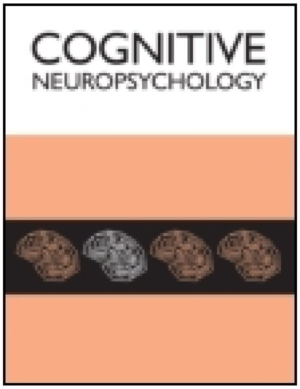
This article was downloaded by: [Valentina Bambini]
On: 02 May 2014, At: 06:40
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Cognitive Neuropsychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pcgn20

Word structure and decomposition effects in reading

Giorgio Arcara[a], Carlo Semenza[ab] & Valentina Bambini[c]
[a] IRCCS Fondazione Ospedale San Camillo, Venice, Italy
[b] Department of Neuroscience, University of Padua, Padua, Italy
[c] Center for Neurocognition and Theoretical Syntax, NeTS, Institute for Advanced Study, IUSS, Pavia, Italy
Published online: 30 Apr 2014.

To cite this article: Giorgio Arcara, Carlo Semenza & Valentina Bambini (2014) Word structure and decomposition effects in reading, Cognitive Neuropsychology, 31:1-2, 184-218, DOI: 10.1080/02643294.2014.903915

To link to this article: http://dx.doi.org/10.1080/02643294.2014.903915

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Cognitive Neuropsychology, 2014
Vol. 31, Nos. 1–2, 184–218, http://dx.doi.org/10.1080/02643294.2014.903915

Routledge
Taylor & Francis Group

Word structure and decomposition effects in reading

Giorgio Arcara[1], Carlo Semenza[1,2], and Valentina Bambini[3]

[1]IRCCS Fondazione Ospedale San Camillo, Venice, Italy
[2]Department of Neuroscience, University of Padua, Padua, Italy
[3]Center for Neurocognition and Theoretical Syntax, NeTS, Institute for Advanced Study, IUSS, Pavia, Italy

Theories on the processing of compound words differ on the role attributed to access to individual constituents. These theories are mostly based on empirical evidence obtained in experimental settings that could induce artificial effects normally not occurring in natural processing. In this study we investigated the processing of compounds as compared to noncompound complex words in Italian through a reading task with eye movement recording. We included both head-initial and head-final compounds, in order to test whether the position of the head may influence the reading process. After ruling out the effects of length and frequency, we observed that pseudocompounds (i.e., words with a segment homograph to a real word in the leftmost part) elicited longer total reading times than all other types of complex words, including compounds. Furthermore, head-final compounds elicited longer total reading times than head-initial compounds. The results suggest that a word structure resembling a compound may induce longer processing, presumably related to unexpected morphological structures. The results also converge with previous evidence that in some cases there is a higher processing costs for head-final as opposed to head-initial compounds, possibly indexing a reanalysis of the stimulus in order to correctly assign the constituent properties. However, a deeper analysis restricted to compounds revealed a more complex scenario where several variables interact with headedness (namely, first and second constituent frequency, compound frequency, and compound length), and future studies are needed to discriminate among possible interpretations. Overall, our findings suggest that longer reading times are related to solving incongruities due to noncanonical structures, rather than to morphologically complexity per se.

Keywords: Morphology; Morphological decomposition; Compound words; Compound headedness; Eye movements.

One of the main issues in studies on morphological processing is whether complex words (i.e., words formed by two or more morphemes, e.g., “cleaner”) are represented and processed differently from simple words (i.e., words formed by one morpheme, e.g., “chain”). Compounds—that is,

---

Correspondence should be addressed to Giorgio Arcara, I.R.C.C.S Fondazione Ospedale San, Camillo, Venice, Italy, Via Alberoni, 70, 30126, Lido-Venice, Italy (E-mail: giorgio.arcara@gmail.com).

Preliminary results of this study were presented at the 51st Annual Meeting of the Academy of Aphasia (Lucerne, 2013). The study was ideated and elaborated jointly by the three authors. Data preparation and collection were conducted by the first and third authors. Data analysis was conducted by the first author. We thank Maria Concetta Morrone (CNR, Pisa) and Pier Marco Bertinetto (Laboratorio di Linguistica, SNS, Pisa) for technical facilities, Chiara Bertini for technical support, and Petar Milin for the very useful comments. We also thank an anonymous reviewer for the help in the interpretation of statistical analyses.

This study has been supported by the University of Padua [grant number STPD11B8HM_004 to Carlo Semenza].

words formed by two or more existing words—represent a special case of morphologically complex words, since their constituent parts can generally occur independently in the language. For example, the constituents in the compound "chainsaw" ("chain" and "saw") may occur independently, whereas the derivational suffix of "cleaner" (i.e., "er") may occur only as a part of a word. Given their special characteristics, compounds have been largely studied, as they offer a unique opportunity to understand the structure of the mental lexicon (Libben, 2006).

Historically, two competing theories have been proposed to account for the representation and processing of compound words. According to the full listing theories (Butterworth, 1983; Bybee, 1995), compounds are listed in the lexicon and processed as simple words, while according to parsing theories (Taft & Forster, 1976), compounds are always decomposed in their processing. A listing approach alone soon proved to be very limited, because it cannot account for common phenomena such as the production and comprehension of novel compounds. Thus, increasing interest was devoted to parsing theories, which were supported also by the first empirical evidence on compound processing. In their seminal study, Taft and Forster (1976) observed, in a lexical decision task, that nonwords where the first constituent is a word (e.g., "*dustworth") were classified more slowly as nonwords than as nonwords whose first constituent is not a word (e.g., "*trowbreak"). By contrast, the presence of an embedded word in the rightmost part of the nonwords did not affect the reaction times. Taft and Forster (1976) interpreted their results as supporting a full-parsing account of compound processing, with a prominent role of the first syllable in lexical access. However, the role of the first constituent in compound processing was not confirmed in later studies, which suggested a crucial role of the second constituent in compound access as well (Andrews, 1986; Andrews, Miller, & Rayner, 2004; Juhasz, Starr, Inhoff, & Placke, 2003; Kehayia et al., 1999; Libben, Gibson, Yoon, & Sandra, 2003).

A third family of models for lexical access is represented by so-called dual-route models (Schreuder & Baayen, 1995). According to these models, both constituent and whole compound representations play a role in compound processing (Pollatsek, Hyönä, & Bertram, 2000). The preference of one route over another (at some point of lexical processing) would be related to several variables. For example, high-frequency (e.g., "birthday") or semantically opaque (e.g., "hogwash") compounds are likely to be accessed via whole-word representation, whereas low-frequency ("trapdoor") and semantically transparent ("carwash") compounds are likely to be accessed via their constituents. Some more recent accounts of compound processing suggested that several sources of information (and not only constituent and whole-word representations) play a role, among which the family size of a constituent, and the conditional probabilities of encountering a constituent in a given position (Kuperman, Bertram, & Baayen, 2008; Kuperman, Schreuder, Bertram, & Baayen, 2009).

From the overview of the literature, one aspect that emerges clearly is that the psycholinguistic approach has focused on understanding what properties of a compound may influence their decomposition, or privilege the access to constituent or to whole-word representations, rather than comparing whether compound and noncompound words are processed differently. Yet this aspect appears as fundamental in order to understand whether a specific effect of compound words exists. A positive answer to this question comes from neuropsychological investigations, which have focused on unravelling different effects of compounds as compared to noncompounds (Semenza & Mondini, 2006; see Semenza & Mondini, 2010, for a review). Aphasic patients may show relatively spared compound processing or noncompound processing (Mondini, Arcara, & Semenza, 2012), suggesting that compound and noncompound are processed differently. Moreover, in naming tasks, aphasic patients tend to substitute a compound target word with another compound (the so-called "compound effect", Semenza & Mondini, 2010), indicating that the knowledge of the morphological status of the words is stored separately from its phonological form.

In this scenario, it becomes of primary importance to search for the specificity of compounds also from the psycholinguistic point of view. In order to address the issue of compound processing, it is useful to adopt a more general approach and test not only compounds as compared to other compounds with different properties, but also compounds as opposed to noncompounds, and especially to other words that are morphologically complex yet not compounds. In the general context of processing morphologically complex words, a relevant question is whether there are features in a word that make it prone to decomposition into its morphemes. The prevailing view comes from studies that focused not on compound words, but rather on derived words (Longtin & Meunier, 2005; Rastle & Davis, 2004, 2008). According to this view, it is the apparently complex morphological structure that elicits decomposition. Thus, words such as "cleaner" (morphologically complex "clean" + "er") and "corner" (morphologically simple, but resembling a morphologically complex word made up by "corn" + "er") are similarly decomposed in their processing. Importantly, words that do not end with a segment homograph to a morphological suffix (e.g., "brothel", where "broth" is homograph to a real morpheme, but "-el" is not a morphological suffix) do not show evidence of decomposition (see Amenta & Crepaldi, 2012, for a recent review on early effects in morphological processing). According to this "form-then-meaning" account (Rastle & Davis, 2008), the semantic properties of a word play a role only in the later stages of processing, after decomposition based on morpho-orthographic characteristics has been carried out. Compatible findings were also found for compounds by Fiorentino and Fund-Reznicek (2009). They replicated a pattern of results supporting the form-then-meaning account in a masked priming study with three kinds of stimulus pairs: Both semantically transparent opaque compounds primed their constituents, whereas primes with a simple orthographic overlap with the target did not elicit a priming effect.

The form-then-meaning view of word recognition has been challenged by some results showing an early role of semantics (Diependaele, Sandra, & Grainger, 2005; Feldman, O'Connor, & Moscoso del Prado Martín, 2009, but see Davis & Rastle, 2010, for a reply). The debate and the alternative models on morphological decomposition stem mostly from results of masked priming experiments in the context of a lexical decision task; it is possible that different tasks produce different results, however. For example, a recent study employing masked priming but in a different task requiring access of the meaning of the target word and using eye movement measures as dependent variable (Marelli, Amenta, Morone, & Crepaldi, 2013) found no evidence of decomposition of pseudo-morphologically complex words (e.g., priming of "corner" on "corn"). On the contrary, early effects of morphosemantics were found (e.g., priming of "cleaner" on "clean"). Notably, in this study the typical pattern of results supporting the form-then-meaning account was replicated with the same set of stimuli when the task was lexical decision. These findings suggest that some morphological effects may be strongly task dependent and underline the importance of gathering data on morphological decomposition also with alternative experimental paradigms to better understand the dynamics of morphological processing.

The issue of headedness in compound processing

One of the key issues in the study of compound processing revolves around the role of the head constituent—that is, the constituent that carries most of the semantic information and that determines the grammatical properties of the whole compound (for a discussion on the head parameters, see Moro, 2000, and Di Sciullo, 2005). Several studies suggest that the head position can influence the processing of compound words (Arcara, Marelli, Buodo, & Mondini, 2013; El Yagoubi et al., 2008; Jarema, Busson, Nikolova, Tsapkini, & Libben, 1999; Jarema, Perlak, & Semenza, 2010; Marelli, Crepaldi, & Luzzatti, 2009; Marelli & Luzzatti, 2012; Semenza et al., 2011).

Position of the head varies across languages: For instance, typical English compounds are head-final (right-headed), while Italian compounds may be both head-initial (left-headed), as in "capobanda" (translation, tr., "band leader") and ufficio reclami (tr. "complain office"; Graffi & Scalise, 2002; Guevara & Pirrelli, 2012; Scalise, 1994), and head-final, as in "astronave" (tr. "spaceship"), "terremoto" (tr. "earthquake"), "scuolabus" (tr. "schoolbus"). These two different types of compounds are presumably originated by two different morphological mechanisms (Radimský, 2013; Scalise, 1994; Schwarze, 2005). Traditionally, the head-initial structure has been considered as dominant in Italian (Scalise, 1994), but recent corpus analyses challenged this claim, suggesting that indeed head-final compounds are numerically dominant, even in Romance languages (Guevara & Scalise, 2009; Marelli & Luzzatti, 2012; Schwarze, 2005). Although head-final structures are undoubtedly more frequent, in theoretical linguistics there is still a debate on whether this implies that the head-final configuration is the most productive mechanism (Delfitto & Melloni, 2012). For example, in their analysis of different types of compounds, Scalise and Fàbregas (2010) concluded that there are more productive mechanisms associated with left-headedness than with right-headedness.

In languages where the head position is fixed, it is not possible to study the effect of headedness separately from the effect of position. For example, in English compounds (almost exclusively head-final), any effect of the head constituent would be also an effect of the final constituent. Thus, Italian provides a good opportunity to study headedness effects in compound processing, since it allows disentangling the role of headedness from that of position.

Results on Italian compounds suggest that the head position may indeed influence processing, and a quite consistent finding is that head-initial and head-final are differently processed. For instance, in a lexical decision task on Italian words, El Yagoubi et al. (2008) compared head-initial and head-final compounds to noncompound words that contained an embedded segment either at the beginning or at the end of the word (e.g., pseudocompounds such as "coccodrillo", tr. "crocodile", where "cocco", tr. "coconut", is homograph to a real word neither morphologically nor semantically related to the whole word). The results showed a difference between head-initial and head-final compounds, with an enhanced P300 effect in right-headed compounds. El Yagoubi et al. (2008) also found an enhanced N400 modulation for pseudocompounds as compared to compounds, suggesting a more effortful semantic integration and lexical access with stimuli including an embedded word-like segment.

Two competing interpretations have been proposed to account for the results documenting differences in processing related to headedness. According to El Yagoubi et al. (2008), the head-initial compounds are the default condition in reading, since the order of constituents reflects the canonical order of head-modifier in Italian syntax (see also results by Arcara et al., 2013). In contrast, according to Marelli and Luzzatti (2012), the default condition is represented by head-final compounds, since the large majority of Italian morphologically complex words (including derived and inflected words) are head-final. However, none of these studies presents conclusive evidence on whether a certain head position (initial or final) is taken as default. If a position is the default, a simple prediction is expected, namely that the stimuli in the default condition would require a less effortful processing than the stimuli in the nondefault condition (being the stimuli matched for all the relevant variables). No study so far has provided this kind of evidence. An advantage of head-initial as compared to head-final compounds (which could be interpreted as supporting a view of head-initial position as default) has been obtained only in two studies with experimental paradigms that strongly induce decomposition in constituents (Arcara et al., 2013; El Yagoubi et al., 2008). Conversely, the default position of head-final compounds has been inferred from results obtained with constituent priming (Marelli et al., 2009) or from different interactions between transparency or constituent frequency effects (Marelli & Luzzatti, 2012).

Rationale of the study

The first goal of this eye-tracking study is to investigate whether compounds are processed differently from noncompounds in ecologically valid reading conditions. To this aim it is important to further explore morphological complexity, in order to isolate complexity due to compounding from complexity due to the presence of multiple parts and suffixes. We thus used compounds as well as several types of noncompound words, including pseudocompounds such as those used by El Yagoubi et al. (2008), all embedded in sentences. We also included other complex words such as words with diminutive or augmentative suffixes and words with long stems followed by inflectional suffixes. This set will allow us to distinguish decomposition related to compounding or to morphological complexity and will overcome a limit of the existing literature, confined to comparison within types of compounds.

According to models such as the form-then-meaning account, a pseudocompound like “coccodrillo” should not trigger decomposition, since it resembles a compound word only in the first part (“cocco”) but has a nonmorphological ending in the second part (“*drillo”). This could be reflected in similar eye movement patterns for pseudocompounds and other complex words, while compounds could elicit different reading strategies. On the contrary, words like “tramezzino” (resembling a word with a diminutive suffix, “tramezzo” + “ino”, lit. “small partition” although with a meaning not related to its part, i.e., “sandwich”) could elicit decomposition, since they are fully decomposable in their parts. Since the meaning of the whole word is incongruent in comparison to the meaning that can be inferred by its parts, diminutives are expected to elicit longer reading times.

The second goal of this study is to explore the headedness effect, by ruling out possible confounds that might have biased previous results—namely, tasks inducing decomposition. We expect to replicate the headedness effect observed in previous neuropsychological study and thus to observe differences between head-initial and head-final compounds. With respect to Marelli and Luzzatti (2012), who found headedness effects in interaction with transparency and constituent frequency, we aim at further exploring potential differences related to the head, by employing a set of compounds that are matched for constituent frequency and predominantly highly transparent.

EXPERIMENT

Method

Participants

Twenty-four students of Università di Pisa and Scuola Normale Superiore participated in the experiment and received a monetary reimbursement for their participation. All participants were native speakers of Italian and had normal or corrected-to-normal vision. Data from three participants were excluded from the analysis because of excessive data loss. The final statistical analyses were performed on the remaining 21 participants.

Procedure

Eye movements (only for the right eye) were monitored with ALS 501 tracked at 240 Hz. A PC displayed the materials in white Courier fonts on a black VGA screen 65 cm from the participants’ eyes. The screen displayed 2 characters per degree of visual angle.

Participants were tested individually in a dimly lit room. They were asked to read each sentence carefully and then perform a semantic task, while their eye movements were recorded. The real goal of the experiment and the nature of the stimuli employed were mentioned only after the experiment was completed. A debriefing session followed the recording, by asking participants to rate a sample of compounds words. After a brief definition of the notion of “head” of a compound, participants were asked to indicate (on each word of the compound sample) whether the head was in first or in final position.

All experimental stimuli (i.e., the morphologically complex words) were embedded in meaningful sentences (see Table 1 for an example). All sentences had the same structure, which included

Table 1. *Example of the sentences used in the experiment*

Example	L'ingegnere / progettava / l'astronave / per la spedizione spaziale.
Translation	The engineer / designed / the spaceship / for the space expedition.
Regions	Subject / Verb / Object / Spillover

Note: The slashes ("/") were not displayed in the experiment and delimit the different regions of the sentence. Analyses were performed on the region of the object.

a subject, a verb, and an object, plus a spillover region. The experimental stimuli were always included as the object of the sentence and were displayed, as close as possible, on the centre of the screen.

Each sentence was preceded by a fixation point, displayed on the left and occupying the position of the first character in the sentence, which allowed for calibration and which lasted on the screen until the experimenter pressed a button to start the presentation. Then the sentence was displayed. The participant was asked to read the sentence and to press the spacebar when finished. After the sentence, the participant had to perform a simple semantic judgement task, to ensure that the sentence was correctly understood. Two words were displayed on the screen, and the participant was asked to decide which out of two words was semantically related to the previous sentence, by pressing one of two buttons. For example, after the sentence "L'ingegnere progettava l'astronave per la spedizione spaziale" (tr. "The engineer designed the spaceship for the space expedition"), the two words "luna" and "birra" (tr. "moon" and "beer), were displayed on the screen. Each participant was asked to decide, as quickly as possible, which of the two words was semantically related to the preceding sentence by pressing the button positioned on the side of the chosen word.

Material

The experimental list consisted of six categories of 20 stimuli each, thus leading to a total of 120 stimuli. Across the set, morphological complexity was varied through both compounding and derivation. Specifically, we selected head-initial and head-final compounds, as well as complex words resembling head-initial and head-final compounds (pseudocompounds). Since most Italian words are formed by a stem and by an inflectional suffix, and only a few are strictly monomorphemic (mainly adverbs), we varied complexity by including words either with additional derivative suffixes or with complex stems. Specifically, opaque diminutives or augmentatives were included as a case of complex words with multiple parts not related to compounds, and long words were included as a case of complexity related mainly to the polysyllabic stems. The stimulus categories were the following: head-initial noun–noun compounds (NN1, e.g., "capobanda", tr. "band leader"); head-final noun–noun compounds (NN2, e.g., "astronave", tr. "spaceship"); pseudocompounds with a segment homograph to a word in the leftmost part (PC1, e.g., "coccodrillo", tr. "crocodile", where "cocco", tr. "coconut", is homograph to a real word neither morphologically nor semantically related to the whole word); pseudocompounds with a segment homograph to a word in the rightmost part (PC2, e.g., "tartaruga", tr. "tortoise", where "ruga", tr. "wrinkle", is homograph to a real word neither morphologically nor semantically related to the whole word); words with a diminutive or an augmentative suffix but semantically opaque (DIM, e.g., "tramezzino", tr. "sandwich", where "-ino" is a diminutive suffix, but the meaning "tramezzino" is not a combination of its root morpheme "tramezzo", tr. "partition" and the suffix "-ino", indicating "small"); long words with a polysyllabic stem followed by an inflectional suffix (LON, e.g., "materasso", tr. "mattress"). Importantly, NN1, NN2, PC1, and PC2 had similar syllabic structure, with similar stress patterns. The psycholinguistic variables considered in the analysis were length and frequency (namely, surface frequency, i.e., the frequency of the word form). Length was calculated as the number of letters composing the stimuli. Frequency was first checked on a database of 3 million words of written Italian, fully lemmatized and annotated (Corpus e Lessico di Frequenza dell'Italiano Scritto, CoLFIS; Bertinetto et al., 2005), available through the web interface EsploraCoLFIS (http://linguistica.sns.it/esploracolfis/home.htm; Bambini & Trevisan, 2012b). However, frequency of compounds was often very low or equal to 0. Therefore, we decided to use a larger corpus of written Italian automatically tagged (La Repubblica, http://dev.sslmit.unibo.it/corpora/corpora.php; Baroni et al., 2004b), where frequency of the compounds remained low but allowed for more fine-grained comparisons. Nevertheless, the data from the two corpora were highly correlated, $r(118) = .85$.

Table 2 reports a summary of the categories of stimuli employed in the study and the relative values with respect to the major psycholinguistic variables.

Differences among stimulus categories were explored by means of two analyses of variance (ANOVAs) with the psycholinguistic measure (length and frequency) as dependent variables and the stimulus category (six levels) as factor. As post hoc analysis, a series of t-tests with false discovery rate correction were carried out.

The ANOVA on frequency showed a significant difference across categories, $F(5,\ 14) = 5.72$; $p < .001$. Post hoc t-tests showed the following significant differences: DIM > NN1 ($p = .001$), PC1 > NN1 ($p = .01$), PC2 > NN1 ($p = .001$), LON > NN1 ($p = .034$), NN2 > NN1 ($p = .027$). In summary, frequency of NN1 compounds was significantly lower than the frequency of all other categories.

The ANOVA on length also showed a significant difference, $F(5,\ 14) = 4.64$; $p < .001$. Post hoc t-tests showed the following significant differences: NN2 > DIM ($p = .03$), NN2 > PC2 ($p = .018$), NN1 > PC2 ($p = .03$), NN2 > PC1

Table 2. *Psycholinguistic variables of the experimental set*

Label	*Category*	*Example*	*Whole word surface frequency*	*Whole word length in letters*	*Surface frequency of first constituent*	*Surface frequency of second constituent*	*Length of first constituent*	*Length of second constituent*
NN1	Head-initial noun–noun compounds	“capobanda”, tr. “band leader”	3.60 (1.94)	10.65 (1.60)	9.70 (1.51)	8.50 (2.21)	4.55 (0.76)	6.40 (1.46)
NN2	Head-final noun–noun compounds	“astronave”, tr. “spaceship”	5.10 (1.32)	10.95 (1.67)	9.57 (1.95)	8.75 (1.66)	5.00 (0.72)	5.65 (1.22)
PC1	Pseudocompounds with a left embedded segment homograph to a real word	“coccodrillo”, tr. “crocodile”, where “cocco”, tr. “coconut” is the embedded segment	5.63 (2.07)	9.45 (1.19)	7.76 (1.69)	—	4.51 (0.51)	—
PC2	Pseudocompounds with a right embedded segment homograph to a real word	“tartaruga”, tr. “tortoise”, where “ruga”, tr. “wrinkle” is the embedded segment	6.34 (2.15)	9.40 (1.05)	—	7.51 (2.12)	—	4.95 (0.51)
DIM	Words with semantically opaque diminutive or augmentative suffix	“tramezzino”, tr. “sandwich” where “-ino” is the diminutive suffix	5.87 (1.09)	9.75 (0.85)	—	—	—	—
LON	Long words with polysyllabic stem	“materasso”, tr. “mattress”	5.15 (1.74)	9.85 (1.46)	—	—	—	—

Note: The table reports mean values of length and frequency of the experimental stimuli employed in the study (standard deviations in parentheses). Frequency values indicate the number of occurrences in the *Repubblica* corpus, logarithmically transformed.

($p = .02$), NN1 > PC1 ($p = .03$). In summary, compound words were slightly longer than almost all other categories.

Differences between NN1 and NN2 compounds were explored by comparing the frequency and the length of the individual constituents. The frequency of the first constituent was not significantly different, $t(38) = 0.24$, $p = .80$, and neither was the frequency of the second constituent, $t(38) = -0.38$, $p = .70$. The length of the individual constituents did not vary significantly, although NN1 compounds had slightly shorter first constituents, $t(39) = -1.91$, $p = .06$, and longer second constituents, $t(38) = 1.75$, $p = .08$, than NN2.

As for the comparison between PC1 and PC2, the embedded word did not differ for frequency, $t(38) = 0.417$, $p = .68$, but differed in length, $t(38) = -2.48$, $p = .02$, with PC2 having a significantly longer embedded word than PC1.

The frequency of embedded segments of PC1 and PC2 was compared to the frequency of constituents of NN1 and NN2. The frequencies of embedded segments of PC1 were significantly lower than the frequencies of first constituents of both NN1, $t(38) = -3.12$, $p = .003$, and NN2, $t(38) = -3.83$, $p = .004$. The frequencies of embedded segments of PC2 were significantly lower than the frequencies of constituents of both NN1, $t(38) = -9.78$, $p < .001$, and NN2, $t(38) = -7.00$, $p < .001$. The frequency of base word forming the leftmost part of DIM stimuli was compared with the frequency of first constituent of NN1 and NN2 and the frequency of embedded word of PC1. In these latter comparisons, DIM was significantly lower than NN1, $t(38) = -10.76$, $p < .001$, than NN2, $t(38) = -12.67$, $p < .001$, and than PC1, $t(38) = -8.16$, $p < .001$.

The structure of the stimuli was further explored from the distributional point of view by investigating the presence of points of discontinuity in the orthographic structure, which could signal the transition from an orthographic separated unit to another. To this aim we referred to the concept of "bigram trough" (Seidenberg, 1987). We adopted a conservative operationalization of "trough", following Rapp (1992): A bigram in a word was considered a "trough" if its two surrounding bigrams show a frequency at least 10 times higher than the frequency of the bigram itself (Rapp, 1992). We searched for troughs in our stimuli by referring to the CoLFIS corpus, using both the number of words containing a given bigram and the cumulative frequency of the words containing a given bigram. This analysis showed that across all the set only very few words contained bigram troughs, both when considering the number of words with that bigram (number of words with "bigram troughs"—raw number of words, NN1: 1; NN2: 2; PC1:2 ; PC2: 1; DIM: 1; LONG: 2), and when considering the cumulative frequency of the words with that bigram (number of words with "bigram troughs"—cumulative frequency, NN1: 1; NN2: 2; PC1: 2 ; PC2: 1; DIM: 1; LONG: 5). In a subsequent analysis, we investigated whether the word boundaries in NN1, NN2, PC1, and PC2 could be seen as a bigram trough. This analysis showed that, in NN1 and NN2, the few bigram troughs resulting from the previous analysis were the word boundaries. In the case of PC1 or PC2, the troughs were not at the boundary between the embedded word segment and other part of the word, but elsewhere.

Data on the transparency of NN1 and NN2 were also collected. As in Marelli and Luzzatti (2012), two types of transparency were taken into account —that is, the transparency of the whole compound and the transparency of the constituents. Two groups of 23 subjects each participated to two separate rating tasks. In the first rating task, participants were asked to rate, on a 4-point rating scale, how the meaning of the whole compound was predictable by its parts (i.e., transparency of the whole compound). In the other task, participants were asked to rate, on a 4-point rating scale, how the meaning of each constituent of the compound contributed to determine the meaning of the compound (i.e., the transparency of the two constituents). Mean rating scores were converted to proportions, ranging from 0 to 1. Table 3 reports descriptive statistics on transparency.

The mean transparency of compounds was considerably high (mean = .75). In the initial selection of compound stimuli for this study, we chose items

Table 3. *Transparency measures for compounds and their constituents*

Type	*Compound semantic transparency*	*First constituent semantic transparency*	*Second constituent semantic transparency*
NN1	.73 (.18)	.81 (.11)	.78 (.13)
NN2	.77 (.15)	.72 (.14)	.79 (.14)

Note: NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds.

that were as transparent as possible. However, it was not possible to satisfy all the constraints imposed by the experiment (e.g., to fall into a given range of frequency and length, to avoid the replication of constituents across stimuli, to avoid bound morphemes within compounds, etc.) and at the same time restrict the set to highly transparent compounds. For this reason, a few low-transparent compounds were included in the set.

NN1 and NN2 were matched for compound semantic transparency, $t(38) = -0.75$, $p = .46$. The difference between NN1 and NN2 in first constituent transparency approached significance, $t(38) = 1.94$, $p = .06$, with NN1 showing slightly more transparent first constituents than NN2. An opposite pattern was found for the second constituent, which showed a higher semantic transparency for NN2 than for NN1, $t(38) = -3.18$, $p = .003$. The comparison of modifier semantic transparency (i.e., second constituent of NN1, first constituent of NN2) showed no significant difference, $t(38) = 0.03$, $p = .97$, as well as a comparison of the transparency of the head constituents (i.e., first constituent of NN1, second constituent of NN2), $t(38) = -1.47$, $p = .14$. Within NN1, the first and the second constituent showed similar transparency, $t(19) = 1.84$, $p = .08$, whereas in NN2, the second constituent was rated as more transparent than the first, $t(19) = -3.49$, $p = .002$. This pattern of differences in transparency is not surprising, as the head of a compound is, by definition, the constituent that contributes more to the meaning of the whole compound (Dressler, 2006).

Overall, the analysis on the stimulus set indicates that NN1 and NN2 were significantly longer than PC1 and PC2, and that NN2 were also significantly longer than DIM. As for frequency, NN1 were significantly less frequent than all other categories of words. The analysis within compounds indicated that NN1 and NN2 were matched for constituent frequency and substantially matched for constituent length and transparency.

Analyses

To ensure that sentences with abnormal reading patterns or with major signal losses were excluded from the analysis, the following rejection criteria were adopted: trials that showed more than two blinks; trials with more than 20% of signal loss; trials in which the pattern of reading skipped two regions (following Frisson & Pickering, 1999); trials in which the behavioural response was wrong. With respect to the behavioural response, it is worth noting that participants reached an overall 96% accuracy with no differences across categories (DIM: 99%; PC2: 96%; PC1: 93%; LON: 95%; NN2: 97%; NN1 99%), indicating that the sentences were easily understood. The above criteria led to the exclusion of 19% of the total number of sentences. Each participant included in the statistical analysis had at least 73% of the total number of trials left after rejection.

After calculating the fixations from the raw data, the following measures were computed on the fixations that fell in the target region, defined as the area delimited by the experimental stimulus (i.e., the complex word) and its determiner,[1] functioning as the syntactic object in the sentence: first pass time (FirstPass), defined as the sum of all fixations occurring within the target region before the first fixation outside the region; total view duration (TotViewDur), defined as the sum of all fixations on the target region; total number of fixations (TotNumFix), defined as sum of the fixations on the target region; probability of making more than one fixations on the target region (ProbMoreFix). FirstPass was included as a

---

[1] As part of the syntactic object, the determiner was included in the region of interest. Fixations landed on the determiner index the processing of the target word as well, and removing them would lead to loss of important information.

measure of analysis related to word access, whereas TotViewDur and TotNumFix were included as measures of the overall effort needed to process the stimulus and integrate its meaning in the sentence (Bertram, 2011). The probability of making more than one fixation (at least 2), ProbMorFix, was included as further measure tapping on early stages of processing, because the need for making a further fixation is made during the first fixation (Hyönä, Bertram, & Pollatsek, 2004; Pollatsek & Hyönä, 2005).

Data were analysed by means of mixed-effect regression (Pinheiro & Bates, 2000). As an additional value with respect to traditional regressions, mixed-effect regressions allow consideration of the whole structure of data in terms of fixed and random effects, thus ensuring enhanced statistical power. A separate model was fitted for each of the measures considered as dependent variable. FirstPass, TotViewDur, and TotNumFix were logarithmically transformed before being entered in the analysis in order to reduce data skewness. ProbMoreFix was codified as a dichotomous variable, with a value of 1 indicating that more than one fixation has occurred and a value of 0 if just one fixation occurred. The ProbMoreFix variable was analysed by means of a logit mixed-effect model (Jaeger, 2008).

The mixed models fitted on data had the following initial structure: one dependent variable, four variables included as fixed effects, and two variables included as random effects. The fixed effects considered were the following: stimulus type (DIM, LON, PC1, PC2, NN1, NN2), word frequency, length of the target stimulus (including its determiner) as number of characters, and the ordinal position of the stimulus in the experimental sequence. The latter variable was included to account for practice effect. All interactions between the stimulus type and frequency and between stimulus type and length were also included. Subject and items were included in the analysis as crossed random effects (Baayen, Davidson, & Bates, 2008). The effect of type and of trial number (but not frequency and length) were modelled initially as random slopes of the subject random intercept term. Models that best fitted the data were chosen by a backward selection procedure. From an initial model including all variables, nonsignificant variables were removed one at time, starting from the one with the highest *p*-value. A likelihood ratio test was carried out between two models, identical except for that given term, to ensure that its presence did not affect the goodness of fit of the model. The same procedure was applied to exclude fixed effects and random effects that did not contribute to the goodness of fit of the model. In all mixed models (except the logit mixed model) the *p*-values reported are computed by mean of Markov chain Monte Carlo sampling (Baayen et al., 2008), a conservative alternative to *p*-value for linear mixed-effect models.

Results

Descriptive statistics on eye movement measures are reported in Table 4. Notably, these results may be misleading because they do not rule out the differences in length and frequency across the stimuli. These confounding effects were ruled out in the inferential analysis.

Analysis including all stimulus types

Results of fixed effects for the statistical models are reported in Tables 5–8. Neither the interaction between type and frequency nor that between type and length contributed significantly to the goodness of fit of the models, and thus only main fixed effects are discussed (the models including the interactions are included in the Appendix in Tables A1, A2, A3, and A4). Moreover, no random slope was included as random effect since the inclusion of random slopes did not improve the goodness of fit of the models. As for model criticism, after fitting the final models, a visual inspection of residual distribution and of correlation between fitted and observed data was performed to ensure that the models fitted satisfactorily the observed data. The influence of outliers was taken into account by refitting the final models excluding the observations whose residuals exceeded two standard deviations of residual distribution. Since in no cases did this outlier removal yield any

Table 4. *Descriptive statistics for eye movement measures*

Type	*FirstPass*	*TotViewDur*	*TotNumFix*	*ProbMoreFix*
NN1	519.08 (254.70)	576.19 (284.84)	2.79 (1.42)	.82
NN2	503.69 (247.65)	583.96 (308.76)	2.83 (2.51)	.85
PC1	456.92 (222.16)	526.89 (269.11)	2.50 (1.45)	.77
PC2	427.43 (198.47)	490.44 (248.50)	2.42 (1.34)	.76
DIM	418.13 (174.92)	470.08 (220.88)	2.30 (1.19)	.75
LON	461.39 (189.83)	512.02 (246.88)	2.39 (1.17)	.78

Note: The table reports the descriptive statistics (mean values, with standard deviations in parentheses) for all the eye movement measures considered. FirstPass = first pass times; TotViewDur = total view duration; TotNumFix = total number of fixations; ProbMoreFix = probability of more than one fixation. DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds.

difference from the models fitted on the full data, only models including all observations are reported.

All mixed-effect models included random intercept terms for subject and items with the following values (FirstPass, subject = 0.27, items = 0.08; TotViewDur, subject = 0.28, items = 0.12; TotNumFix, subject = 0.26, items = 0.10; ProbMoreFix, subject = 0.90, items = 0.29). These random intercepts improved significantly the goodness of fit of the models, as attested by a likelihood ratio tests.

In the model on FirstPass (see Table 5) a significant effect of length was found: As the length of the stimulus increases, the FirstPass increases. As expected, a negative effect of frequency was found: As the frequency of the stimulus increases, the FirstPass decreases. The trial number was also significant, suggesting that taking into account the practice effect led to a better model. Importantly, a significant effect of type was found. A significant effect means that a difference is expected for the significant term as compared to the reference stimulus—that is, the intercept—which in all models was the DIM category. The pattern of results suggests that NN2 and PC1 were associated with higher values of FirstPass than for the reference, whereas no difference was found for all other categories (PC1, LON, NN1).

Table 5. *Fixed effects for FirstPass*

Effects	*Estimate*	*Standard error*	t	p*MCMC*
DIM (intercept)	5.53	0.13	42.21	<.001
NN2	0.07	0.04	2.01	.04*
PC2	0.04	0.04	1.14	.25
PC1	0.07	0.04	1.99	.04*
LON	0.05	0.04	1.44	.14
NN1	0.04	0.04	1.04	.29
LENGTH	0.05	0.008	6.49	<.001*
FREQUENCY	−0.04	0.006	−6.44	<.001*
TRIAL NUMBER	−0.0007	0.0002	−3.26	<.01*

Note: The table reports the results for the fixed effects of the model fit on gaze duration measure. Asterisks indicate significant p-values ($p < 0.05$). FirstPass = first pass times; MCMC = Markov chain Monte Carlo; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds.

Table 6. *Fixed effects for TotViewDur*

Effects	*Estimate*	*Standard error*	t	p*MCMC*
DIM (intercept)	5.54	0.17	33.28	<.001
NN2	0.11	0.05	−2.22	.03*
PC2	0.08	0.05	−0.47	.07
PC1	0.12	0.05	0.19	.01*
LON	0.04	0.05	−1.39	.39
NN1	0.06	0.05	−1.50	.50
LENGTH	0.06	0.009	6.01	<.001*
FREQUENCY	−0.04	0.008	−5.35	<.001*
TRIAL NUMBER	−0.001	0.0002	−5.85	<.001*

Note: The table reports the results for the fixed effects of the model fit on the total view duration measure. Asterisks indicate significant p-values ($p < 0.05$). TotViewDur = total view duration; MCMC = Markov chain Monte Carlo; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds.

The difference between the coefficients of all categories was tested by means of an ANOVA on the estimated coefficients (with the aovlmer.fnc function, implemented in the languageR R package, Baayen, 2011). This test did not show significant results for any contrast (all ps > .05).

The model on TotViewDur (see Table 6) showed results similar to the model on FirstPass. A significant effect of length and frequency was found, with very similar values to those of FirstPass. Again, a significant effect of type was observed, with PC1 and NN2 showing significant differences with respect to the reference. In this case, the ANOVA between tests yielded a significant result for the contrast between NN1 and NN2 ($p = .04$), and between PC1 and all other categories (all ps < .05), and a difference approaching significance between NN2 and PC1 ($p = .06$). These results suggest a difference between compounds with different headedness and a more difficult processing for the PC1 stimuli than for the other stimuli.

The model on TotNumFix (see Table 7) showed a trend to the previous models: The

Table 7. *Fixed effects for TotNumFix*

Effects	*Estimate*	*Standard error*	t	p*MCMC*
DIM (intercept)	−0.07	0.16	−0.48	.62
NN2	0.10	0.05	2.13	.03*
PC2	0.09	0.05	1.92	.06
PC1	0.08	0.05	1.77	.07
LON	0.01	0.05	0.26	.79
NN1	0.03	0.05	0.67	.50
LENGTH	0.07	0.01	7.55	<.001*
FREQUENCY	−0.03	0.008	−4.46	<.001*
TRIAL NUMBER	−0.001	0.0003	−4.45	<.001*

Note: The table reports the results for the fixed effects of the model fit on the total number of fixation measure. Asterisks indicate significant p-values ($p < 0.05$). TotNumFix = total number of fixations; MCMC = Markov chain Monte Carlo; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds.

Table 8. *Fixed effects for ProbMoreFix*

Effects	*Estimate*	*Standard error*	*z-value*	p
DIM (intercept)	−1.93	0.85	−2.28	.02*
NN2	0.53	0.26	2.05	.04*
PC2	0.34	0.24	1.41	.15
PC1	0.26	0.24	1.07	.28
LON	0.21	0.24	0.90	.37
NN1	0.06	0.26	0.24	.80
LENGTH	0.30	0.05	5.55	<.001*
FREQUENCY	−0.11	0.04	−2.56	.01*

Note: The table reports the results for the fixed effects of the model fit on the probability of more than one fixation measure. Asterisks indicate significant *p*-values ($p < 0.05$). ProbMoreFix = probability of more than one fixation; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds.

length was positively associated with the number of fixations, while frequency and trial number were negatively associated. The effect of type was significant, with PC1 and NN2 differing from the baseline. The ANOVA tests for compound estimates showed a significant difference between NN2 and PC2 ($p = .04$) and only approached significance ($p = .06$) for the difference between NN1 and NN2, as well as the ANOVA for the difference between PC1 and PC2 ($p = .07$).

Finally, the model on the ProbMoreFix (see Table 8) showed results similar to those of the linear models. A significant effect of length and frequency was found, with the same pattern as that observed in all other models. Trial number was not a significant variable, and hence it was dropped from the final model. In this model, only NN2 had a significantly different estimate as compared to the reference value. To test the difference between NN1 and NN2, and between PC1 and PC2 in the logit mixed model, a different approach was used. Two more models were fitted on data, changing the reference level to allow the meaningful comparison. The difference between NN1 and NN2 approached significance ($p = .08$), whereas the difference between PC1 and PC2 was not significant ($p = .73$).

Details on the *p*-values of all contrasts are reported in the Appendix, in Tables A5, A6, A7, and A8.

The overall results suggest that, when taking into account length and frequency of the stimuli, a difference across categories emerges. NN2 and PC1 were consistently more difficult to elaborate in terms of time necessary for their processing (as attested by the results of FirstPass and TotViewDur), number of fixations elicited, and probability of eliciting more than one fixation. Figure 1 shows examples of single trial data for each of the stimulus categories, providing a pictorial representation of the reading patterns and TotViewDur effects. The results of the offline debriefing questionnaire indicate that the effect related to headedness cannot be traced back to the metalinguistic analysis of the stimuli. In the debriefing questionnaire, participants performed worse in attributing correctly the head of head-initial compounds (60% correct) than the head of head-final compounds (78% correct), $t(36) = -23.45, p < .001$.

Analysis including only compound words

The reliability of the headedness effect on Total View Duration was explored in an additional mixed-effect model that was fitted only on compound words. This analysis was run in order to allow for a better comparison of our findings with the results by Marelli and Luzzatti (2012). The procedure was the same as that in the previous analysis on all stimuli, but it initially included several variables (compound type, transparency,

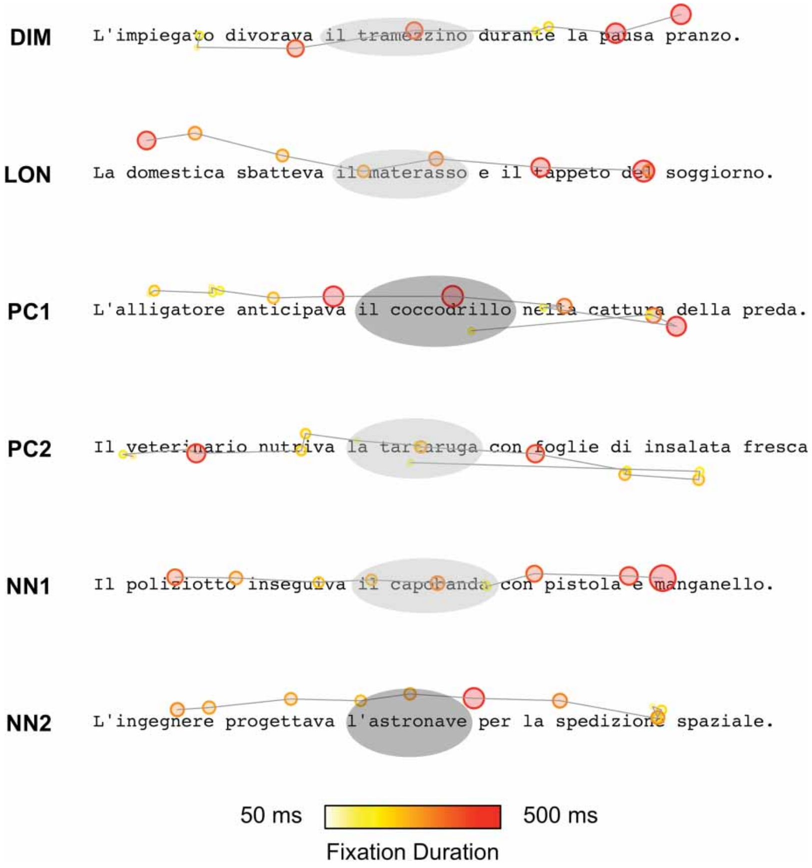


Figure 1. *Main results on total view duration in the analysis on all stimulus types. The figure illustrates examples of single trial data for each stimulus category, showing fixation positions over the sentences as coloured circles. Border colour and dimension code the duration of the fixations (with larger circles indicating longer fixations). Brighter colours indicate shorter fixation times, whereas darker colors indicates longer fixations times. The grey ellipses approximately delimit the target regions. The length of the vertical axis of each ellipse is proportional to the effect of TYPE as found in the mixed model on total view duration. Darker grey ellipses highlight the categories for which the effects of TYPE were associated with significantly longer reading times. DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. For the English translation of the target words see Table 2. [To view this figure in colour, please see the online version of this Journal.]*

compound frequency, constituent frequency, and compound length) and all their interactions as predictors. Nonsignificant effects were excluded from the model by backward elimination, following the same procedure as that used in the previous analyses. Differing from the previous analysis and following the Marelli and Luzzatti (2012) approach, the region of interest was here limited to the compound, excluding its determiner. As expected, transparency as a continuous variable did not improve the goodness of fit of the model of our stimuli (since they were mostly transparent) and was dropped during model selection. The final model of the analysis confined on compounds is reported in Table 9.

The model showed several significant effects, but a discussion on a term can be meaningful only taking into account the highest order interaction in which

Table 9. *Mixed model on compound words*

Effects	*Estimate*	SE	t	p
TYPE = NN2 (intercept)	27.23	9.88	2.76	.014
TYPE = NN1	−14.64	6.34	−2.31	.034*
FREQ.COMP	−3.32	1.69	−1.96	.068
FREQ1	−2.08	0.97	−2.16	.046*
FREQ2	−2.25	1.25	−1.79	.090
LENGTH	−1.64	0.81	−2.02	.061
TYPE = NN1 × FREQ.COMP	1.47	0.46	3.20	.005*
TYPE = NN1 × FREQ1	1.26	0.62	2.02	.059
FREQ.COMP × FREQ1	0.31	0.16	1.88	.078
TYPE = NN1 × FREQ2	2.81	1.35	2.08	.052
FREQ.COMP:FREQ2	0.44	0.23	1.90	.073
TYPE = NN1 × LENGTH	0.93	0.44	2.10	.052
FREQ.COMP × LENGTH	0.25	0.15	1.70	.108
FREQ1 × LENGTH	0.17	0.08	2.09	.053
FREQ2 × LENGTH	0.22	0.11	1.96	.066
TYPE = NN1 × FREQ.COMP:FREQ1	−0.10	0.04	−2.70	.015*
TYPE = NN1 × FREQ.COMP × FREQ2	−0.60	0.27	−2.27	.036*
TYPE = NN1 × FREQ.COMP × LENGTH	−0.04	0.03	−1.25	.230
TYPE = NN1 × FREQ1 × LENGTH	−0.08	0.04	−1.79	.092
FREQ.COMP × FREQ1 × LENGTH	−0.02	0.01	−1.69	.111
TYPE = NN1 × FREQ2 × LENGTH	−0.28	0.12	−2.27	.036*
FREQ.COMP × FREQ2 × LENGTH	−0.04	0.02	−2.07	.053
TYPE = NN1 × FREQ.COMP × FREQ2 × LENGTH	0.06	0.02	2.47	.025*

Note: Fixed effects for total view duration (TotViewDur) on the model taking into account all frequency measures. The table reports the model selected after backward elimination of nonsignificant variables. Random effects associated with the model were the random intercept for subject (*SD* = 0.26) and a random intercept for items (*SD* = 0.08). The p-values were calculated by means of lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2013). Asterisks indicate significant p-values ($p < 0.05$). FREQ.COMP = compound frequency; FREQ1 = frequency of the first constituent; FREQ2 = frequency of the second constituent; LENGTH = compound length; TYPE = compound headedness.

that term is involved. Thus, only significant highest order interactions are discussed.

The interaction of TYPE = NN1*COMP.FREQ*FREQ1 shows that first constituent frequency interacts with compound frequency differently for NN1 and NN2 (see Figure 2). In NN1, as the compound frequency increases, the facilitation related to first constituent frequency also increases (Figure 2, left panel). With low values of compound frequency, the pattern is reversed, with a slight inhibitory effect of first constituent frequency. In NN2, low first constituent frequency is associated with longer reading times than high first constituent frequency (Figure 2, right panel). However, this discrepancy becomes smaller as the compound frequency increases. In the highest values of compound frequency, the effect reverses: Lower first constituent frequencies are associated with faster reading times than higher first constituent frequencies.

To allow for a better comprehension of the effect of first constituent frequency, its interaction with length and headedness is reported in Figure 3. This interaction is not significant ($p = .09$).

The effect of TYPE = NN1*COMP.FREQ*FREQ2*LENGTH shows a complex interaction of compound frequency, second constituent frequency, and compound length, modulated by headedness (see Figure 4). In NN1, low second constituent frequency is associated with faster reading times: As the compound frequency increases, this beneficial effect increases, and this effect is higher for longer than for shorter compounds (Figure 4, Panels a, c, and e). The effect reverses for the lower range of compound frequency

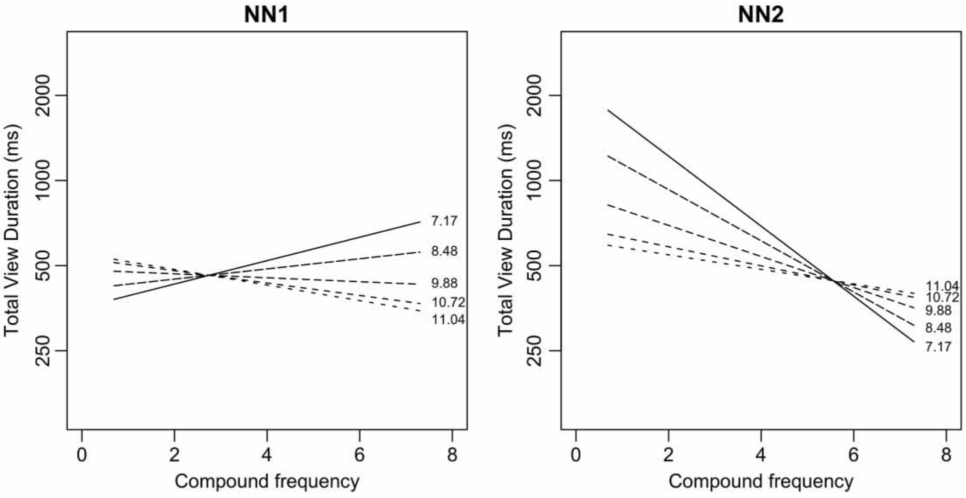


Figure 2. *Analysis on compounds: Results of the interaction between first constituent frequency, compound frequency, and headedness on total view duration in the analysis restricted to compounds. In the left panel, results for head-initial compounds are reported. In the right panel, results for head-final compounds are reported. NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. The compound frequency (log-transformed) is reported in the* x*-axis. The total view duration is reported in the* y*-axis (the original model was fitted on log measures, whereas the figure shows results back transformed to raw durations in ms). Different lines refer to the effects predicted for different values of first constituent frequency (for the 10th, the 30th, the 50th, the 70th, and the 80th percentiles of the values observed in the data). Values are adjusted for median value of length and second constituent frequency.*

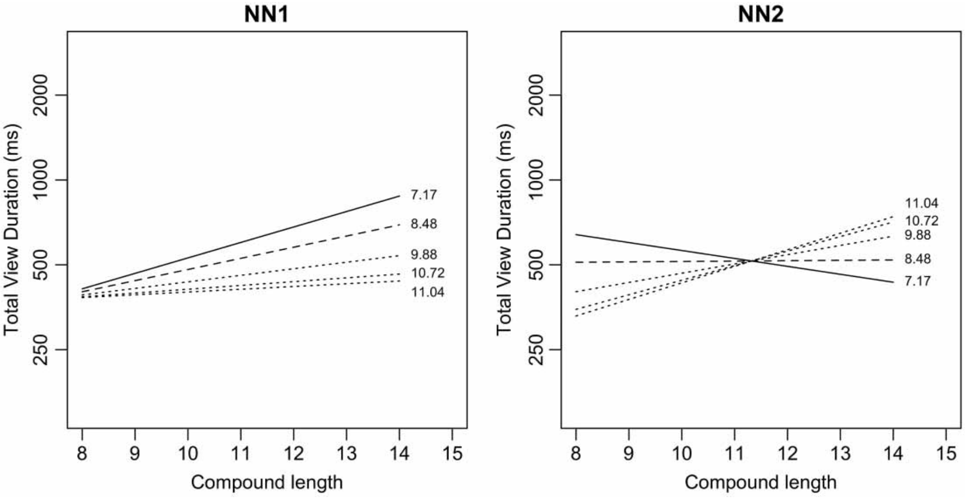


Figure 3. *Analysis on compounds: Results of the interaction between first constituent frequency, compound length, and headedness on total view duration in the analysis restricted to compounds. In the left panel, results for head-initial compounds are reported. In the right panel, results for head-final compounds are reported. NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. The compound length is reported in the* x*-axis. The total view duration is reported in the* y*-axis (the original model was fitted on log measure, whereas the figure shows results back transformed to raw durations in ms). Different lines refer to the effects predicted for different values of first constituent frequency (for the 10th, the 30th, the 50th, the 70th, and the 80th percentiles of the values observed in the data). Values are adjusted for median values of compound frequency and second constituent frequency.*

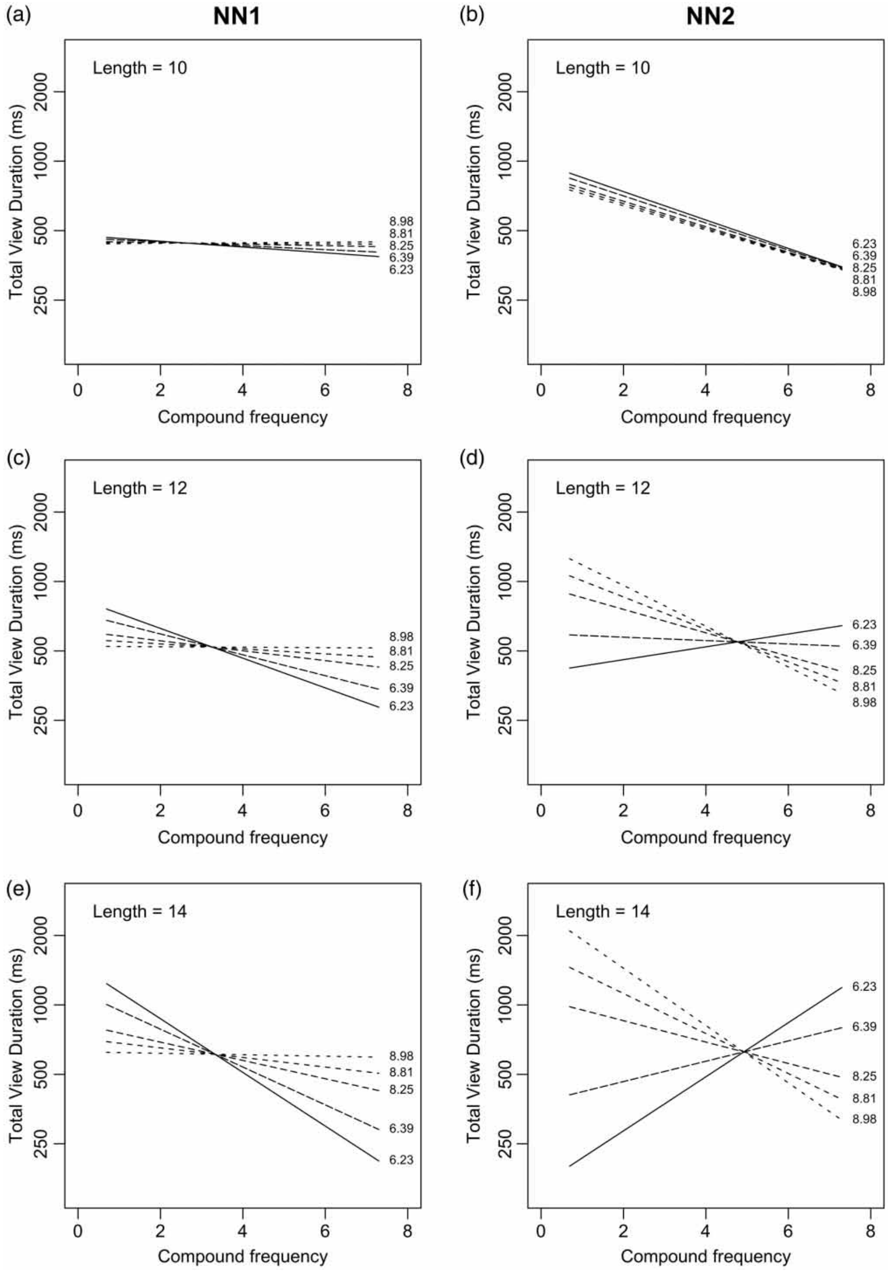


Figure 4. *Analysis on compounds: Results of the interaction between second constituent frequency, compound frequency, compound length, and headedness on total view duration, in the analysis restricted to compounds. In the left panels, results for head-initial compounds are reported. In the right panels, results for head-final compounds are reported. NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. Upper panels refer to compound with length = 10, middle panels refer to compounds with length = 12, and lower panels refer to compound with length = 14. The compound frequency is reported in the* x*-axis. The total view duration is reported in the* y*-axis (the original model was fitted on log measure, whereas the figure shows results back transformed to raw durations in ms). Different lines refer to the effects predicted for different values of second constituent frequency (for the 10th, the 30th, the 50th, the 70th, and the 80th percentiles of the values observed in the data). Values are adjusted for median values of first constituent frequency.*

(especially in longer compounds), in which low second constituent frequency is associated with slower reading times (Figure 4, Panels c and e). This effect is mostly visible when focusing on the lowest values of second constituent frequency, and it is almost negligible with the highest values of second constituent frequency.

In NN2, the pattern is different. In short compounds, as the compound frequency increases, the reading times are faster, with almost no effect of second constituent frequency (Figure 4, Panel b). For longer compounds, the effect of second constituent frequency depends on compound frequency. In the upper range of frequencies included in the experiment, a facilitation effect related to constituent frequency is observed (especially for the longest compounds). In the lower range of constituent frequency, the pattern is reversed, with an inhibitory effect of second constituent frequency (Figure 4, Panels d and f).

To sum up the results of this model, several interactions between compound frequency, constituent frequencies, and headedness were found. In general, there is a beneficial effect of first constituent frequency on NN1 and NN2. Low second constituent frequency had a detrimental effect on NN1 (especially with high frequency compounds), whereas the effect of second constituent frequency on NN2 were beneficial or detrimental, depending on compound frequency.

In order to understand whether the advantage of NN1 as compared to NN2 found in the Analysis Including All Stimulus Types section also holds in this analysis, we used the model fitted on compounds (Table 9) to make predictions on a new set of simulated compounds. The aim of this exploratory analysis is to understand whether it possible to conclude that, in general, there are faster reading times for one headedness condition than for the other. We generated a series of combinations of the variables considered in the model (headedness, compound frequency, compound length, constituent frequencies), and we observed the predictions of the model in these hypothetical scenarios. All the possible combinations were considered, and all the variables included in the analysis evenly spanned from the minimum to the maximum values observed in the experiment. Two sets of compounds were generated, which were identical except for the headedness: In one set all the compounds were head-initial, whereas in the other they were all head-final. These two sets allowed us to make a comparison between headedness across a wide range of variable combinations, even those not observed in our data (but within the range of values of our data). In this analysis, in 58% of the cases the head-initial compounds were predicted to have faster reading times than head-final compounds. This result highlights that the pattern of advantage of NN1 as compared to NN2 reported in the previous analysis (see Analysis Including All Stimulus Types section) is not always true. For several variable combinations the pattern reverses—that is, NN2 show faster reading times than NN1 (see, for example, Figure 4, Panel f, in the highest range of compound frequency). However, any inference should be drawn with caution, because this exploratory analysis does not take into account the actual distribution of stimulus properties (as constituent frequencies) in Italian, but rather it assumes that the properties are evenly distributed. Systematic differences may indeed occur in real language (for example, in Italian compounds first constituent frequency may be systematically higher than second constituent frequency, or the opposite), and thus the results from this exploratory analysis may not reflect what is expected when real data are tested.

GENERAL DISCUSSION

This study investigated compound and noncompound processing through eye movement recording. The goal of the study was two-fold: comparing compounds with pseudocompounds and other morphologically complex words, and investigating whether the head effect arises also in ecologically valid tasks such as sentence reading. Exploring these two aspects may provide a better understanding of the dynamics of processing morphologically complex word in a natural context and to address a highly debated issue in the literature on compounds—namely, the effect of headedness. In what follows, we consider the main results obtained on the whole set (see Analysis Including All Stimulus Types in the Results section) by discussing the higher reading times of pseudocompounds with a word-like segment in the left part (PC1) over all other categories and the higher times for head-final (NN2) as compared to head-initial compounds (NN1). We interpret these findings as reflecting the processing of noncanonical word structures. An additional paragraph focuses on the analysis of compounds (based on the section Analysis Including Only Compound Words in the Results section), in comparison with the study of Marelli and Luzzatti (2012), and highlighting other potential interpretations of the results.

Word structure effects in reading

Results on all the measures considered (first pass time, FirstPass; total view duration, TotViewDur; total number of fixations, TotNumFix; and probability of eliciting more than one fixation, ProbMoreFix) revealed significant differences between categories and especially highlighted two principal effects (see Analysis Including All Stimulus Types in Results section). First, pseudocompounds with a segment homograph to a real word in the leftmost part (PC1, e.g., "coccodrillo") elicited longer TotViewDur than all other categories. Second, head-final compounds (NN2) elicited longer TotViewDur than head-initial compounds (NN1). Similar effects (but with a less clear pattern of results) were found for FirstPass, TotNumFix, and ProbMoreFix. A difference approaching significance was also found for NN2 as compared to NN1 in the probability of eliciting more than one fixation. Importantly, all the effects were obtained after ruling out the effects of frequency and length, the two psycholinguistic variables that are expected to mostly affect reading (Kuperman et al., 2008). Given the convergent results of the measures employed, the discussion will focus on FirstPass and TotViewDur measures. Even if these measures encompass late stages of processing (mostly related to word access and integration into the sentence), we believe that they should show potential decomposition processes in word access, if any.

The first important effect observed in our study is the higher reading time required for PC1 than for all other types of words, indexed by TotViewDur. PC1 are words like "coccodrillo", which are not compounds and where the first part is a segment homograph to a word. More specifically, PC1 stimuli are formed by a pseudoconstituent and a nonmorphological ending (e.g., in PC1 stimulus "coccodrillo", "cocco" is homograph to a real world, but "drillo" is neither a word nor a suffix). PC1 stimuli thus resemble the orthographic condition ("brothel") employed, for instance, in the study by Rastle and Davis (2004), in which a decompositional effect is not observed and which is taken as supporting the form-then-meaning account. According to this model, only words whose structure resembles existing morphemes in all their parts are decomposed in their processing (e.g., "brother" is decomposed, "corner" is decomposed, whereas "brothel" is not; Rastle & Davis, 2004, 2008). By contrast, our results on PC1 seem to be in contradiction to this body of evidence (assuming that the effects in TotViewDur reflect the consequence of a decompositional process). The discrepancy can be traced back to several factors. First, the large majority of the results supporting the form-then-meaning account come from masked priming paradigms, whereas here subjects were engaged in a sentence reading task. Second, in the studies supporting the form-then-meaning account the dependent variable is reaction times in lexical decision, whereas in the present study the dependent variables are eye movement measures. The importance of these two differences was underlined also in the recent study by Marelli et al. (2013), which showed that, when the task is not lexical decision and when the dependent variable is related to eye movements, the typical pattern observed in studies on morpho-orthographic decomposition is not found (only a priming effect of "brother" on "broth" was found); yet the typical effects are confirmed in the lexical decision task (i.e., a priming effect for "brother" on "broth" and "corner" on "corn", but not of "brothel" on "broth"). Thus, the results by Marelli et al. (2013)

highlighted the contextual dependency of the effect on morpho-orthographic decomposition and suggest the importance of investigating different experimental settings. In this light, it is possible that the ecological reading paradigms employed here enhanced the decomposition of PC1, of course in relation to specific properties of the stimuli.

Another reason for the difference observed in the decompositional effects may indeed reside in the properties of the stimuli. Although our pseudocompounds resemble the words like "brothel" used by Rastle and Davis (2004) in that both are not composed by real morphemes, our stimuli differed notably from those typically employed in tasks designed to study morphological decomposition. Our stimuli are consistently longer than those in Rastle and Davis (2004; respectively, mean length of 10 letters versus mean length of 4.78 letters), and longer stimuli are more likely to be decomposed in their processing (Bertram & Hyönä, 2003). Furthermore, the prosodic and syllabic structure may have played a role. Prosody has been showed to influence the behavioural and brain responses to compounds (Koester, Gunter, Wagner, & Friederici, 2004), and its role is not confined to spoken recognition but rather prosodic features are also implicitly processed in reading (Stolterfoht, Friederici, Alter, & Steube, 2007). This may have influenced the processing of PC1, since these words had the same prosodic structure and stress patterns as compounds, thus enhancing decomposition. In the context of this study, it is not possible to further disentangle the role of prosody and length in eliciting decomposition. Prosody was kept constant across several stimulus categories (NC1, NC2, NN1, NN2), and other stimulus types with the same prosody as PC1 (i.e., PC2, NN1) never showed longer reading times. Length was kept relatively constant across stimuli and did not show significant interaction with stimulus types (see Table A1). The discrepancies in prosody and length between our stimuli and those traditionally used in the experiments testing the form-then-meaning account could be two reasons underlying the different results, but only a study encompassing stimuli with wide differences in prosodic structure and length could confirm these hypotheses.

In terms of eye movement measures, although a difference between PC1 and the reference level (i.e., DIM) was found for the measure of initial processing (FirstPass), the clearer pattern of results was found in the TotViewDur, in which PC1 differed significantly from all other stimulus categories. These findings suggest that additional processing costs are required for those stimuli that at a first glance resemble a head-initial compound word but that are not compounds, and these costs reflect in the global analysis of the stimulus and its integration with the sentence. Since TotViewDur is a global index of processing costs that might occur at different phases, it is possible to hypothesize that the costs for PC1 unfold in the initial word analysis (already observed in the FirstPass measure) and in later reanalysis. Late effects in reading of stimulus characteristics expected to influence earlier stages were already observed, for example, in Hyönä and Pollatsek (2000), who found that the effect of the first constituent frequency of compounds (typically affecting early stages of processing) influences also later measures of eye movements.

Based on the literature, the additional processing costs observed for PC1 can be interpreted as reflecting the decomposition of the stimulus in its parts and its reanalysis. The process for PC1 might be similar to the processing of structurally ambiguous sentences like garden path constructions, which are known to elicit early effects and later selective reanalysis of critical portions of the sentence (Frazier & Rayner, 1982). In addition, since the effect was found only for PC1 and not for PC2—that is, only when the embedded word was in the first position—it is possible that the serial order of the embedded word plays a crucial role. The presence of a word-like segment as first part probably generates the expectation that the whole word is a compound, which produces the reanalysis of the stimulus. Compatible evidence for a serial access to compound constituents was found for long Finnish compounds (Hyönä et al., 2004; Hyönä & Pollatsek, 1998).

To summarize, our results indicate higher costs for the processing of words that include a word-like segment in the leftmost part, which seem to undergo decomposition, and to generate a sort of morphological garden path effect. The difference between the results on PC1 and previous results reporting no decomposition for other morphologically complex words may be related to differences in the task, in the dependent variable used, and in the stimuli properties.

The results observed for the stimulus categories other than PC1 suggest that a complex morphological structure per se does not necessarily lead to longer reading times. PC2 did not differ from the reference level and from most of the other stimulus categories (DIM, LON, NN1), and, as explained before, this may reflect a serial access to constituent order. When a word-like segment is found in the final part, this does not elicit decomposition. The category of DIM was one of the categories that elicited the lowest reading times (together with LON, NC2, and NN1). The absence of effects on DIM is of special interest, as it contradicts the expectations that could be derived from the form-then-meaning account. Being fully decomposable in their parts, DIM would be expected to undergo decomposition during reading, leading to a semantic incongruity, as the composition of their parts gives a meaning that is not correct, as in "corner". Although the form-then-meaning account does not make explicit predictions on the semantic incongruity effect, but rather on pure (and early) morphological effects, it is reasonable to hypothesize that in DIM a decomposition could lead to the activation of word part meanings. If these meanings are not congruent with the whole word meaning, this could hamper the whole word access. This effect, typically not observed in lexical decision tasks, could indeed have been observed for the DIM category in our experiment, with different experimental task and different dependent variables, but this was not the case.

One could ask why the effect is found for PC1 but not for DIM. DIM are different from PC1 structures in two main aspects. First, DIM are formed by two bound morphemes—that is, a stem and a suffix that may appear only as a part of a morphologically complex word. On the contrary, the embedded segments in PC1 (e.g., "cocco", in "coccodrillo") are homograph to words that may occur freely in language, and they can be part of complex words only in compounds. Thus, there is a greater degree of similarity between PC1 and compounds than between DIM and compounds. The presence of "cocco" as first part induces the expectation that the whole word is a compound, generating the previously mentioned morphological garden path effect, while "tramezz" does not produce such expectation. Second, the prosodic structure of DIM is not the same as that of compounds, whereas that of PC1 is, and this could contribute to generate expectations.

Finally, a compound structure did not necessarily lead to longer reading times either. NN1 did not differ from other noncompound categories (DIM, LON, PC2), whereas NN2 did. Given that the two types of compounds were matched for the relevant variables, this suggests that NN1 are processed as the canonical condition and leads us to the next point.

The second main result of the study concerns the headedness effect. The effect arises already in the FirstPass measure, with longer reading times for NN2 than for DIM, and shows a clearer pattern in TotViewDur, in which NN2 exhibit significantly longer reading times than NN1, DIM, PC2, and marginally LON. Focusing on compounds, the result of longer TotViewDur for NN2 than for NN1 converges with evidence found in other studies that investigated the effect of headedness in different experimental paradigms. In particular, our data highlight an overall more effortful processing for NN2 than for NN1. Similar results were found in the event-related potential (ERP) study by El Yagoubi et al. (2008), in a study with neglect patients by Semenza et al. (2011), and in a recent study with ERP by Arcara et al. (2013). In the case of NN2, a reanalysis seems to be necessary in order to correctly assign the constituent properties—that is, to determine which constituent is the head of the compound (El Yagoubi et al., 2008). The analysis on all stimuli supports previous evidence that, at least in some cases, higher processing costs for head-final compounds are observed, providing new data from a more ecologically valid task such as sentence reading. Furthermore, the results of the offline debriefing questionnaire, where the participants performed better with head-final than with head-initial compounds, suggest that the eye movement results reflect differences in automatic processing, rather than different metalinguistic elaboration for head-initial as opposed to head-final compounds.

Conversely, our data partly differ from the eye movement data reported in Marelli and Luzzatti (2012), where the effect of headedness emerged only in interaction with semantic transparency and constituent frequency. This difference is probably motivated by the different materials across the two studies. The stimuli employed here are consistently more transparent than the stimuli in Marelli and Luzzatti (2012): Transparent compounds are more likely than opaque compounds to undergo decomposition during their processing (as pointed out also by Marelli & Luzzatti, 2012) and thus to exhibit headedness effects. Accordingly, the decomposition effect observed here is expected to be found for transparent compounds but not (or to a lesser degree) in opaque compounds.

Moreover, the overall set of stimuli employed here was different from Marelli and Luzzatti (2012), mainly because we included noncompound words: Here all partial effects of length and frequency were calculated by taking into account noncompounds as well, whereas in Marelli and Luzzatti (2012) only compounds were entered in the analyses. Another relevant difference is on the variables included in the statistical analysis. In our data, an advantage of NN1 as compared to NN2 was evidenced when taking into account the effect of length and frequency, whereas in Marelli and Luzzatti several constituent and whole word measures (and their interactions) were considered. To allow for a better comparison and to perform a deeper investigation of the effect, we reproduced the analysis in Marelli and Luzzatti on our data, limited to compound words, and the results are discussed in the section Constituent Effect and Headedness in Compound Processing (see below).

Nevertheless, there are converging findings across our analysis and that in Marelli and Luzzatti (2012). In Marelli and Luzzatti (2012), the headedness effect arose in similar eye movement measures (FirstPass, ProbMoreFix), supporting the importance of morphosemantic features in affecting Italian compound reading. This is in line with a recent extension of multiple-route models (originally introduced by Kuperman et al., 2008), according to which semantic features also may play an important role in compound processing (Marelli & Luzzatti, 2012). Thus, we believe that the results of the two studies must be taken as complementary. Whereas the study of Marelli and Luzzatti showed that the position of the head may interact with several variables (transparency and frequency) in influencing reading times, the current study shows that, when focusing on mostly transparent compounds, head-final compounds may exhibit longer processing than head-initial ones.

Our interpretation of the results is also in line with findings reported by Arcara et al. (2013). These authors showed that, when decomposition is elicited in compound processing, head-final compounds are more difficult to process than head-initial compounds. Since transparent compounds are presumably more prone to constituent detection (and our compound stimuli were mostly transparent), it is more likely that decomposition and access to constituents played an important role in reading our stimuli, thus leading to the detrimental effect observed for head-final compounds.

The reason why decomposition may have a negative effect in head-final compounds is probably related to a different mental representation of head-initial and head-final compounds and to the canonical order of head-modifier in Italian.

One reason for a different mental representation of head-initial and head-final compounds comes from theoretical linguistics, in relation to how compounds are generated. It was suggested that head-initial compounds are generated in analogy with syntactic order of elements, in which the noun precedes the modifier (Scalise, 1994), whereas head-final compounds are generated according to the structure of morphologically complex word (Di Sciullo & Williams, 1987). This view has been recently confirmed also by the corpus analysis by Radimský (2013), who found that, in novel compound production, Italian head-initial noun–noun compounds tend to be represented orthographically as two separated words, whereas head-final ones tend to be represented as a single word. Thus, head-initial compounds would be stored in the lexicon like word juxtapositions, whereas head-final would be stored as single (although morphologically complex) words. It is possible to hypothesize that, if a compound word is decomposed in its constituents, then it is treated as a sequence of words in a sentence. In this case, in Italian, a head-modifier structure would be the canonical one. According to this hypothesis, head-initial compounds would benefit from processing via a decompositional route, because the canonical head-modifier order is met. On the contrary, if head-final compounds are accessed via a decompositional route, an anomaly is found, because the typical head-modifier order is not met. This would explain the results observed in our study, consistently with Arcara et al. (2013). One could argue that a flat or hierarchical structure difference could explain the difference in reading times, without assuming that there is a default position for a condition. However, if this was the case, an advantage (in terms of processing) of NN1 as compared to NN2 would always have been observed. Indeed, literature suggests that only when decomposition is elicited does NN1 show an advantage as compared to NN2 (Arcara et al., 2013; El Yagoubi et al., 2008; Marelli and Luzzatti, 2012).

Importantly, the advantage of NN1 as compared to NN2 cannot be traced back to the distributional properties of headedness. Recent corpus linguistics studies show a quite reverse pattern, with a higher number of head-final compounds than head-initial ones, even in Italian (Guevara & Scalise, 2009, Marelli & Luzzati, 2012). Thus, based on distributional evidence, it is not possible to conclude that NN1 are likely to be considered as a default for compounds, because they are not the more common structure. The expectation of a canonical position cannot be related to the pure distribution of headedness, but rather on more general properties of the language, such as the syntactic order of elements. We argue that, in the case of Italian, since the head generally precedes the modifier, a head-initial compound can be (at least in some cases) the expected structure. Future studies could disentangle whether the distributional characteristics of headedness are more relevant than structural properties of language.

To summarize the results on headedness, a more effortful processing was found for head-final compounds (when partialling out the effect of whole-word frequency and length), presumably because they are processed via their constituents and do not meet the expected canonical order of elements, thus requiring reassignment of the constituent properties. This also provides additional explanations for the results on PC1 discussed above: The additional processing costs observed for PC1 and not for PC2 might be interpreted in light of the expectation of a canonical—that is, head-initial—compound, while word-like segments in the second part do not generate this sort of incongruity.

Overall, combining the results observed for PC1 and NN2, our findings suggest that complex words undergo decomposition, and, when the initial segment is a word or a segment homograph to a word, the stimulus is expected to be a compound, with the head in the first position. If the stimulus does not meet these expectancies, as in the case of PC1 and of NN2, then additional processing costs are necessary in order to reinterpret the structure or reassign the constituent properties. The incongruity in the structure of the word is probably detected in the first analysis of the compound, as indicated by the FirstPass results, and additional processes are carried out in later stages, as attested by the effect on TotViewDur.

Taken together, our findings suggest that it is not the morphological structure per se that elicits decomposition and longer reading times, but rather some feature of the words, possibly related to noncanonical structures: either a perceived noncanonical compound headedness (as for NN2) or an unexpected word structure (as for PC1). Presumably, this effect is not a simple consequence of distributional properties of compounds or PC1, but the result of structural properties. NN2, and compounds in general, are generated by means of a productive morphological process while the existence of PC1 is simply the consequence of accidental embedding of segments homograph to real words and do not reflect a mechanism of complex word formation. Yet it seems that both NN2 and PC1 appeared as noncanonical configuration for the processing system. We can also reasonably exclude some of the possible explanations in terms of internal distributional properties, for instance associated to "bigram troughs"—that is, statistical discontinuities in the bigram frequency. Our stimuli showed a very low number of bigram troughs (defined as in Rapp, 1992), and thus the presence of troughs cannot motivate stimulus decomposition. However, bigram trough is just one of the many statistical properties that could be investigated, and more recent models that do not posit any explicit morphological processing (i.e., "amorphous" models) may be able to capture all the effects described in the study (Baayen, Milin, Đurđević, Hendrix, & Marelli, 2011).

Finally, our results fit also with the most recent models of word reading that suggest that multiple routes are activated for morphologically complex words, taking into account several sources of information available (Kuperman et al., 2008), including semantic and headedness information (Marelli & Luzzatti, 2012). Words that show some unexpected features at different levels are solved by spending more time in reading and reanalysing the constituent parts. Along these lines, future studies may explore the influence of other variables, such as imageability or family size, that may reasonably play a role in compound processing.

It is important to notice that the analysis discussed so far has an intrinsic limitation. Given that several stimuli with different internal structure were included, we were able to partial out only the effect of the two main variables that are expected to affect the processing: word length and word frequency. With such an approach it is impossible to detect other effects related to word parts, and those effects might affect processing as well. Thus, these results above capture only the main effects, while they may hide more complex patterns interactions that could arise when taking into account the influence of word parts. We try to overcome the limitation of this approach in a deeper discussion restricted to compounds with different headedness (see the following section, Constituent Effect and Headedness in Compound Processing). Finally, the considerations made so far are valid for Italian, on which the experiment was carried out. In languages other than Italian, different expectations, possibly related to different word structures and different canonical orders of elements in the sentences, could lead to different patterns of morphological decomposition and, in turn, to different reading patterns.

Constituent effect and headedness in compound processing

In discussing the results of the analysis on all the stimulus categories we highlighted the longer reading times for NN2 than for NN1, and we interpreted this evidence as reflecting the effort in reassigning the constituent properties of compounds that are processed as noncanonical. Our findings and our interpretation are in contrast with those reported by Marelli and Luzzatti (2012), who suggested that when compounds are processed the head is searched in the final position. The difference observed between our results and those reported by Marelli and Luzzatti (2012), who focused on compounds, may be related to the different variables included in the analysis. To allow for a better comparison of the two studies, a further analysis on total view duration restricted to compounds was performed, reproducing as close as possible the procedure in Marelli and Luzzatti (2012; see Analysis Including Only Compound Words in the Results section). This analysis showed that several constituent and whole-word measures interact with headedness in influencing reading times.

First, we examined transparency as a continuous variable. In our results, transparency was not a relevant variable in influencing reading times, and this does not replicate the findings in Marelli and Luzzatti (2012). We argue that the absence of the effect of transparency (as a continuous variable) does not exclude a role of transparency in mediating the headedness effect (as showed by Marelli &

Luzzatti, 2012), but rather it is a consequence of the skewed distribution of transparency of the items included in our experiment, which were mostly highly transparent compounds (mean transparency = .75 in a 0–1 scale, whereas stimuli in Marelli & Luzzatti, 2012, had a mean transparency of .49 in a 0–1 scale). Other effects observed in our analysis slightly differ from those of Marelli and Luzzatti (2012), presumably because the effects in the present experiment refer to transparent compounds, and transparency was dropped from the analysis.[2]

Importantly, the results showed several effects of headedness, in interaction with compound frequency and both first and second constituent frequency. The role of constituent frequencies corroborates our assumption that compounds were processed via their constituent parts. The interaction of frequency measures with headedness strengthens the conclusion that head-initial and head-final compounds are processed differently. However, this analysis showed also that the advantage of NN1 as compared to NN2 is expected only in some conditions. In the following discussion, we interpret constituent frequency effects and compound frequency effects in terms of activation of, respectively, constituent representations or of whole-word representations[3] (see, for example, Kuperman et al., 2008).

A first important effect concerns the activation of the first constituent and its different role for NN1 and NN2. In NN1, as the access to whole word representation is likely to play an important role (when compound frequency increases), the activation of first constituent representation plays an important role as well. The effect is slightly inhibitory if the activation of first constituent representation is weaker (i.e., with relatively low values of first constituent frequency), and it is beneficial if the activation is stronger (i.e., with high values of first constituent frequency). In NN2, the pattern is different. Easier access to first constituent (as attested by the first constituent frequency effect) has almost always a beneficial effect for compound access, although this effect is less strong when the access to the whole-word representation is available (i.e., with high-frequency compounds).

The second important effect concerns the second constituent frequency modulation, which was different for NN1 and NN2. The role of the activation of the second constituent was visible mostly with longer compounds. A weaker activation of the second constituent was associated with faster reading times for NN1, when the compound frequency was high, but the opposite effect was found with low compound frequency. The more likely the compound is to be accessed also via a whole-word representation, the more inhibitory the effect of the second constituent becomes. In NN2, the access to the second constituent is beneficial with frequent compounds but it may be detrimental with less frequent compounds.

These results point to an interactive access to both whole-word and constituent representations, and their interactions suggest that the two representations are not serially accessed, but rather activated in parallel and interacting, at least in the stages reflected in the total view duration measure. These results replicate (in a different eye movement measure) the findings by Kuperman et al. (2008) and do not easily fit the strictly sublexical and supralexical models, according to which word parts and whole-word representations are thought to be serially accessed. Since total view duration includes both first pass and regressions, it is possible to argue that the interaction reflects the effects occurring at different serial stages merged together. However, if this were the case, pure additive effects would be expected, and not an interaction such as the one observed here. The interaction of

---

[2] Given that the study in Marelli and Luzzatti (2012) included also high-transparency items, one may ask why we do not replicate exactly their results here with a different distribution of transparency. We believe that there are many statistical reasons that can explain the discrepancies. For example, the mixed-effect model used by Marelli and Luzzatti (as the one used here) assumes linearity: If nonlinear effects are present for highly transparent compounds (more numerous in our data), some nuances may fail to be evidenced.

[3] Another interpretation of whole-word frequency effect has been suggested in the literature (Baayen, Wurm, & Aycock, 2007), but a discussion of this alternative explanation is not relevant to the aims of the current study.

the effect of length with the role of access to the second constituent is consistent with the results by Bertram and Hyönä (2003), who found that the length of the compounds influences the decomposition and the access to the constituents.

An extensive discussion of all the activation effects described above transcends the aims of this paper. What is important to stress here is that the overall pattern of results suggests that both the activation of constituent representation and whole-word representation differently affect the processing of head-initial and head-final compounds. This aspect was further investigated in an exploratory analysis in which reading times of compounds were predicted from the model fitted only on compounds, taking into account several possible values of length and frequencies. In this analysis, NN1 were predicted to have faster reading times in 58% of the cases. In light of these findings, we develop the following considerations, by integrating all the analyses carried out in the paper.

The complex pattern of results emerging from the analysis restricted to compounds jeopardizes our interpretation of NN1 as the expected structure as compared to NN2 (when considering only the main effect of length and frequency). The results highlighted in the section Analysis Including All Stimulus Types captured only the main effects, and a closer inspection of constituent effects reveals that head-initial compounds are not always expected to show faster reading times than head-final ones. A crucial aspect concerns the effect of second constituent frequency. Similarly to the study of Marelli and Luzzatti (2012), we found that the activation of the second constituent may boost head-final compound recognition, but almost always hamper head-initial compound recognition (although we observed this pattern only in some ranges of compound frequency). Below we focus on the effects of second constituent frequency, and we explain how this fits with the theory proposed by Marelli and Luzzatti, but might fit with a different interpretation as well. Specifically, we argue that the effect of second constituent frequency might reflect serial access, rather than headedness.

Marelli and Luzzatti (2012) argued that, given the different pattern of effects of second constituent frequency for NN1 as compared to NN2, the head is initially "searched" in the second position. Although we agree that this is a sound explanation, we believe that a different interpretation is possible as well, which would be compatible with the hypothesis that NN1 is (at least in some cases) the default condition. In our view, the important effect of second constituent frequency may arise as a consequence of the fact that it is the latter to be processed in a serial access to compound. The role of serial access to constituents for Italian compounds is corroborated by the results of Marelli and Luzzatti (2012) themselves, who in very early measures of word processing (i.e., first fixation duration) found only an effect of the first constituent and in later measures (including total view duration) found mostly effects of the second constituent. The measure considered in our analysis (i.e., total view duration) is likely to encompass later processing stages, in which the second constituent may have a bigger influence than the first one. Thus, the prominent effect of the second constituent in the total view duration may not necessarily be a consequence of "searching" the head in the second position, but it may reflect the integration of the representation of the second constituent with the representation of the first one (accessed in earlier stages).

The interpretation discussed so far could motivate the importance of second constituent frequency observed in our study and in Marelli and Luzzatti (2012), but it does not fully explain why the effect of this frequency may be detrimental for NN1 as compared to NN2 (as showed by Marelli & Luzzatti, 2012, and also observed in the highest range of compound frequency in this study, see Figure 4). The activation of the second constituent could indeed be inhibited in the case of head-initial compounds, because the meaning of the compound, and its syntactic properties, is mainly related to the first constituent and not to the second one. Following the proposal by Semenza et al. (2011), this could be explained in terms of conflicting attentional capture, driven by the frequency or by the headedness. Importantly, the serial access to constituents does not exclude a parallel influence of whole-word form, whose effects are consistently reported in the literature as well as in this study (Kuperman et al., 2008; Marelli & Luzzatti, 2012). To sum up, our results confirm that the second constituent plays an important role (in line with previous evidence: Andrews et al., 2004; Juhasz et al. 2003; Pollatsek et al., 2000, among others). In our interpretation, the role observed for second constituent frequency does not necessarily imply that the default condition of a compound is head-final. Rather, compound processing (especially the later stages) is mostly influenced by the integration of constituent meanings and the retrieval of morphosyntactic properties that occur differently for head-initial and head-final compounds. Presumably, the second constituent often exhibits a more important effect in the integration process, since it is serially accessed after the first. This alternative explanation fits both our results and the results in Marelli and Luzzatti (2012). Put in other words, while Marelli and Luzzatti (2012) argue that the difference in pattern of constituent effects (facilitatory or inhibitory) can be considered as diagnostic of which condition is taken as default, here we suggest that the same effects reflect the beneficial or detrimental influence of activating a constituent representation in a given processing stage.

The considerations made on this analysis, restricted to compounds and taking into account several measures, highlight that the advantage of NN1 as compared to NN2 is indeed questionable. When several covariates are considered, straightforward predictions become difficult to make with respect to which condition (head-initial or head-final) is taken as default, and the same results may lead to different interpretations (such as the one proposed here versus the interpretation by Marelli & Luzzatti, 2012). The detailed analysis on compounds, also including constituent measures, shows that the advantage for either head-initial or head-final compounds may depend on the stimulus properties. For example, long head-initial compounds with a relatively high compound frequency and low second constituent frequency (e.g., “pallacanestro”, tr. “basketball”) are expected to show faster reading times than head-final compounds with similar properties (e.g., “autonoleggio”, tr. “car rental”). Conversely, long head-initial compounds with relatively high compound frequency and high second constituent frequency (e.g., “guardiacaccia”, tr. “game keeper”) are expected to show slower reading times than head-final compounds with similar properties (e.g., “mondovisione”, tr. “worldwide broadcast”; see Figure 4, Panel f).

To sum up, there is not a simple answer to the experimental question concerning a default headedness condition. Indeed, there is not even a univocal answer to the straightforward prediction that, if a condition is a default, this condition would be processed faster, by excluding the influence of other variables. Taking into account the results of this study as well as the results of previous studies on compound headedness (Arcara et al., 2013; Marelli et al., 2009; Marelli & Luzzatti, 2012; Yagoubi et al., 2008), the advantage of (and expectations related to) one headedness condition compared to the other seems to depend on the characteristics of the stimulus. In some cases, a head-initial structure is expected, while in other cases, a head-final configuration is expected. In our overall stimulus set of complex words, the head-initial position was expected in most cases (as attested by the results of analysis including all stimulus types), and this is related to the canonical order of the head-modifier sequence in Italian syntax and to the specific properties of the experimental items (e.g., transparency). The results of the analysis including only compound words showed that, in some cases, faster reading times are expected for head-final compounds than for head-initial, possibly because in those cases the expectation is to find the head in the rightmost constituent. We believe that only a direct experimental manipulation, coupled with the study of frequency and transparency effects, could shed further light on this issue.

CONCLUSIONS

The present study confirmed—in a more ecologically valid task such as sentence reading—previous evidence of the difference between head-initial and head-final compounds, highlighting the decomposition mechanisms and, in some conditions, the additional processing cost needed for head-final structures. We interpreted these results as related to a difference that can be traced back to a perceived noncanonical position of the head in NN2, which requires a more effortful process to correctly assign the head properties, at least when mostly transparent compounds are used and when decomposition is enhanced. However, the interaction of several covariates (length and both whole word and constituent frequencies) with headedness precludes the possibility of drawing certain conclusions on the default position of the head, suggesting that the expectation of the head in a given position depends on several stimulus properties.

This study also suggests that it is not the morphological structure per se that elicits additional processing costs, but rather the analysis unexpected structures. This happens, for instance, for words embedding a segment homograph to a word in the left part. These words generate a sort of morphological garden path effect, as they are expected to be compounds (and, specifically, head-initial compounds) and need to be reanalysed accordingly. Discrepancies with previous results obtained for other morphologically complex words (e.g., derived words, Rastle & Davis, 2008) may be related to the type of task, the dependent variables measured, and the stimuli employed in addressing the domain of morpho-orthographic decomposition. As a final consideration, the use of ecological experimental paradigms such as sentence reading and the inclusion of stimuli of different morphological and structural complexity would allow for a more detailed description of the processing dynamics of morphologically complex words.

REFERENCES

Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology*, *3*, 1–12.

Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects?. *Journal of Memory and Language*, *25*, 726–740.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, *16*, 285–311.

Arcara, G., Marelli, M., Buodo, G., & Mondini, S. (2013). Compound headedness in the mental lexicon: An event-related potential study. *Cognitive Neuropsychology*, 1–20. doi:10.1080/02643294.2013.847076

Baayen, R. H. (2011). languageR: Data sets and functions with “Analyzing Linguistic Data: A practical introduction to statistics”. R package version 1.4. Retrieved from http://CRAN.R-project.org/package=languageR

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

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*, 438–481.

Baayen, R. H., Wurm, L. H., & Aycock, J. (2007). Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities. *The Mental Lexicon*, *2*, 419–463.

Bambini V., & Trevisan (2012a). *EsploraColLFIS*. Retrieved from http://linguistica.sns.it/esploracolfis/home.htm

Bambini, V., & Trevisan, M. (2012b). EsploraCoLFIS: Un’interfaccia web per le ricerche sul Corpus e Lessico di Frequenza dell’Italiano Scritto (CoLFIS). *Quaderni del Laboratorio di Linguistica*, *11*, 1–16.

Baroni, M., Bernardini, S., Comastri, F., Piccioni, L., Volpi, A., Aston, G., & Mazzoleni, M. (2004a). *La repubblica corpus*. Retrieved from http://dev.sslmit.unibo.it/corpora/corpus.php?path=&name=Repubblica

Baroni, M., Bernardini, S., Comastri, F., Piccioni, L., Volpi, A., Aston, G., & Mazzoleni, M. (2004b). Introducing the la Repubblica corpus. A large, annotated, TEI(XML)-compliant corpus of newspaper Italian. In Proceedings of the 4th international conference on language resources and evaluation LREC (pp. 1771–1774).

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton, A. M. (2005). *Corpus e Lessico di Frequenza dell’Italiano Scritto (CoLFIS)*. Retrieved from http://linguistica.sns.it/CoLFIS/Home.htm

Bertram, R. (2011). Eye movements and morphological processing in reading. *Mental Lexicon*, *6*, 83–109.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language*, *48*, 615–634. doi:10.1016/S0749–596X(02)00539–9

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 257–294). London: Academic Press.

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

Davis, M. H., & Rastle, K. (2010). Form and meaning in early morphological processing: Comment on Feldman, O'Connor, and Moscoso del Prado Martin (2009). *Psychonomic Bulletin & Review*, *17*, 749–755.

Delfitto, D., & Melloni, C. (2012). Compounding as a symmetry-breaking strategy. In Bambini, V., Ricci, I., Bertinetto, P. M., & Bertoncin, C. (Eds.), *Linguaggio e cervello - Semantica / Language and the brain - Semantics. Atti del XLII Congresso Internazionale di Studi della Società di Linguistica Italiana (Pisa, SNS, 2008)* (Vol. 1, pp. 155–177). Roma, Italia: Bulzoni.

Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes*, *20*, 75–114.

Di Sciullo, A. M. (2005). *Asymmetry in morphology*. Cambridge, MA: MIT Press.

Di Sciullo, A. M., & Williams, E. (1987). *On the definition of word* (Vol. 14). Cambridge, MA: MIT Press.

Dressler, W. U. (2006). Compound types. In G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 23–44). Oxford: Oxford University Press.

El Yagoubi, R. Y., Chiarelli, V., Mondini, S., Perrone, G., Danieli, D., & Semenza, C. (2008). Neural correlates of Italian compounds and potential impact of headedness effect: An ERP study. *Cognitive Neuropsychology*, *25*, 559–581.

Feldman, L. B., O'Connor, P. a, & Del Prado Martín, F. M. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, *16*, 684–691.

Fiorentino, R., & Fund-Reznicek, E. (2009). Masked morphological priming of compound constituents. *The Mental Lexicon*, *4*, 159–193.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178–210.

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology Learning Memory And Cognition*, *25*, 1366–1383.

Graffi, G., & Scalise, S. (2002). *Le lingue e il linguaggio. Introduzione alla linguistica*. Bologna: il Mulino.

Guevara, E., & Pirrelli, V. (2012). Understanding NN compounds, In V. Bambini, I. Ricci, P. M. Bertinetto, & C. Bertoncini (Eds.), *Linguaggio e cervello - Semantica / Language and the brain - Semantics. Atti del XLII Congresso Internazionale di Studi della Società di Linguistica Italiana (Pisa, SNS, 2008)* (Vol. 2, pp. 1–17). Roma, Italia: Bulzoni.

Guevara, E., & Scalise, S. (2009). Searching for universals in compounding. In S. Scalise, E. Magni, & A. Bisetto (Eds.), *Universals of language today* (pp. 101–128). Dordrecth, Netherlands: Springer.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & cognition*, *32*(4), 523–532.

Hyönä, J., & Pollatsek, A. (1998). Reading finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(6), 1612–1627.

Hyönä, J., & Pollatsek, A. (2000). Processing of finnish compound words in reading. In Kennedy, A., Heller, D., Pynte, J., & Radach, R. (Eds.), *Reading as a perceptual process* (pp. 65–87). Oxford: Elsevier.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.

Jarema, G., Busson, C., Nikolova, R., Tsapkini, K., & Libben, G. (1999). Processing compounds: A cross-linguistic study. *Brain and Language*, *68*, 362–369.

Jarema, G., Perlak, D., & Semenza, C. (2010). The processing of compounds in bilingual aphasia: A multiple-case study. *Aphasiology*, *24*, 126–140.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, *94*, 223–244.

Kehayia, E., Jarema, G., Tsapkini, K., Perlak, D., Ralli, A., & Kadzielawa, D. (1999). The role of morphological structure in the processing of compounds: The interface between linguistics and psycholinguistics. *Brain and Language*, *68*, 370–377.

Koester, D., Gunter, T. C., Wagner, S., & Friederici, A. D. (2004). Morphosyntax, prosody, and linking elements: The auditory processing of German nominal compounds. *Journal of Cognitive Neuroscience*, *16*, 1647–1668.

Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes*, *23*, 1089–1132.

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 876–895.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B., (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0-3. Retrieved from http://CRAN.R-project.org/package=lmerTest

Libben, G. (2006). Why study compound processing? An overview of the issues. In Libben, G., Jarema, G. (Eds.), *The representation and processing of compound words* (pp. 1–22). Oxford, UK: Oxford University Press.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language*, *84*, 50–64.

Longtin, C., & Meunier, F. (2005). Morphological decomposition in early visual word processing. *Journal of Memory and Language*, *53*, 26–41.

Marelli, M., Amenta, S., Morone, E. A., & Crepaldi, D. (2013). Meaning is in the beholder's eye: Morpho-semantic effects in masked priming. *Psychonomic Bulletin & Review*, *20*, 534–541.

Marelli, M., Crepaldi, D., & Luzzatti, C. (2009). Head position and the mental representation of nominal compounds. A constituent priming in Italian. *The mental Lexicon*, *4*, 430–454.

Marelli, M., & Luzzatti, C. (2012). Frequency effects in the processing of italian nominal compounds: Modulation of headedness and semantic transparency. *Journal of Memory and Language*, *66*, 644–664.

Mondini, S., Arcara, G., & Semenza, C. (2012). Lexical and buffer effects in reading and in writing Noun-Noun compound nouns. *Behavioural Neurology*, *25*, 245–253.

Moro, A. (2000). *Dynamic antisymmetry*. Cambridge, MA: MIT Press.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*, *20*, 261–290.

Pollatsek, a, Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology. Human Perception and Performance*, *26*, 820–833.

Radimský, J. (2013). Position of the head in Italian N + N compounds, the case of "mirror compounds". *Linguistica Pragensia*, *1*, 41–52.

Rapp, B. C. (1992). The nature of sublexical orthographic organization: The bigram trough hypothesis examined. *Journal of Memory and Language*, *31*, 33–53.

Rastle, K., & Davis, M. H. (2004). The broth in my brother's brothel : Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*, 1090–1098.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, *23*, 942–971.

Scalise, S. (1994) *Morfologia*. Bologna: Il Mulino.

Scalise, S., & Fábregas, A. (2010). The head in compounding. *Cross-disciplinary issues in compounding* (pp. 109–126). Amsterdam, Netherlands: John Benjamins.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Schwarze, C. (2005). Grammatical and paragrammatical word-formation. *Lingue e linguaggio*, *2*, 137–162.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 245–263). Hillsdale, NJ: Lawrence Erlbaum Associates.

Semenza, C., Arcara, G., Facchini, S., Meneghello, F., Ferraro, M., Passarini, L., … et al. (2011). Reading compounds in neglect dyslexia: The headedness effect. *Neuropsychologia*, *49*, 3116–3120.

Semenza, C., & Mondini, S. (2006). Neuropsychology of compound words. In Libben, G., Jarema, G., (Eds.) (2006). *The representation and processing of compound words* (pp. 71–95). Oxford: Oxford University Press.

Semenza, C., & Mondini, S. (2010). Compound words in neuropsychology. *Linguistische Berichte*, *17*, 331–348.

Stolterfoht, B., Friederici, A. D., Alter, K., & Steube, A. (2007). Processing focus structure and implicit prosody during reading: Differential ERP effects. *Cognition*, *104*, 565–590.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, *15*, 607–620.

APPENDIX

Table A1. *Fixed effects for FirstPass between type and frequency and between type and length*

Effects	*Estimate*	*Standard error*	t	p*MCMC*
DIM (intercept)	5.38	0.4	13.44	<.001
PC2	0.24	0.48	0.49	.63
PC1	0.17	0.54	0.32	.75
LON	0.27	0.49	0.54	.59
NN2	0.12	0.49	0.24	.81
NN1	0.32	0.46	0.69	.49
FREQUENCY	−0.01	0.02	−0.56	.57
LENGTH	0.05	0.03	1.85	.07
TRIAL NUMBER	0.0006	0.0002	−2.91	<.001
PC2 × FREQUENCY	−0.02	0.03	−0.77	.44
PC1 × FREQUENCY	−0.04	0.03	−1.28	.2
LON × FREQUENCY	−0.01	0.03	−0.27	.79
NN2 × FREQUENCY	−0.05	0.03	−1.54	.13
NN1 × FREQUENCY	−0.03	0.03	−1.02	.31
PC2 × LENGTH	0	0.03	−0.14	.89
PC1 × LENGTH	0.01	0.04	0.23	.82
LON × LENGTH	−0.01	0.03	−0.33	.74
NN2 × LENGTH	0.02	0.03	0.52	.61
NN1 × LENGTH	−0.01	0.03	−0.25	.81

Note: FirstPass = first pass times; MCMC = Markov chain Monte Carlo; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. Random effects associated with the model were the random intercepts for subject (*SD* = 0.26) and for items (*SD* = 0.08).

Table A2. *Fixed effects for TotViewDur between type and frequency and between type and length*

Effects	*Estimate*	*Standard error*	t	p*MCMC*
DIM (intercept)	5.37	0.52	10.38	<.001
PC2	0.36	0.63	0.57	.57
PC1	0.23	0.69	0.33	.74
LON	0.21	0.64	0.32	.75
NN2	0.19	0.64	0.29	.77
NN1	0.43	0.6	0.72	.47
FREQUENCY	0.06	0.03	1.65	.1
LENGTH	−0.01	0.03	−0.44	.66
TRIAL NUMBER	0.0013	0.0002	−5.82	<.001
PC2 × FREQUENCY	−0.01	0.04	−0.22	.83
PC1 × FREQUENCY	0	0.05	0.1	.92
LON × FREQUENCY	0	0.04	−0.1	.92
NN2 × FREQUENCY	0.02	0.04	0.44	.66
NN1 × FREQUENCY	−0.02	0.04	−0.4	.69
PC2 × LENGTH	−0.03	0.04	−0.74	.46
PC1 × LENGTH	−0.03	0.04	−0.85	.4
LON × LENGTH	−0.02	0.04	−0.45	.65
NN2 × LENGTH	−0.06	0.04	−1.46	.15
NN1 × LENGTH	−0.03	0.04	−0.77	.44

Note: TotViewDur = total view duration; MCMC = Markov chain Monte Carlo; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. Random effects associated with the model were the random intercepts for subject (*SD* = 0.28) and for items (*SD* = 0.12).

Table A3. *Fixed effects for TotNumFix between type and frequency and between type and length*

Effects	*Estimate*	*Standard error*	t	p*MCMC*
DIM (intercept)	−0.47	0.52	−0.91	.36
PC2	0.35	0.63	0.56	.58
PC1	0.46	0.7	0.65	.51
LON	0.44	0.64	0.69	.49
NN2	0.52	0.64	0.82	.42
NN1	0.77	0.6	1.28	.21
FREQUENCY	0.1	0.03	2.75	.01
LENGTH	−0.02	0.03	−0.56	.57
TRIAL NUMBER	0.0013	0.0003	−4.58	<.001
PC2 × FREQUENCY	−0.01	0.04	−0.22	.83
PC1 × FREQUENCY	−0.02	0.05	−0.44	.66
LON × FREQUENCY	−0.03	0.04	−0.63	.53
NN2 × FREQUENCY	−0.02	0.04	−0.38	.7
NN1 × FREQUENCY	−0.05	0.04	−1.13	.26
PC2 × LENGTH	−0.02	0.04	−0.62	.54
PC1 × LENGTH	−0.01	0.04	−0.41	.69
LON × LENGTH	−0.01	0.04	−0.26	.79
NN2 × LENGTH	−0.04	0.04	−0.96	.34
NN1 × LENGTH	−0.01	0.04	−0.4	.69

Note: TotNumFix = total number of fixations; MCMC = Markov chain Monte Carlo; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. Random effects associated with the model were the random intercepts for subject (*SD* = 0.25) and for items (*SD* = 0.11).

Table A4. *Fixed effects for ProbMoreFix between type and frequency and between type and length*

Effects	*Estimate*	*Standard error*	t	p*MCMC*
DIM (intercept)	−3.37	2.47	−1.36	.17
PC2	1.04	3.03	0.34	.73
PC1	0.76	3.36	0.23	.82
LON	1.08	3.1	0.35	.73
NN2	2.65	3.21	0.83	.41
NN1	4.21	2.89	1.45	.15
FREQUENCY	−0.08	0.15	−0.56	.58
LENGTH	0.39	0.17	2.33	.02
PC2 × FREQUENCY	−0.06	0.17	−0.38	.71
PC1 × FREQUENCY	0.03	0.17	0.15	.88
LON × FREQUENCY	0.05	0.18	0.28	.78
NN2 × FREQUENCY	−0.27	0.22	−1.25	.21
NN1 × FREQUENCY	0.01	0.17	0.07	.94
PC2 × LENGTH	−0.02	0.21	−0.09	.93
PC1 × LENGTH	−0.05	0.23	−0.21	.84
LON × LENGTH	−0.08	0.21	−0.4	.69
NN2 × LENGTH	−0.05	0.2	−0.24	.81
NN1 × LENGTH	−0.29	0.2	−1.49	.14

Note: ProbMoreFix = probability of more than one fixation; MCMC = Markov chain Monte Carlo; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. Random effects associated with the model were the random intercepts for subject (*SD* = 0.25) and for items (*SD* = 0.11).

Table A5. P-*values of contrasts for model on FirstPass*

	FirstPass				
Type	*NN1*	*NN2*	*PC1*	*PC2*	*LON*
NN1	—	—	—	—	—
NN2	.10	—	—	—	—
PC1	.11	.40	—	—	—
PC2	.37	.10	.11	—	—
LON	.29	.09	.30	.26	—
DIM	.29	.04*	.04*	.25	.14

Note: FirstPass = first pass times; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. The *p*-values that are not referred to the contrasts with the reference level (i.e., DIM) were calculated with the aovlmer.fnc from the R package languageR.

Table A6. P-*values of contrasts for model on TotViewDur*

	TotViewDur				
Type	*NN1*	*NN2*	*PC1*	*PC2*	*LON*
NN1	—	—	—	—	—
NN2	.04*	—	—	—	—
PC1	.02*	.06*	—	—	—
PC2	.15	.04*	.02*	—	—
LON	.63	.05	.01*	.15	—
DIM	.50	.03*	.01*	.07	.14

Note: TotViewDur = total view duration; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. The *p*-values that are not referred to the contrasts with the reference level (i.e., DIM) were calculated with the aovlmer.fnc from the R package languageR.

Table A7. P-*values for model on TotNumFix*

	TotNumFix				
Type	*NN1*	*NN2*	*PC1*	*PC2*	*LON*
NN1	—	—	—	—	—
NN2	.07	—	—	—	—
PC1	.16	.05	—	—	—
PC2	.11	.04*	.07	—	—
LON	.78	.05	.02*	.04*	—
DIM	.50	.03*	.01*	.07	.39

Note: TotNumFix = total number of fixations; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. The *p*-values that are not referred to the contrasts with the reference level (i.e., DIM) were calculated with the aovlmer.fnc from the R package languageR.

Table A8. P-*values for model on ProbMoreFix*

	ProbMorefix				
Type	*NN1*	*NN2*	*PC1*	*PC2*	*LON*
NN1	—	—	—	—	—
NN2	.08	—	—	—	—
PC1	.47	.27	—	—	—
PC2	.32	.47	.73	—	—
LON	.55	.23	.62	.86	—
DIM	.80	.04*	.28	.15	.37

Note: ProbMoreFix = probability of more than one fixation; DIM = words with semantically opaque diminutive or augmentative suffix; LON = long words with polysyllabic stem; PC1 = pseudocompounds with a left embedded segment homograph to a real word; PC2 = pseudocompounds with a right embedded segment homograph to a real word; NN1 = head-initial noun–noun compounds; NN2 = head-final noun–noun compounds. The *p*-values were obtained by fitting several models, changing the intercept.