on the nature of the following consonant (in Italian for example *lapsus* and *ictus* syllabifies as *lap.sus* and *ic.tus*, McCrary, 2002). Therefore, L-clusters or N-clusters can never be found at word onset in any of these languages<sup>4</sup>, contrary to other types of clusters. When facing an N- or L-cluster, the French or Italian listener can assume safely that this cluster onset does not match a word onset, giving them an unambiguous cue for syllabic segmentation.

Table 1 shows the distribution of consonant clusters in Italian and French, revealing that in French L-clusters are nearly 10 times more frequent than N-clusters, whilst in Italian the distribution of these clusters are approximately equal. Therefore, whilst it would appear that L-clusters might have a privileged status in French because of their relative unique position as a heterosyllabic cluster, this role is equally filled by both L- and N-clusters in Italian. In light of these statistics it seems unlikely that Tabossi et al. (2000) failure to report any syllable effect would be entirely due to their use of a mix of N- and L-clusters, in contrast to our only using L-clusters.

Finally, touching upon the modulation of syllable effects due to stress location, we found a similar pattern of results in Italian as

observed by Sebastià-Gallés et al. (1992) in Catalan. That is, the observation of a syllable effect when the to-be-analyzed segment is unstressed, and a phonemic effect when the same segment is stressed.

We suggested earlier that the disappearance of a syllabic effect in Experiment 2 is due to participants moving to a phonemic matching process, a strategy allowed by the lack of catch trials in the experiment. However, when catch trials are present, as in Experiment 1, why would participants tend to revert to a phonemic strategy with stressed initial segments, and a syllabic strategy with unstressed segments? A possibility for this divergence could be in the preponderance of information arising from prelexical processing leading up to the decision process. In Experiment 2, in which it is not necessary to “listen carefully” given that there are no catch trials, priority can be given to a simple pattern matching strategy, thus the longer the overlap between the prime and the target, in terms of duration and/or in terms of amount of information, the longer it takes to respond. In contrast, in Experiment 1 it is necessary to “listen carefully” because of the catch trials, and in that case, information coming from the prelexical syllabic segmentation level is taken into account as well as information coming from phonemic matching. This competition results in the observation of a syllabic effect when the information about prelexical syllabic processing is available *before* that given by pattern matching. Why does this happen for unstressed segments only? It is likely that it is because unstressed segments are shorter, allowing relatively rapid syllable segmentation and syllabic processing. In contrast, the reliability of phonemic matching increases as acoustic evidence is accumulated. In this context, like Dupoux (1993), we also advocate for the mandatory aspect of syllabic segmentation in the language under scrutiny. However we suggest that the nature of the fragment-detection task encourages pattern matching strategy, as it represents a simple and efficient heuristic alternative to more analytic, resource-consuming, segmentation processes.

Thus far our interpretation has rested upon the assumption that listeners rely upon a unique prelexical unit by default. However, whilst this has the benefit of simplicity it also requires the intervention of many additional processes related to attention, task demands, quality of the stimuli, and the listener’s language. It is the variability inherent in the combination of these processes

**Table 1 | Distribution and mean frequency of consonant clusters in Italian and French.**

	Italian		French	
	Count	Mean frequency	Count	Mean frequency
N-clusters	23,999	18.6	1709	29.4
F-clusters	18,837	20.6	18,490	33.5
P-clusters	28,412	18.3	37,450	26.5
L-clusters	23,830	15.4	19,035	24.2

Frequency is obtained by a count in lexique for French (New et al., 2004) and in COLFIS for Italian (Laudanna et al., 1995). Frequency counts are derived from written corpora (the only available in COLFIS). The count was made from 55,000 lemmas in French and 89,352 in Italian.

<sup>4</sup>There are some very rare exceptions, such as in French the NN cluster found at the start of “mnésique.”

that makes the identification of a unique prelexical building block so difficult as to lead us to abandon the concept.

An interesting alternative provided by Goldinger and Azuma (2003) is based upon Grossberg's ART model of speech perception (Grossberg, 2003), where multiple perceptual units in speech arise when bottom-up sensory information coalesce with top-down knowledge. Top-down knowledge refers to chunks of information in working memory, perhaps prototypes, corresponding to any combination of features including phonemes, syllables, and words. This "resonance" is achieved through a self-perpetuating feedback cycle between the bottom-up patterns and input-consistent chunks. This model will always privilege larger units, so that syllables will typically win over phonemes (as seen in Savin and Bever, 1970). Resonance can also be facilitated by directing participants' attention toward particular units (see also McNeill and Lindig, 1973, for a first formulation of this idea). Revisiting our results in the light of this interpretation leads to a rather different conclusion than that stated previously. Our finding that the syllable effect could only be observed when participants were slowed down (as compared to Tabossi et al., 2000) can be explained by the dynamic of activation of multiple chunks in the course of speech processing. As the signal unfolds over time, larger and larger chunks resonate with the input, allowing syllable chunks to win the race in the slower reactions. That the syllable effect could only be found in unstressed initial syllables can also be interpreted in a similar fashion, in that short, unaccented syllables could trigger less resonance amongst phoneme chunks than longer and better defined accented syllables. Finally, and most interestingly, the ART model is particularly well suited to an explanation of the dependence that the syllable effect has on the presence of catch trials. Within this framework particular speech units can be perceptually favored if bottom-up and/or top-down knowledge is manipulated through task demands. If experience with the task reveals that a specific stimulus

is anticipated, the corresponding top-down chunks will be pre-activated, resulting in accelerated resonance. In our experiments, this would mean that listeners in the condition without catch trials would pre-activate phonemic rather than syllabic chunks.

In summary, our findings show that the syllable effect can be observed in Italian in unstressed segments, with greater prominence in slow responses. However, when catch trials are removed from the procedure participants would appear to be encouraged to use a phonemic response strategy. This leads to a disappearance of syllable effects even when response latencies are equivalent to those found when catch trials are present. The impact of strategic demands on the observation of syllabic effects are possibly best accounted for by Goldinger and Azuma's (2003) proposal that perceptual speech units result from a resonance process between sensory input and multiple chunks of linguistic units, following Grossberg's (2003) ART model of speech perception. Under this perspective the identity of the prelexical unit varies constantly as a function of the information in the input and the level of activation in the chunks. Further research would allow an examination of how this proposal could accommodate for the language-specificity of results obtained in previous fragment-detection tasks, and from the diversity of results arising from other paradigms such as word spotting (Dumay et al., 2002) or the on-line cross-modal semantic fragment priming task (Tabossi et al., 2000; Tagliapietra et al., 2009).

## ACKNOWLEDGMENTS

This research was supported by a Human Frontier and Science Program project (ERBCHRXCT920031) and an ARC project n°96/01-203 entitled "The structure of the mental lexicon: A multilevel approach to the multiple representations of words". Many thanks to Luca Bonatti, Alessandra Sansavini, Giovanni Marvulli and Elena Morone.

## REFERENCES

- Aquil, R. (2011). The role of the syllable in the segmentation of Cairene spoken Arabic. *J. Psycholinguist. Res.* 41, 141–158.
- Bradley, D. C., Sanchez-Casas, R. M., and Garcia-Albea, J. E. (1993). The status of the syllable in the perception of Spanish and English. *Lang. Cogn. Process.* 8, 197–233.
- Brown, G. (1977). *Listening to Spoken English*. London: Longman.
- Content, A., Meunier, C., Kearns, R. K., and Frauenfelder, U. H. (2001). Sequence detection in pseudowords in French: where is the syllable effect? *Lang. Cogn. Process.* 16, 609–636.
- Cutler, A., Eisner, F., McQueen, J. M., and Norris, D. (2010). "How abstract phonemic categories are necessary for coping with speaker-related variation," in *Laboratory Phonology*, Vol. 10, eds C. Fougerson, B. Kühnert, M. D'Imperio, and N. Vallée (Berlin: de Gruyter), 91–111.
- Cutler, A., Mehler, J., Norris, D., and Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *J. Mem. Lang.* 25, 385–400.
- Cutler, A., and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 113–121.
- Dumay, N., Frauenfelder, U. H., and Content, A. (2002). The role of the syllable in lexical segmentation in French: word-spotting data. *Brain Lang.* 81, 144–161.
- Dupoux, E. (1993). "The time course of lexical processing: the syllabic hypothesis revisited," in *Cognitive Models of Speech Processing*, eds G. T. M. Altman and R. C. Shillcock (Cambridge MA: MIT Press), 81–114.
- Gaskell, G., and Marslen-Wilson, W. D. (1997). Integrating form and meaning: a distributed model of speech perception. *Lang. Cogn. Process.* 12, 613–656.
- Goldinger, S. D., and Azuma, T. (2003). Puzzle-solving science: the quixotic quest for units in speech perception. *J. Phon.* 31, 305–320.
- Goslin, J., and Floccia, C. (2007). Comparing French syllabification in preliterated children and adults. *Appl. Psycholinguist.* 28, 341–367.
- Grossberg, S. (2003). Resonant neural dynamics of speech perception. *J. Phon.* 31, 423–445.
- Jusczyk, P. W. (1986). "Towards a model for the development of speech perception," in *Invariance and Variability in Speech Processes*, eds J. S. Perkelland and D. H. Klatt Hillsdale (NJ: Lawrence Erlbaum Associates), 1–19.
- Kolinsky, R. (1998). Spoken word recognition: a stage-processing approach of language differences. *Eur. J. Cogn. Psychol.* 10, 1–40.
- Kolinsky, R., Morais, J., and Cluytens, M. (1995). Intermediate representations in spoken word recognition: evidence from word illusions. *J. Mem. Lang.* 34, 19–40.
- Laudanna, A., Thornton, A. M., Brown, G., Burani, C., and Marconi, L. (1995). "Un corpus dell'italiano scritto contemporaneo dalla parte del ricevente," in *III Giornate internazionali di Analisi Statistica dei Dati Testuali*, Vol. I, eds S. Bolasco, L. Lebart, and A. Salem (Roma: Cisu), 103–109.
- Mattys, S., and Melhorn, J. F. (2005). How do syllables contribute to perception of spoken English? Insight from the migration paradigm. *Lang. Speech* 48, 223–253.
- McClelland, J. L., and Elman, J. L. (1986). The TRACE model of speech perception. *Cogn. Psychol.* 18, 1–86.
- McCrary, K. (2002). "Syllable structure vs. segmental phonotactics: geminates and clusters in Italian revisited," in *Proceedings of the Texas Linguistics Society* (Austin: University of Texas).
- McNeill, D., and Lindig, K. (1973). The perceptual reality of phonemes, syllables, words, and sentences. *J. Verbal Learn. Verbal Behav.* 12, 419–430.
- McQueen, J. M., Cutler, A., and Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cogn. Sci.* 30, 1113–1126.

- Mehler, J., Dommergues, J. Y., Frauenfelder, U. H., and Segui, J. (1981). The syllable's role in speech segmentation. *J. Verbal Learn. Verbal Behav.* 20, 298–305.
- New, B., Pallier, C., Brysbaert, M., and Ferrand, L. (2004). Lexique 2: a new French lexical database. *Behav. Res. Methods Instrum. Comput.* 36, 516–524.
- Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition* 52, 189–234.
- Norris, D. G., and Cutler, A. (1988). The relative accessibility of phonemes and syllables. *Percept. Psychophys.* 43, 541–550.
- Obleser, J., and Eisner, F. (2009). Prelexical abstraction of speech in the auditory cortex. *Trends Cogn. Sci. (Regul. Ed.)* 13, 14–19.
- Pallier, C., Colomé, A., and Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: exemplar-based versus abstract lexical entries. *Psychol. Sci.* 12, 445–449.
- Pallier, C., Dupoux, E., and Jeannin, X. (1997). EXPE: an expandable programming language for on-line psychological experiments. *Behav. Res. Methods Instrum. Comput.* 29, 322–327.
- Pallier, C., Sebastián-Gallés, N., Feiguera, T., Christophe, A., and Mehler, J. (1993). Attentional allocation within the syllabic structure of spoken words. *J. Mem. Lang.* 32, 373–389.
- Savin, H. B., and Bever, T. G. (1970). The nonperceptual reality of the phoneme. *J. Verbal Learn. Verbal Behav.* 9, 295–302.
- Sebastián-Gallés, N., Dupoux, E., Segui, J., and Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *J. Mem. Lang.* 31, 18–32.
- Tabossi, P., Collina, S., Mazetti, M., and Zoppello, M. (2000). Syllables in the processing of spoken Italian. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 758–775.
- Tagliapietra, L., Fanari, R., Collina, S., and Tabossi, P. (2009). Syllabic effects in Italian lexical access. *J. Psycholinguist. Res.* 38, 511–526.
- Treiman, R., and Danis, C. (1988). Syllabification of intervocalic consonants. *J. Mem. Lang.* 27, 87–104.
- Vroomen, J., and de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 98–108.
- Wiltshire, C., and Maranzana, E. (1998). “Geminates and clusters in Italian and Piedmontese: a case for OT ranking,” in *Formal Perspectives in Romance Linguistics, LSRL XXVII*, eds M. Authier, B. Bullock, and L. Reed (Amsterdam: Benjamins), 289–303.
- Zwitzerlood, P., Schriefers, H., Lahiri, A., and van Donselaar, W. (1993). The role of the syllable in the perception of spoken Dutch. *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 260–271.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 December 2011; accepted: 20 April 2012; published online: 10 May 2012.

Citation: Floccia C, Goslin J, Morais JJD and Kolinsky R (2012) Syllable effects in a fragment-detection task in Italian listeners. *Front. Psychology* 3:140. doi: 10.3389/fpsyg.2012.00140

This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Floccia, Goslin, Morais and Kolinsky. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.

## APPENDIX

**Table A1 | List of test stimuli, with English translation in italics and raw frequency under brackets (from COLFIS, Laudanna et al., 1995). Numbers in italic correspond respectively to the duration of the CV portion and the CVC portion in each carrier.**

Words stressed on the first syllable			Word stressed on the second syllable		
CV	CVC	CVG	CV	CVC	CVG
Balia (14)	Baldo (2)	Ballo (99)	Baleno (6)	Balcone (38)	Balletto (39)
<i>Nursemaid 391–565</i>	<i>Proud 330–470</i>	<i>Ball 390–527</i>	<i>Flash 195–269</i>	<i>Balcony 200–291</i>	<i>Ballet 173–290</i>
Bara (29)	Barca (7)	Barra (7)	Barile (7)	Barcone (3)	Barrito (1)
<i>Coffin 391–503</i>	<i>Boat 373–470</i>	<i>Bar 392–488</i>	<i>Barrel 260–326</i>	<i>Pontoon 226–279</i>	<i>Trumpeting 260–363</i>
Calo (11)	Caldo (155)	Callo (3)	Calore (127)	Calvizie (2)	Calloso (1)
<i>Decrease 313–427</i>	<i>Hot 251–417</i>	<i>Corn 257–407</i>	<i>Heat 126–205</i>	<i>Thinning 121–220</i>	<i>Callous 140–298</i>
Caro (12)	Cargo (3)	Carro (38)	Carena (1)	Cartello (51)	Carrello (13)
<i>Dear 379–434</i>	<i>Freight 289–362</i>	<i>Cart 277–394</i>	<i>Keel 193–266</i>	<i>Sign 131–198</i>	<i>Trolley 200–284</i>
Male (366)	Malto (2)	Mallo (1)	Malore (10)	Malvagio (6)	Malloppo (4)
<i>Pain 414–511</i>	<i>Malt 329–448</i>	<i>Husk 317–438</i>	<i>Illness 236–298</i>	<i>Bad 203–274</i>	<i>Loot 217–319</i>
Mira (27)	Mirto (1)	Mirra (6)	Marino (44)	Martello (21)	Marrone (13)
<i>Aim 333–448</i>	<i>Myrtle 353–448</i>	<i>Myrrh 386–464</i>	<i>Marine 235–289</i>	<i>Hammer 200–279</i>	<i>Brown 225–312</i>