Stem and fragment priming on verbal forms of Italian

Alessandro Laudanna
University of Salerno
Via Giovanni Paolo II, 132 -
84084 - Fisciano (SA)
alaudanna@unisa.it

Giulia Bracco
University of Salerno
Via Giovanni Paolo II, 132 -
84084 - Fisciano (SA)
gcbracco@unisa.it

Abstract

English. In this paper we investigate the respective roles of orthographic and morphological structure in the processing of Italian verbal forms by using the masked priming paradigm.
According to the morphology-based view, in a priming condition the recognition of an inflected word should be facilitated by the presentation of the stem. The cohort model, instead, postulates that the orthographic material from the word's onset to the uniqueness point should be sufficient for the activation of the morphological family.

Italiano. *In questo lavoro indaghiamo il ruolo della struttura ortografica e della struttura morfologica nella elaborazione delle forme verbali dell'italiano, usando il paradigma del priming mascherato. Secondo i modelli basati sulla morfologia, in una condizione di priming il riconoscimento di una forma flessa dovrebbe essere facilitato dalla preventiva presentazione della radice. I modelli coorte, invece, propongono che il materiale ortografico dall'inizio della parola fino al punto di unicità sia sufficiente per attivare la famiglia morfologica.*

1 Introduction

According to the cohort models of visual word recognition (Johnson and Pugh, 1994), all the sources of information that contribute to the identification of a target word proceed from left to right: a single word of the cohort becomes unique at the uniqueness point, when it remains the only candidate corresponding to the orthographic configuration of the stimulus. At that point the recognition takes place. Cohort models use the neighborhood size and the frequency distribution of neighbors as predictors of the competition between candidates and of the consequent recognition latencies. The same models not always deal with the internal morphological structure of the word (but see Marslen-Wilson, 1987). In Italian, verbal families have orthographically similar members, but these words often become 'unique' only at the end, since information about mood, tense and number is carried by affixes in the final part of the word. The N-count is not the best measure to describe their relatedness. On the other hand, morphological parsing accounts (e.g., Baayen, Dijkstra, and Schreuder, 1997; Burani and Caramazza, 1987; Colé, Beauvillain, and Segui, 1989; Taft, 1979) support the argument that the stem is the key of access to the lexicon. Morphological priming shows that, even though morphologically related pairs share some orthographic material (verbs sharing the stem also share letters in initial position, unless they are prefixed), the role played by the morphological structure is different from the one played by the orthographic structure (Pastizzo and Feldman, 2002). In the following experiments, we used the priming paradigm to preactivate both the morphological and the orthographic keys of access. Many studies employing the priming paradigm used an orthographically similar baseline, while others (e.g., Feldman and Soltano, 1999; Marslen-Wilson, Tyler, Waksler, and Older, 1994) employed orthographically and morphologically dissimilar baselines, and others, finally, (Giraudo and Grainger, 2000; Grainger, Colé, and Segui, 1991) estimated morphological facilitation in comparison with both types of baselines. To our knowledge, no study employed the stem priming in Italian, by using both an orthographic and a morphological baseline. We use the term 'stem' to denote the residual part of the word when all inflectional affixes are removed and that can, or cannot be, complex. In contrast, the root is not analysable (Bauer, 1988). The stem does not end in a vowel, and if presented in isolation, it appears as an incomplete word, an orthographic fragment, even if it carries lexical information. By this token, this feature could also turn out to be a benefit: studies employing the priming paradigm usually have to deal with the objection that the base form is a word. The stem primes provide an abstract lexical information in a non-lexicalized form. In order to ascertain if the initial string of letters up to the uniqueness point is the orthographic access code, in Experiment 1 we used Italian verbs with the uniqueness point occurring before the stem (e.g., 'ABBAN' is a fragment present only in the morphological family of 'abbandonare', *to abandon.* All the words starting with the fragment 'ABBAN' (e.g. 'abbandonai', *I abandoned*, 'abbandonato', *abandoned*) belong to the same morphological family). Then, in Experiment 2, we started from the consideration that, when the initial fragment before the stem is shared by more than one family, only a non homographic stem is the true 'uniqueness point': we used verbal forms with the uniqueness point only at the stem boundary (e.g., the fragment 'DISTRI' is shared by two morphological families: 'distribuire', *to distribute*, and 'districare', *to unravel*, whose stems, 'DISTRIB' and 'DISTRIC', respectively, have no homograph). The masked priming paradigm was used in order to avoid intuitions or response strategies in participants (Forster and Davis, 1984; Forster, Davis, Schoknecht and Carter, 1987): this technique avoids the overt detection of any relation between prime and target. Moreover, it has been argued that lexical decision latencies associated with masked priming also reflect the organization of the lexicon in the mind, rather than representing the mechanisms directly involved during single words access (Baayen, 2014). In Experiment 2 cohorts defined by the fragment and by the stem had different frequency distribution, with fragments matching the initial part of lower or higher frequency morphological families. By this token, on the one hand we expected to detect the morphological vs. orthographic nature of the key(s) of access to the lexicon; on the other hand, we indirectly tested the stem frequency effect.

2 Experiment 1

Italian verbs such as 'abbandonare' (*to abandon*) or 'scivolare' (*to slip*) contain fragments ('ABBAN' and 'SCIVO'), that, although shorter than the respective stems 'abbandon*' and '*scivol' , can be considered 'morphological uniqueness points', because they belong to just one morphological family. Those stimuli are relevant to decide whether the stem is the necessary key of access to lexical information, or the fragment is sufficient to contact the lexicon. According to the morphology-based view, in a priming condition the recognition of an inflected word should be strongly facilitated by the presentation of the stem. According to the cohort model, instead, the orthographic material from the word's onset to the uniqueness point should be sufficient for the activation of the morphological family.

2.1 Method

Stimuli We selected 16 inflected forms of verbs with a 'unique' initial fragment (e.g., 'abbandonare', to abandon), which served as targets in three different experimental conditions: A) primed by the stem, (e.g., 'ABBANDON'); B) primed by the initial fragment up to the uniqueness point (e.g., 'ABBAN'); C) preceded by an orthographically unrelated fragment which shared no letter with the prime (e.g., 'COTRU'). Mean values for length were 6.9 letters for Stems and 5.3 for Unrelated Prime and Fragments; prime-target orthographic overlap was 67% in Stem Condition and 55% in Fragment Condition. Target mean frequency was 13; root frequency was 462 and initial stem cohort frequency was 245. Three hundred eighty-four items were included in the list as fillers. One hundred eighty-four were words, (40 adjectives, 106 nouns, 38 inflected verbal forms). Those words, together with those in the experimental list, displayed a distribution similar to the one of written Italian (see CoLFIS, Bertinetto et al., 2005). The filler words were matched with experimental targets for their mean length in letters and for their surface frequency. The list included two-hundred items as pseudoword targets. The whole list was composed of 200 words and 200 pseudoword targets preceded in turn by 100 existing primes and 100 non existing primes.

Participants Fifty-four participants, all students of the University of Salerno, and native speakers of Italian, took part into the experiment. They served for a session lasting about 40 minutes. The whole experiment was arranged in three different sessions and each session contained all the targets in one of the three experimental condition (either preceded by the fragment, or preceded by the stem, or preceded by the unrelated fragment). Each participant was submitted to a single experimental session, for a total of 18 'superparticipants'. Each superparticipant was composed of 3 participants, and constituted one data point in the statistical analyses.

Equipment Response box, connected to an IBM PC running the E-Prime 1.1 software (Version 1.1).

Procedure Participants had to press the button corresponding to their dominant hand for the decision 'word', and another one for the decision 'non word'. When the participants reached the 70 % of correct responses in a practice session, the experiment started. All the stimuli appeared in Courier New font, 18 point size in the centre of the computer screen. The fixation was 51 ms, followed by a 51 ms pause. Primes appeared for 51 ms, followed by a 12 characters backward mask ############ (150 ms). The targets remained on the computer screen for a maximum of 1 second. If the participants did not produce any answer within 1 second, the feedback 'Fuori tempo' (*Out of time*) appeared on the screen. The reaction times (RT) were measured from target's onset to subject's response, and the lack of a response was scored as an error.

2.2 Results and Discussion

In Table 1 the mean reaction times and percentage of errors are shown. Table 2 shows the size of Stem and Fragment Priming effects in response latencies and percentage of errors. For 'size of priming effect' we mean the difference between mean Reaction Times (or number of errors) in Stem Condition (or in Fragment Condition) and mean Reaction Times (or number of errors) in Control Condition (Unrelated Fragment Condition).

Condition	Stem	Unrelated Fragment	Fragment
Reaction Times	626 ms	650 ms	626 ms
Errors	12%	15%	13%

Table 1: Mean correct lexical decision latencies and percentage of errors in each priming condition.

Stem Priming Efffect	- 24 ms (-3%)
Fragment Priming Effect	- 24 ms (-2%)

Table 2: Priming effects in response latencies. In parentheses the effect in percentage of errors.

As shown in Table 1, the conditions 'Fragment' and 'Stem' were faster than 'Unrelated Fragment' (Control) condition and they did not differ from each other despite stems were on average longer and with more letters in common with the target word than fragments. The ANOVA on error data did not reveal any significant result (ANOVA by participants $F_{(2,34)}$=.50; p>.6; ANOVA by item $F_{(2,30)}$=.66; p>.5).

The ANOVA on response latencies showed a main effect of prime type only in the analysis by participants ($F_{(2,34)}$= 7.01; p<.002; ANOVA by item $F_{(2,30)}$=1.80; p>.2). Post-hoc analyses based on the ANOVA by participants showed a significant difference between the conditions 'Fragment' vs. 'Unrelated Fragment' (p<.002) and between the conditions 'Stem' vs. 'Unrelated Fragment' (p<.002), but not between 'Fragment' vs. 'Stem' (p >.9). The results are inconsistent with predictions of morphologically-based view: the orthographic uniqueness point is sufficient to contact lexical information.

3 Experiment 2

In Experiment 2 we selected fragments (e.g., DISTRI) shared by verbal families with different frequencies (e.g. 'districare', *to unravel*, lower frequency, LF, and 'distribuire', *to distribute,* higher frequency, HF). Both the accounts (orthographic and morphological) suggest that the fragment is not sufficient for the activation of the morphological family, while the stem, which is also the uniqueness point, should determine a stronger facilitation. The aim of the Experiment 2 was also to address the stem frequency effect.

3.1 Method

Stimuli We selected 32 inflected forms of verbs in 16 pairs with the same initial fragment (e.g., 'distribuito', and 'districato'). One half of the list was composed of 16 targets belonging to higher frequency morphological families (HF, e.g. distribuire, *distributed*); the other half was composed of 16 targets belonging to lower frequency morphological families (LF, e.g., districato, *unraveled*). Target frequency was 2 for LF words, 15 for HF words, 50 for LF Stems and 216 for HF Stems; initial stem cohort frequency was 216 for HF words and 50 for LF words, root frequency was 676 for HF words and 60 for LF words; prime- target orthographic overlap was 72% in Stem condition and 53% in Fragment condition. The same three different experimental conditions of Experiment 2 were arranged. Three hundred sixty-eight items were included in the list as fillers. One-hundred sixty-eight were words, two-hundred items as pseudoword targets. The whole list was composed of 200 words and 200 pseudowords targets preceded in turn by 100 existing primes and 100 non existing primes.

Participants Fifty-four participants, all students of the University of Salerno, and native speakers of Italian, took part into the Experiment. Each participant was submitted to a single session (like in Experiment 1), for a total of 18 superparticipants. Each superparticipant was composed of 3 participants.

Equipment and procedure They were the same as in Experiment 1.

3.2 Results and Discussion

In Tables 3 and 4 the mean reaction times and percentage of errors are shown. Table 5 shows the size of Stem and Fragment Priming effects in response latencies and percentage of errors.

LF			
Condition	Stem	Unrelated Fragment	Fragment
Reaction Times Errors	652 ms 22%	644 ms 24%	647 ms 28%

Table 3: LF verbal forms: mean correct lexical decision latencies and percentage of errors in each priming condition.

HF			
Condition	Stem	Unrelated Fragment	Fragment
Reaction Times Errors	615 ms 14%	627 ms 9%	621 ms 18%

Table 4: HF verbal forms: mean correct lexical decision latencies and percentage of errors in each priming condition.

	LF	HF
Stem Priming Effect	+ 8 ms (-2%)	- 12 ms (-2%)
Fragment Priming Effect	+ 3 ms (+4%)	- 6 ms (+ 3%)

Table 5: Priming effects in response latencies. In parentheses the effect in percentage of errors.

The ANOVA on error data showed an effect of frequency in analyses on both participants ($F_{(1,16)}=22.33$; $p<.0005$) and items ($F_{(1,30)}=3,91$; $p<.05$): LF frequency words elicited higher error rates; we also found an effect of prime type in analyses on both participants ($F_{(2,32)}=5.00$; $p<.01$) and items ($F_{(2,60)}=4.42$; $p<.01$), but no interaction (ANOVA by participants $F_{(2,32)}=1.23$; $p>.3$; ANOVA by item $F_{(2,60)}=1.29$; $p>.2$). The ANOVA on RT showed a main effect of frequency in analyses on both participants ($F_{(1,17)}=25.80$; $p<.0001$) and items ($F_{(1,30)}=4.41$; $p<.04$), no effect of prime type (ANOVA by participants: $F_{(2,34)}=.07$; $p>.9$, ANOVA by item: $F_{(2,60)}=.18$; $p>.8$), and no interaction (ANOVA by participants $F_{(2,34)}=1.07$; $p>.3$; ANOVA by item $F_{(2,60)}=.15$; $p>.8$). On average, HF targets were recognized better than LF targets (621 ms Vs. 647 ms), with faster latencies and a lower percentage of errors (13% Vs. 24%). The lack of priming effect for the Stem condition as compared with the Unrelated Fragment condition is the most surprising result. Post-hoc correlations were performed using main lexical and orthographic variables as predictors, and size of stem and fragment priming effects for RT and errors as criteria. The correlations on results in Fragment condition showed a significant length effect for the fragment prime on HF words ($r=-58$, $p<.02$). More interestingly, correlations in Stem conditions showed that the ratio between the surface frequency and the frequency of the stem in initial position is inversely correlated with the size of stem priming ($r= -.36$, $p<.04$). The higher the ratio, the faster the latencies: the “relative frequency” of the form in its cohort determines the direction of the effect.The correlation was reliable on LF words ($r=-.60$, $p<.01$), while it was not significant on HF words ($r=.31$ $p>.2$). The effect did not occur in the Fragment condition, and this might suggest that the effect occurs at the point where the morphological family is selected: the more frequent the cohort, the stronger the inhibition for a verbal form that has a low surface frequency. No effect of cumulative root frequency occurred when the frequency count was obtained by including words embedding the stem in any position (for instance prefixed words), and this allows us to assume that the effect is orthographic in nature. We conclude that not only the word surface frequency, but also the “relative frequency” of the word with respect to its cohort is responsible for recognition.

4 General Discussion

Results of Experiment 1 show that when orthographic information about initial part of the word is exhaustive, it is as reliable as stem priming, and these results are difficult to reconcile with the morphologically- based view which postulates that the stem is critical for lexical access. In addition, the ‘relative frequency’ effect (Experiment

2), which arises in Stem Condition, suggests that, during recognition, when a low frequency word shares the stem with higher frequency members, it is disadvantaged. In order to gain lexical access, the word has to sustain a harder competition with other members according to their frequency distribution in the morphological family. This effect has been largely described for orthographic neighborhood (Grainger, O'Regan, Jacobs, and Segui, 1989), but, again, it is difficult to reconcile with the stem frequency effect (Burani, Salmaso, and Caramazza, 1984) which states the opposite, with cumulative frequency of morphologically related words facilitating the recognition of a low frequency word. These results are in line with previous data on Italian, which failed in replicating root frequency effects (Laudanna and Bracco, 2009). Root morpheme frequency effects are crucial in the general issue of whether words are accessed through decomposition rather than as full forms, since it has been widely used as the strongest evidence in favor of the hypothesis of the root morpheme representation.

Nevertheless, the morphological parsing account is not the unique explanation for root frequency effects: also full listing models (see Giraudo and Grainger, 2001) can provide a general outline in which this effect is explained, for instance, in terms of lexical connections. This view has been inspected also in linguistics: Bybee (1995) proposed that the activation level of a word is the result of lexical connections and lexical strength, with the first one determined by the frequency of the lexical item, and lexical connections corresponding to the pattern of weight connections associated with the links between related items. For high frequency words the individual lexical strength is elevated; low frequency words have a weak lexical activation and need the support of the activation of lexical connections. This theory is consistent with the claim that whole word frequency effects arise over a precise threshold (Alegre and Gordon, 1999). If the measure for these connections is the stem frequency, the more frequent the stem, the faster the recognition (Burani et al., 1984, Traficante and Burani, 2003, Colombo and Burani, 2002). Meunier and Segui (1999) found that the relative frequency of family members affects the recognition of auditory stimuli: words with high-frequency suffixed candidates derived from the same stem were recognized more slowly than words with morphological family members of a lower frequency. The effects discussed in this paper are also consistent with previous data on stem priming and the "relative frequency" effect in Italian (Bracco and Laudanna, 2012), suggesting that the relations between words in the paradigm need to be taken into account, even if we maintain that whole word representations are the keys for lexical access.

In summary, the results presented in this paper about the processing of Italian verbal forms suggest that it is performed sequentially and it proceeds from left-to-right. Morphological structure does not play a deterministic role, and recognition is guided by the information carried by the initial part of the word, whether it matches a morpheme or not.

Priming induced by 'unique' fragments is as reliable as stem priming. In addition, stem priming is not explainable in terms of a stem frequency effect.

Furthermore, the observation of a 'relative frequency' rather than a 'stem frequency' effect, suggests that we are tapping into a phenomenon concerning connections among whole words.

References

Alegre, M., and Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory & Language*, 40, 41-61.

Baayen, R.H. (2014) Experimental and psycholinguistic approaches to studying derivation. In: R. Lieber and P. Stekauer (Eds), *Handbook of derivational morphology*. Oxford: Oxford University Press, 95-117.

Baayen, H., Dijkstra, T., and Schreuder, R. (1997). *Singulars and plurals in Dutch: Evidence for a parallel dual-route model*. Journal of Memory and Language, 37, 94-117.

Bauer, L. (1988). *Introducing Linguistic Morphology*. Edinburgh, Edinburgh University Press.

Bertinetto, P.M., Burani C., Laudanna, A., Marconi L., Ratti n , D., Rolando, C. and Thornton A.M. (2005). *Corpus e Lessico di frequenza dell'Italiano Scritto* (CoLFIS) http://alphalinguistica.sns.it/CoLFIS/CoLFIS_Presentazione.htm

Bracco G., and Laudanna, A. (2012). "Meccanismi di competizione tra forme verbali nell'accesso al lessico mentale". In: Pier Marco BERTINETTO, Valentina BAMBINI and Irene RICCI e Collaboratori (a cura di). *Linguaggio e cervello / Semantica, Atti del XLII Convegno della Società di Linguistica Italiana* (Pisa, Scuola Normale Superiore, 25-27 settembre 2008). Roma: Bulzoni. Volume 2.

Burani, C., and Caramazza, A. (1987). *Representation and processing of derived words*. Language and Cognitive Processes, 3, 217-227.

Burani, C., Salmaso, D., and Caramazza, A. (1984). *Morphological structure and lexical access*. Visible Language, 18, 348-358.

Bybee, J. L. (1995). *Regular morphology and the lexicon*. Language and Cognitive Processes, 10, 425-455.

Colé, P., Beauvillain, C. and Segui, J. (1989). *On the representation and processing of prefixed and suffixed derived words: A differential frequency effect*. Journal of Memory and Language, 28, 1-13.

Colombo, L., and Burani, C. (2002). *The influence of age of acquisition, root frequency and context availability in processing nouns and verbs*. Brain and Language, 81, 398-411.

Feldman, L.B., and Soltano, E.G. (1999). *Morphological priming: The role of prime duration, semantic transparency and affix position*. Brain and Language, 60, 33-39.

Forster, K. I., and Davis, C. (1984). *Repetition priming and frequency attenuation in lexical access*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 680–698.

Forster, K. I., Davis, C., Schoknecht, C., and Carter, R. (1987). *Masked priming with graphemically related forms: Repetition or partial activation?*Quarterly Journal of Experimental Psychology, 39, 211–251.

Giraudo, H. and Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. Psychonomic Bulletin & Review, 8, 127-131.

Giraudo, H., and Grainger, J.(2000). *Effects of prime word frequency and cumulative root frequency in masked morphological priming*. Language and Cognitive Processes 15, 421-444.

Grainger, J., Colé, P., and Segui, J. (1991). *Masked morphological priming in visual word recognition*. Journal of Memory and Language, 30, 370-384.

Grainger, J., O'Regan,K., Jacobs,A., and Segui,J. (1989). *On the role of competing word units in visual word recognition: The neighborhood frequency effect*. Perception and Psychophysics, 45, 189-195.

Johnson, N. F. and Pugh, K. R. (1994). *A cohort model of visual word recognition.* Cognitive Psychology, 26, 240-346.

Laudanna A., and Bracco, G. (2009). “Family Frequency, Stem Frequency and Family Size in Visual Word Recognition of Italian Verbal Forms”. In: Proceedings of the 6th International Morphological Processing Conference. Turku, June 14-17, 2009, Turku: Morproc, p. 63-64.

Marlsen-Wilson, W.D., Tyler, L., Waksler, R., and Older, L. (1994). *Morphology and meaning in the English mental lexicon*. Psychological Review, 101, 3-33.

Marslen-Wilson, W.D. (l987). *Functional parallelism in spoken word-recognition*. In U. Frauenfelder and L.K. Tyler, Spoken word recognition. Cambridge. MA, MIT Press, 71-102.

Meunier, F. and Segui, J. (1999) Morphological priming effect: the role of surface frequency. Brain and language, 68, 54-60.

Pastizzo, M. J. and Feldman, L. B. (2002) *Discrepancies between orthographic and unrelated baselines in masked priming undermine a decompositional account of morphological facilitation*. Journal of Experimental Psychology: Learning, Memory and Cognition 28, 244–249.

Taft, M. (1979). *Recognition of affixed words and the word frequency effect*. Memory and Cognition, 7, 263-272.

Traficante, D., and Burani, C. (2003). *Visual processing of Italian verbs and adjectives: The role of the inflectional family size*. In R. H. Baayen, R. Schreuder, Morphological Structure in Language Processing, Berlin, Mouton de Gruyter, 45-64.