UNIVERSITÀ DEGLI STUDI DI PADOVA

TESI DI DOTTORATO

Sede Amministrativa Università degli Studi di Padova

Dipartimento di Psicologia Generale

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE PSICOLOGICHE
INDIRIZZO PSICOBIOLOGIA
CICLO XXII

NEURAL CORRELATES OF MORPHOLOGICAL PROCESSING: THE CASE OF ITALIAN NOUN-NOUN COMPOUNDS

Direttore della Scuola : Ch.ma Prof.ssa Clara Casco
Coordinatore d'indirizzo: Ch.mo Prof. Angelo Bisazza
Supervisore : Ch.ma Prof.ssa Sara Mondini

Dottorando: Giorgio Arcara

Introduction

Noun-Noun compounding is a particular case in Italian morphology. Its peculiarity resides mostly on the ambiguity of *head* position. Generally, in a given language, the *head* of a compound (the word of the compound that determines semantic, syntactic and lexical feature of the whole compounds) is always in the same position. However Italian compounds can be both right and left headed.

How the cognitive system faces this ambiguity? How words with different headedness are represented in the lexicon? When information about head position is accessed? How the information about headedness is encoded?

Two psychophysiological studies and two neuropsychological studies have been carried out in order to answer to these questions.

The present dissertation that is organized as follows:

In part 1 a theoretical introduction to compounds is provided

In Chapter 1, compounds are analyzed in linguistic terms . Some specific issue of Italian compounding are raised: the interplay between syntax and morphology in Italian compounding and the difference between centre and periphery of word formation rules.

In Chapter 2, a survey of psycholinguistic and neurolinguistic studies of compounds is provided. Furthermore, a description of the main theories on mental representation and processing of compounds is given.

In part 2 four experiments on Italian compounds are described.

In Experiment 1 the early activation of head information of Noun-Noun compounds was studied through a morphological masked priming paradigm with ERP recording. Both behavioural and electrophysiological results suggest an early activation of head, particularly of left-headed compounds

In Experiment 2 the mental representation of Verb-Noun compounds and Noun-Noun compounds was investigated in through a task in which compounds were presented as whole words or as separated constituents. ERP were recorded during the task. An analogy between left headed Noun-Noun compounds and Verb-Noun compounds, compared to right headed compounds was found in behavioural data. The results suggest that word formation rules influence the way in which words are represented in the lexicon. ERP data suggested an greatest semantic relation between constituents of left headed Noun-Noun compounds compared to right headed Noun-Noun compounds.

In Experiment 3 Neglect dyslexic patients were asked to read Noun-Noun and Verb-Noun compound words. A lowest accuracy in reading was observed for right headed Noun-Noun compounds. Results suggest an interplay between attentional resources allocation and headedness.

In Experiment 4 two aphasic patients that showed peculiar behaviour with compounds were studied. Results are discussed underlining the interplay between lexical and peripheral aspects of compound processing

Results from all Experiments suggest that Headedness is indeed a feature that plays an important role in compound processing. *Head* can be seen as property that emerges from complex interaction of lexical, morpho-semantic and morpho-syntactic features. Through the activation of this information, the cognitive system is able to accomplish the task of dealing with compounds with different head positions.

The linguistic analysis of chapter 1 will specifically focus on Italian compounds. However comparisons with compounds of other languages (especially with English) will be given. *Palatino Italic* will be used to mark all examples in the text. After every example a translation in English will be given, enclosed in single quotes. When a crucial term will be described, and in general introduced for the first time, it will be in **Times New Roman Bold**. When necessary the same terms will be highlighted by *Times New Roman Italic*. Since the majority of the examples given will be in Italian and English, the language of the examples will be specified (enclosed in parenthesis) only if belonging to a different language. Since Italian compounds are mainly composed by two members, the majority of the example will be with two member compounds.

PART 1

1. COMPOUNDS IN LINGUISTICS

1.1 The morphological process of compounding

Compounding is the grammatical process that forms new words by the combination of existing words. Together with *derivation* (the grammatical process that generate derived words), it represents the way in which, in a language, it is possible to create new linguistic elements from the existing material. Words generated by compounding and derivation are morphologically complex words.

There is a main characteristic that makes compounding different from derivation. In derivation a free form is concatenated with a bound form (for example the English word *guitarist* is formed by the free form *guitar* and a bound form, the derivational affix *–ist*, or the Italian word *orchestrale*, 'orchestral' is formed by the combination of the free form *orchestra* and the derivational affix *-ale*[1]). In compounding, instead, two free forms are concatenated to form a new word (for example *baseball* is formed by free forms *base* and *ball*, or the Italian word *capogruppo* 'group leader' (lit. 'head-group') is formed by *capo* and *gruppo*. Under an historical perspective, compounding appears to be the fundamental word formation process and it seems very likely that its origin preceded that of derivation. This conclusion may be drawn by two main arguments: Firstly, it is unlikely that derivation could have started without an intermediate step of grammaticalization. Compounding could have represented this step with roots replaced successively, in derivation, by affixes. Secondly, as pointed out by Dressler (2006), languages may have compounding without affixation, but almost no language has affixation without compounding. This strongly suggests that

[1] The correct segmentation in morphemes would be *orchestr+al+e*, where *-e* is the morpheme for the inflectional affix indicating the traits [+singular, +feminine] of the whole word. Since to our analyses inflectional affixes don't play a fundamental role, here, and in the rest of the text, they will be disregarded and incorporated in the derivational affixes.

compounding might have preceded derivation and that could be the most ancient word formation process in languages.

1.2 Compounds in formal terms

Compounds are grammatical combinations of words, that is lexical item or lexemes, to form new words[2] (Dressler, 2006; Bauer, 2003). The two (or more) words that form a compound are called **constituents**. In more formal terms, a compound is a word formed by the concatenation of two (or more) words labelled with a given lexical category. The output consists in a new word labelled with a lexical category that may, or not, be the same of the lexical categories of constituents. This process may be represented as in 1:

1. $[\ ]_X [\ ]_Y \rightarrow [[\ ]_X \# [\ ]_Y]_Z$

X, Y and Z are lexical categories, X is the lexical category of the first constituent, Y is the lexical category of the second constituent and Z is the lexical category of the whole compounds. The hash mark (#) indicates a word boundary that keeps distinct the phonological and semantic identities of the two constituents (see later in this paragraph, criterion 3.d)). So for example, these are the structures of the A+ N English compound *blackboard* and of the Italian V+N compound *grattacielo* 'skyscraper' (lit. scrape sky').

2.

a) $[\text{ black }]_A [\text{ board }]_N \rightarrow [[\text{ black }]_A \# [\text{ board }]_N]_N$

b) $[\text{ gratta }]_V [\text{ cielo }]_N \rightarrow [[\text{ gratta }]_V \# [\text{ cielo }]_N]_N$

In the example 2.a) the input categories are Adjective and a Noun while the output category is a noun. In the example 2.b) the input categories are a Verb and a Noun, while the output category is again a noun. In Italian, as in many other languages,

---

2 As suggested by Lieber and Štekauer (2009) is hard to come up with a satisfying and universally accepted definition and the one presented is just one of the mainly acknowledged.

compounding generate mostly nouns. Other languages, however, are rich in other categories. Many Modern Greek compounds, for example, are verbs (Ralli, 1992). The output category of Italian compounds can be easily predicted: only if both the constituents are adjective the output is an adjective (e.g. *dolce amaro* 'bitter sweet', lit. 'sweet bitter'), any other combination generates a noun (see par. 1.4.1, for all possible combinations in Italian compounding).

A main issue in theoretical Linguistics is the distinction between *compounds* and *syntactic phrases*, multilexical units that are considered as a whole and that can be very similar to compounds (phrases are called *polirematiche* in Italian, Grossman & Rainer, 2004), but that, unlike compounds, include function words, as prepositions, or other link elements (they're sometimes called *prepositional compounds*). Some examples of syntactic phrases are *freno a mano* 'hand brake' (lit. 'brake by hand') or *papier à lettres* (in French, 'letter paper' lit. 'paper for letters'). Some linguists consider syntactic phrases as compounds (Di Sciullo, 2006), while other linguists consider the two phenomena as distinct (Dressler 2006).

Reasonably, regardless of the theoretical labels given, they represent different phenomena along the same linguistic continuum.

The criteria presented below thus represent only some of the possible linguistic criteria that can be used for identifying compounds.

3.

a) The resulting word is an "atom".
b) The resulting word denotes a unique entity.
c) It is possible to establish a semantic relationship between the constituents.
d) The constituents within the compound keep separate identity.

3.a) The first criterion indicate that compound words are impermeable to syntax. Compounds are "atoms" and other linguistic elements cannot be inserted within the compound without losing its original meaning. Thus, the V + N compound *portaombrelli* (lit. 'carry umbrellas') cannot be modified with the quantifier *molti* (many) in **porta**molti**ombrelli*. Compounds are furthermore "anaphorical island" (Bisetto & Scalise, 1999), no syntactic coreference to one of the constituents of the compounds can be made. So for example, using the compound *truck driver* it is ungrammatical says ‘*Truck drivers do not fill *them up* ‘ with the pronoun *them* that corefers to the compound internal word *truck* (Dressler, 2006). The criterion of syntactic atomicity is however also applicable to syntactic phrases that in this sense, behave just like compounds (e.g., *freno a mano,* see above). For example, it is not possible to say **freno* ***difettoso*** *a mano* (‘lit. brake defective by hand’), by inserting an adjective within the phrase.

3.b) With the second criterion is meant that compound words denotes always a unique referent. For example the compound *policeman* even if composed by two words denotes only one entity: a member of a police force.

3.c)The semantic criterion, according to which it is possible to identify a semantic relation between the compound and its constituents, posits some difficulties and thus, is not always applicable. In novel compounds the semantic criterion is fundamental, for example, let’s consider the novel compound *documento chiave* ‘key document’ (lit.’document key’), generated in analogy to other compounds as *parola chiave* ‘key word’ (lit. ‘word key’). In this case a *documento chiave* can be easily understood as an important document with a crucial role for something. A metaphorical semantic relation can be thus found between the constituents *documento* and *chiave*. A semantic relation should be transparent in order to make a novel compound understandable, but this is not true for already existing compounds. Compounds that reside for long in a language may become linguistically **lexicalized**, that is they may become a whole unit treated as other noncomplex word and the semantic relation between the constituent and the whole compounds can be difficult to be determined (see par. 1.4.4).

3.d) As outlined in 1., the constituents of compounds usually keep separated identities (as formally indicated by the sign #). This aspect can be tested in Italian by some phonological tests. For example, according to phonological rules of Northern Italy it is expected the sonorization of the /s/ in intervocalic position (Scalise, 1994). This normally happens in simple and suffixed words. This sonorization however is not realized in the case of compounds.

4.

a) riso → ri[z]o (riso, ‘rice’)

b) prendisole → *prendi[z]ole (prendisole, ‘sundress, lit. take sun’)

While in the example in 4.a), that is a simple word, the phonological rule is applied, the same is not true for 4.b) in which a phonological separation of the constituents *prendi* (lit. 'take') and *sole* (lit 'sun') is maintained.

The other criterion that allows us to identify compounds is the stress assignment. In English for example the compound *blackboard*, can be distinguished from the syntagma *black board* (indicating a board whose colour is black) relying on the different stress assignment.

5.

a) blàck bóard

b) bláckboard

c) sálva génte ('save people')

d) sàlvagénte ('safety jacket')

In 5.a) and 5.b) are represented the different stress patterns in the syntagms and in the compounds both formed by the two words *black* and *board*. While in the nominal syntagms there is a minor stress in *black* followed by a major stress in *board*, in the compound there is only one stress applied to the first member *black*.

In 5.c) and 5.d) analogous examples are given for Italian. While the phrase *salva gente* 'save people' has two major stresses, the compound *salvagente* has one major stress in the second constituent and one minor stress in the first constituent. Usually German languages (English, Dutch, German etc.) tend to have the major stress on the first constituent, while Romance languages (Italian, Spanish, Portuguese) on the second constituent. This however is not always true, since it is possible to find many exceptions to this rule (eg. *Apple píe* has the major stress on the second constituent and not, as expected, on the first one).

Summarizing, although many criteria can be found for the identification of *compounds*, these criteria are not without any problem and several exceptions for each criterion can be found. The first criteria introduced, the 'syntactic' criteria, seems to be the most reliable although it is valid also for syntactic phrases.

Finally, it is important to underline that the orthographic features play no role in determining the status of a compound. A common error of "Folk Linguistics" is in fact that compounds are words formed by the combination of two or more words, *written as a single word*. As seen in the previous analysis, the linguistic identification of a compound never relies on its graphical representation. Several examples of compounds written as separated words have been already given. Orthographic representation of compounds, in fact, is not univocal. Italian and English behave similarly in this aspect since compounds can be graphically represented in three possible forms:

6.

a) as conjunct words.

blackboard, astronave (lit. 'spaceship')

b) as words separated by an hyphen.

long-term, *verde-bottiglia* (lit. 'green bottle')

c) as separated words

black hole, *uomo ragno* (lit. 'man spider')

Let's consider example in 6.c). A *black hole* (if considered as the astronomical term) is a compound word. It is possible to say *an **enormous** black hole*, still referring to the same concept. It is not however possible to say **a black **enormous** hole* if the intention is still to refer to the astronomical name. *Black hole* thus respects the criterion 3.a) given previously for identifying compound words. Hence, the presence of a space between the constituents has no influence on the characteristics of the whole compound.

1.3 Head of a compound

The **head** of a compound is the most important member of the compound. The concept of head is crucial in compounding (and in general in morphology), has an important role in compound classification (see par. 1.4) and, as we will see in the next chapter, is very important in psycholinguistic literature for several reasons.

To give a more rigorous definition, the head is the constituent that:

7.

a) determines the lexical category.

b) determines the syntactic properties (e.g. number and gender).

c) determines mainly the semantic properties of the compound as a whole.

By examining some examples it is possible to notice some important characteristic of headedness. Let's consider the example in 8

8.

a) $[[\text{black}]_A \# [\text{hole}]_N]_N$

b) $[[\text{croce}]_N \# [\text{rossa}]_A]_N$

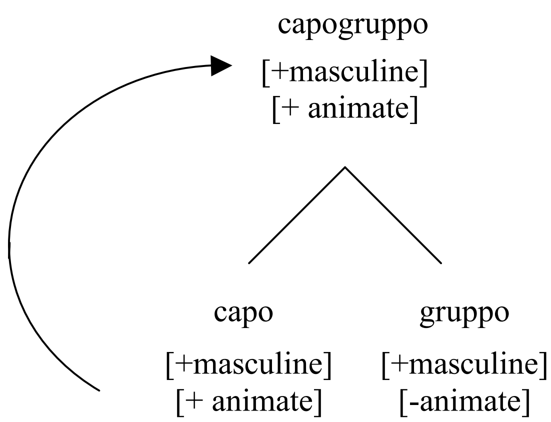
c) $[[\text{space}]_N \# [\text{ship}]_N]_N$

d) $[[\text{capo}]_N \# [\text{gruppo}]_N]_N$

According to the criterion 7 the head of a compound can be identifyied by analyzing the lexical category of the whole compound and of the constituent. According to this criterion the head is easily identifiable in 8.a) and 8.b). In the English compound *black hole*, the lexical category of the output is a noun, so the head must be the noun *hole* the second constituent and cannot be the adjective *black*. The same is true for the Italian compound *croce rossa* ('red cross, lit. cross red') where, since the whole compound is a noun, the head is *croce* 'cross' , the only element that is a noun. In the examples in 8.c) and 8.d), the situation is different: the output category is a noun and both constituent are nouns. According to the criteria 7 both members of each compound are potential candidate for being head, and so it is necessary to rely on the other criteria. The criterion 7.a) cannot help 8.c) because syntactic features are not marked in English. The same criterion cannot help us also in 8.d) since, even if Italian has different inflectional suffix that could allow us to distinguish the head, both constituents in *capogruppo* 'group leader' (lit. 'head group') are singular/masculine and thus marked with the thematic vowel *–o*. Hence, we must rely on 7.b) that tell us that the head constituent is the one with that have the greatest influence in determining the semantic properties of the whole compounds.

The simplest way for testing this is the application of the semantic test «IS A». Since *spaceship* «IS A» *ship* and not a *space*, *ship* must be the head of the compound. The same rule can be easily applied for *capogruppo* that «IS A» *capo* and not a *gruppo*.
The criterion 7.b) can be also applied via the analysis of the semantic traits of the constituents and the whole compounds. For example *capogruppo* is [+masculine] [+animate], *capo* is [+masculine] [+animate], and gruppo is [+masculine] [-animate]. Since the head is the constituent that mostly determines the semantic traits of the whole compound, this analysis converge in assigning to the word *capo* the role of head. The property of headedness can be graphically represented expressed as in 9.

9.



The linguistic movement that brings all the information from the head to the whole compounds is called **percolation**.

So far, several criteria for identifying the compound headedness have been introduced. However in some compounds the identification of the head posits some problems.

10.

a) mother-child

b) pellerossa ('redskin, lit. skin red')

c) saca punta (Spanish, 'pencil sharpener, lit. get top')

Trying to verify the criteria in all compounds given in 10. neither of the constituents can be identified as the head. Let's focus on the semantic criterion 7.b). In the English compound in 10.a) *mother-child* neither «IS A» *mother*, nor «IS A» *child*, but it is referred to the relationship between mother and child. Let's consider the Italian compound *pellerossa* in 10.b). A *pellerossa* neither «IS A» *pelle* ('skin') nor «IS A» *rossa* ('red'), but it is a person with red skin. The same is for the Spanish compound in 10.c) *saca punta* neither «IS A» way of *sacar* nor «IS A» kind of *punta*, but is an object whose function is that of sharpening pencils. In some compounds, so, is not possible to identify the head within the compound, that it is "outside" the compound itself. These compounds are called **exocentric** and are opposed to **endocentric** compounds: the compounds in which the head is one of the constituents. Some

categories of compounds are systematically exocentric. For example, almost all V + N Italian compounds are exocentric.
There are also compounds in which more than one constituent satisfies the conditions of being the head.

11.
a) bitter sweet
b) attore-regista ('actor – director')

In the examples in 11. constituents carries the same amount of information to the whole compounds and is not possible to determine an asymmetry between the role of the constituents[3]. So it is as the compounds have two heads. These compounds are called **coordinate compounds**.

To summarize, the head is an important concept in morphology and in compounding. Compounds may have a constituent that is the head, more than one constituent that is the head (the *coordinate* compounds), or the head "outside" the compound (the *exocentric* compounds).

Although useful for an explanation of the concept of head, the one given is not a satisfactory classification of compounds. In the next paragraph a survey of some possible classifications of compounds will be made. The concept of headedness will be further analyzed in par 1.7 in which the particular case of Italian Noun-Noun compounds will be discussed.

1.4 Classification of compounds

Compounds can be classified in several ways. Every classification is useful to underline important aspect of compounding and give us information about linguistically relevant properties of compounds. Five types of classification will be

[3] it has been argued that an asymmetry is present: the first constituent is the one whose influence is more prominent for the meaning of the whole compound. An *attore – regista*, in fact, is not exactly the same of a *regista – attore*.

introduced in order to give a complete picture of compound types and characteristics, with a specific focus on Italian compounds.

The first classification (1.4.1) is made according to the lexical categories of the constituents. The second classification (1.4.2) divides compounds according to the relation between constituents. The third (1.4.3) classifies compounds according to the surface structure (although this structure may reflect deeper features of the compound). The fourth (1.4.4) classifies compounds according to the transparency of the relation between the meaning of the constituent and the meaning of the whole compound.

1.4.1 Compound classification based on constituent categories

Compounds can be obtained by the combination of words belonging to several categories. However, in any language, not every possible combination of words generates grammatical compounds. In Table 1 (adapted from Scalise, 1994) are represented the main combinatorial possibilities of Italian words.

Categories	**existence**	**productivity**	**example**	**translation**
N + N	yes	yes	*crocevia*	crossroad (lit.'cross way')
A + A	yes	yes	*dolceamaro*	'bittersweet' (lit. 'sweet bitter')
V + V	yes	no	*giravolta*	'spin' (lit .'spins turns')
P + P	no		**senzaper*	(lit.'without for')
Adv +Adv	yes	no	*sottosopra*	'upside down' (lit.'down upside')
V + N	yes	yes	*cantastorie*	'storyteller' (lit. 'sings stories')
V + A	no		**vedibuono*	(lit. 'see good')
V+ P	no		**saltasopra*	(lit.'jump on')
V + Adv	yes	no	*buttafuori*	'bouncer' (lit. 'throws outside')
N +A	yes	no	*camposanto*	'cemetery' (lit. 'field holy')
N +V	yes	no	*manomettere*	'tampering' (lit. 'hand put')
N + P	no		**scalasotto*	(lit. 'stair under')
N + Adv	no		**casabene*	(lit. 'house well)
A + N	yes	no	*gentiluomo*	gentleman (lit. 'gentle man')
A + V	no		**caropaga*	(lit.'expensive pays')
A + P	no		**biancosenza*	(lit. 'white without')
A + Adv	no		**nerodomani*	(lit. 'black tomorrow')
P + N	yes	no	*sottopassaggio*	'subway' (lit. 'under passage')
P + A	no		**soprastrano*	(lit 'over strange')
P + V	yes	no	*sottomettere*[4]	'submit' (lit. 'under put')

Tab . 1. Combinatorial Possibilities in Italian compounding (from Scalise, 1994). A=Adjective, Adv=Adverb, N=Noun, V=Verb, P=Preposition. The column labelled "existence" indicate if there are compounds generated by the combination, while the column labelled "productivity" indicate if current Italian the category is productive.

As can be seen from Tab. 1 not every combination is possible, since some of them produce ungrammatical compounds. It is important to notice that not all categories are of equivalent importance. Every language shows preference for some combination

[4] P + V is a problematic category. Another explanation of these constructions is that the preposition is a prefix and not a word, thus making the whole word a prefixed word and not a compound.

sover the other. In general, N + N compounds represent the largest subclass of compounds in most languages and the preferred form of compounding). The recognition of this interesting aspect can be traced back to the definition of compounding itself. The term compounding (or composition) goes back to the Latin *vocabulorum genus quod appellant compositicium* ‘the word class which is called composite (Varro) and *figura nominum composita* of some Ancient Roman grammarians (among which Priscianus and Donatus). Here the Latin *com-positum* is a literal translation of Greek συνθετον *syn-theton* (Dressler, 2006). Thus, compounds have been for long mostly associated with the prototypical class of N + N, suggesting the greatest importance represented by this class.

The grammaticality of combinatorial possibilities reported in Tab. 1, shows some interesting patterns and can give us some important insights. Let’s consider for example the two categories P + N and * N + P. The ungrammaticality of *N + P (e.g., **passaggiosotto*, lit.‘passage under’) opposed to the grammaticality of P + N (e.g., *sottopassaggio*, ‘subway’, lit.‘under passage’), seems not to be arbitrary. An interesting analogy can, in fact, be seen with Italian syntax in which a preposition cannot follow a noun, but can precede it. So combinatorial possibilities of compounding appear to be related with syntactic rules of the language.

Although compounding belongs to the domain of morphology it interacts with syntax. This is what could be normally expected in a language in which different linguistic domains are not completely separated one form each other. As we will se later (1.5), this syntax-morphology interaction can play an important role for Romance languages compounding (Di Sciullo, 2006; Di Sciullo, 2009).

1.4.2 Compound classification based on constituent relation

The most relevant linguistic classification of compounds is based on the type of relation that connects constituents and that connects constituents and the whole compound. The most ancient classification of compounds goes back to the Indian grammarian Panini (6^{th} century B.C.). Most of his terms are still used today, although several other classification methods have been suggested.

A traditional, and widespread, classification distinguishes between *root compounds* and *synthetic compounds*. The formers are compounds made by two roots (as the name suggests) and that specifically don't contain any word derived from a verb. Examples of root compounds are *sand box* or *post stamp*. *Synthetic compounds*, on the contrary, are compounds in which a constituent (usually the second) is a noun derived from a verb, as *sky scraper* and *dish washer*. The classification in *root* and *synthetic compounds* however is not applicable to all languages. Italian (as the other Romance languages) does not have compounds based on roots. Romance lexemes, beside the root, contain other elements (in Italian a thematic vowel). For example in the Italian compound *capobanda,* 'bandleader' (lit. 'head-band') the two words *capo* and *banda* are formed *by* the roots *cap*- and *band*- plus the thematic vowels *–o* and *–a* respectively (that carries syntactic information). Again in Romance languages in Romance languages the class of *synthetic compounds*, furthermore does not seems to exist. This classification, born for Anglophone compounds, presents several limitations in the applicability to other languages, and thus is very limited and almost abandoned.

Bisetto & Scalise (2005) have recently suggested a new classification, intended to overcome the limits of the past classifications. In their new classification there is an explicit focus on the notion *endocentricity* and *exocentricity* of compounds (see also par. 1.3), and also on the classification of traditionally neglected categories. This classification is graphically represented in fig .1

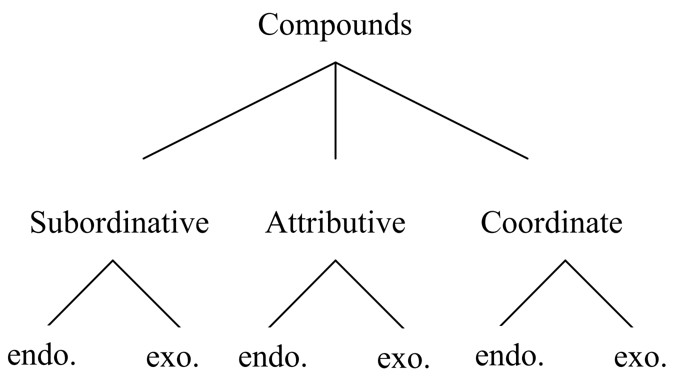


Fig. 1 Classification of compounds based on relational properties. From Bisetto and Scalise (2005).

According to this classification Compounds relation are classified in *Subordinate, Attributive,* and *Coordinate*.

Compounds are classified as *subordinate* when there is a "complement" relation between the two constituents. In the compound *dish washer*, *dish* is clearly the "complement" of the deverbal[5] constituent *washer*. An example, in Italian is the V + N compound *posacenere* ,'ash trasher' (lit. put-ashes') in which *cenere* is clearly subordinate to the verb constituent *posa*.

Compounds classified as *attributive* are formed either by a noun and an adjective (in which the adjective expresses a property and is a modifier of the noun) or by two nouns, where a constituent is used somehow metaphorically exspressing an attribute of the other constituent (specifically the *head* constituent, see par. 1.3). Thus, the compound *studio pilota* 'pilot study, lit. study pilot' is attributive, since one of the constituents modify the other by defining a property.

The last category, of *coordinate* compounds, includes compounds whose constituents are tied by the conjunction "and"[6]. Some examples are English *poet-painter* or Italian *attore regista* ,'actor-director'.

Each macro-category described is subdivided in two sub-categories according to the *endocentricity* and *exocentricity*, and hence on the possibility, or not, of identifying the head within the compound itself.

1.4.3 Compounds classification in strict and loose compounds

Another classification suggested for Italian compound (Scalise, 1992), distinguishes between *strict* and *loose* compounds. The labels of this classification refer to a typical surface characteristic that may be used to distinguish the compounds belonging to these classes. As described in par. 1.2 the formal structure of a compound can be described as in 12:

12. $[[\ ]_X \# [\ ]_Y]_Z$

---

[5] The term *Deverbal* indicate constructions are derived by appropriately suffixing verb bases.

[6] The conjunction usually is not "legible at the phonetic interface" (Di Sciullo, 2009), in other words it is not "visible", although is implied. E.g. *poet-painter → poet (AND) painter*

The hash mark # indicates the boundary between the words. This boundary, for old formed compounds or for compounds with high frequency of usage can weaken (formally # → +). Let's consider the compound in 13.

13.

a) gentile uomo → gentiluomo

b) [[gentile]$_A$ # [uomo]$_N$]$_N$ → [[gentil(e)]$_A$ + [uomo]$_N$]$_N$

The strong boundary in *gentile uomo* has been lost by a vowel deletion (the *–e in gentile*) and the two constituents have been "melted" together in *gentiluomo*, as formally represented in 13.b). Compounds that show this behaviour are called *strict compounds*. Other properties that characterize *strict compounds* are the non *componentiality* of their meaning (the meaning cannot be derived easily from the constituents, but see also par 1.4.4 for some considerations on componentiality), the fact that they generally have an ancient origin, and the *markedness*[7] of the constituent order (Scalise, 1994). In respect of the latter property, since the head of Italian compounds is expected to be on the left, the strict compounds, being mostly right headed (e. g., in *gentiluomo* the head is *uomo*) are marked. However the marked position of head seems not to be a characteristic that allows us to identify strict compounds. Many strict compounds are V + N, and thus *exocentric* (see par. 1.3) *porta + aerei → portaerei,* 'aircraft carrier' (lit. carry aircrafts'). Furthermore some Italian right headed compounds are not strict compounds. For example the Italian A+N compound *bassorilievo*, 'bas-relief' (lit. "low-relief) is a right headed word but present no amalgam and so is not a strict compounds Strict compounds, anyway, represent only a small set of Italian compounds.

Loose compounds are, on the contrary, the most common compound words, those generated by the main productive rules of compounding (as the example *documento chiave,* given before). They don't allow a phonological amalgam and the semantic relation between the constituents is transparent (1.4.4). The head in loose compounds is generally in the leftmost position , as in *uomo ragno,* 'spiderman, lit. man spider' where the head is *uomo*). However this is not true for all *loose compounds*. For

[7] The term *marked,* in Linguistics, indicates the less natural form of a linguistic construction

example the compound *video conferenza* ('video conference'), is a right headed loose compound (at list in term of surface characteristic).
The observations on headedess given so far, suggest that issue of headedness may represent a particular case in Italian, since we have so far provided examples of both right headed and left headed compounds. This aspect will be further analyzed in par 1.5., 1.6 and 1.7).

1.4.4 Compounds classification based on transparency of constituents

In paragraph 1.2, in relation to the criteria for identifying compound words we introduced the notion of *transparency*. We underlined that, although a criterion for identifying compounds is the possibility of establishing a semantic relationship between the constituents (see criterion 3.c)), this criterion is often not applicable.

The *transparency* of a compound can be defined as the degree of intelligibility of the compound. The more clear is the interpretation of a compound, the more it is said to be *transparent*, while the less clear is the interpretation, the more it is said to be *opaque*. As suggested by Libben, Gibson Yool and Sandra (2003), however transparency is best viewed as a property of the constituents of the compound rather then a property of the whole compound.
From this standing point it is possible distinguish several kinds of compounds, according to the analysis of transparency (Libben, 1998). All the examples in (14 from Libben, 2003) suggest a possible classification based on transparency:

14.
a) T - T (transparent – transparent), e.g. car-wash
b) O - T (opaque – transparent), e.g. strawberry
c) T - O (transparent – opaque), e.g. jailbird
d) O - O (opaque – opaque), e.g. hogwash

It is important to notice that constituents are not transparent or opaque by themselves, but only when considered within a given compound (compare *wash* in example 14.a) and 14.d)). Usually compounds with opaque constituents are words that reside in a language for long time, and the original meaning of their constituent can be no more accessible to the speaker. The Italian compound *caciocavallo* (that is a kind of cheese, lit. ‘cheese horse’) is a case of Italian T-O compounds. Apparently it is not possible find the relation between the second constituent *cavallo* and the meaning of the hole compound. However, the term *cavallo* presumably derives from the fact that during its preparation this type of cheese is left to dry by placing it *a cavallo*, ‘straddling’, upon a horizontal stick or branch.

Transparency is often associated with *productivity*. As already said in par. 1.2) they are intrinsically related: in all novel compounds the constituents must be, in order to make the compound intelligible, transparently related to the meaning of the whole construction. Another concept related to transparency, is that of *componentiality*, which indicates the possibility of deriving the meaning of the whole construction by the “sum” of its parts (this terms have been already been used to distinguish *loose* and *strict* compounds, see par 1.4.3). It is important to notice that a compound is never completely transparent and a certain degree of opacity, at least in terms of componentiality, it is always present.

15. uomo scimmia, ‘monkey man’ (lit. man monkey’)

The novel compound in 15 could be classified as a T-T compound (and according to the distinction in 1.4.3 as a loose compound). Even if both the constituents are transparent, several interpretations of the compound can be made (Gagné. & Spalding, 2006). A *uomo scimmia* could be a man hairy as a monkey, a man with the agility of a monkey, or a man whose facial traits resemble those of a monkey, a man with long arms similar to those of a monkey, a man who likes eating bananas etc. So contextual and pragmatic information must be taken into account in order to make a correct interpretation of a compound, and a certain degree of ambiguity is always present in the case of a fully transparent compound. Using a definition borrowed from psychology of Gestalt, compounds are word whose meaning is “more then the sum of their parts”. In conclusion, classification in 14 give us important suggestions on how different kinds of compounds may be differently represented in the mental lexicon. The notion of transparency, in fact, has been of central interest in Psycholinguistics theories, rather than in Linguistics (see par 2.4).

1.5 Compounding: between morphology and syntax

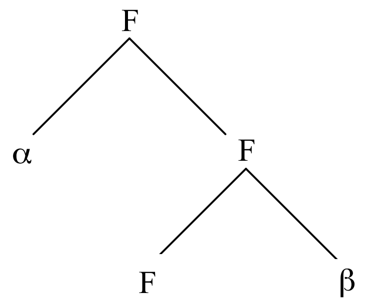
In par. 1.2 we have already underlined that although compounding is a domain of morphology, it also is influenced by syntax. This is not surprising because morphology and syntax are strictly related and interact in a language. The analogies between compounding and syntax, especially for Italian compounding, are however noteworthy.

We have already seen that some forms of compounding (P + N compounds, see par. 1.4.1) seem to reflect conventional syntactic order of words present in Italian, according to which a preposition must precede a noun. The same line of reasoning can be followed for the very productive category of V + N compounds (e.g. *cantastorie,* see Tab. 1). Although N + V construction are also present in Italian (e.g. *manomettere*, see Tab. 1), these are not productive and presumably they originate from Latin or Greek forms (Scalise, 1994, see also par. 1.6). The V + N category is instead really widespread in Italian and very productive. The preferred order of V + N again seems not to be casual. Italian syntactic canonical order is in fact S(ubject), V(erb), O(bject), as in *Valeria lava i piatti* 'Valeria washes the dishes' . The order is reflected in the V + N compound *lavapiatti* ,'dish washer' (lit. 'wash dishes'). In V + N compounds the verb and the noun thus behave as the V and the O in a syntactic SVO construction. The same doesn't happen in English, that is another language in which the syntactic canonical order is still SOV. While a phrase would be *Valeria washes the dishes,* the "corresponding" compound is *dishwasher,* in which the order of constituents doesn't reflect the syntactic order.

Is the relation between syntax and compounding stronger in Italian (or in other Romances languages) compared to English? There are different interpretations of this phenomenon. According to Scalise (1994), although there is a relation between syntactic order and order of the constituent in a compound, compounding shouldn't be considered as having a phrasal origin, but is a truly morphological process and these influences can be seen as collateral. A different perspective is suggested by Di Sciullo (2006, 2009). Before describing its theory is necessary to introduce some terms. Di Sciullo (2006) analyzes compounds within the framework of Asymmetric Morphology. Unlike in syntax, the order of the members in a compound (as in any morphological complex word) cannot be changed without changing the meaning of the compound itself (*e.g. core engine* vs *engine core*). This reflects a fundamental property of morphologically complex words: the asymmetry of relations (Di Sciullo, 2005).

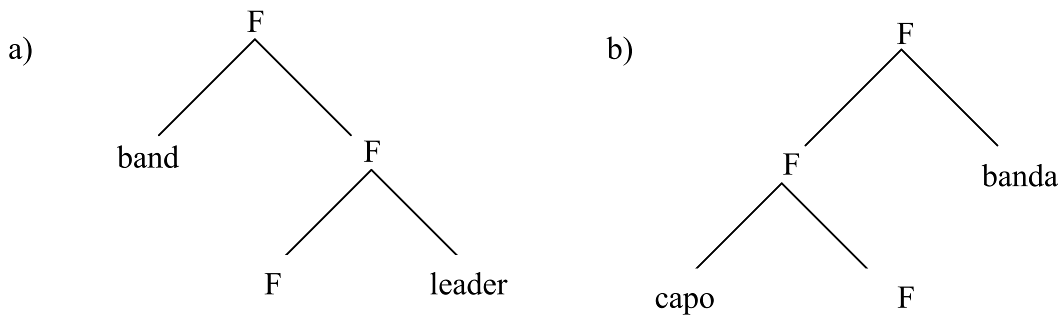
Di Sciullo (2006) claims that all compounds have an F-tree structure as represented in 16. F is the functional head while the other constituents of the compound may occupy the specifier position of the F-tree.

16.



Now let's compare the structure of the two compounds that are *cognates*[8], the English *bandleader* and the Italian *capobanda ,*'bandleader' ('lit. head band').

17.



The structure of 18a) and 18b) are one the flipped version of the other. Basically, according to Di Sciullo this is related to the different origin of compound in the two languages. The conclusions are here clear-cut: In German languages (like German, English, Dutch) compounds are generated from the *Morphological Plane* (DM), while

---

8 With the term *cognates* here we indicate translation equivalent compounds, both in the whole word meaning and in the meaning of the constituents but, potentially, with a different constituent order.

in Romance languages (Italian, French, Spanish etc.) compounds are derived, in part, from the *Syntactic Plane* (DS) and then transferred to the *Morphological Plane* (DM)[9]. Once transferred in the DM their internal structure is no longer accessible to the operations of DS, i.e. they become impermeable to syntactic rules.

Unlike Scalise (1994), that underlines the analogies between syntax and morphology in Italian compounds but that considers the two domain as separated, Di Sciullo thus suggests the possibility of a different origin (at list in part) for compound words, that is valid for Italian. However an important difference can be seen between the point of view of Di Sciullo and Scalise. Di Sciullo uses examples given by *syntactic phrases* (see also par. 1.2) as further argument that supports the syntactical origin of Romance compounds. She indeed argues that romance compounds may include a phrasal constituent and that this fact would be unexpected if these constructs were derived in the morphological plane DM. The fact that they include a phrasal constituent follows if they are derived in DS and then transferred to DM (Di Sciullo, 2006).

Scalise, although recognizing the similarity of compounds and syntactic phrases, does not consider them as belonging to the domain of compounding.

The theory of Di Sciullo, can be better understood by making a comparison between the compounding and derivation in correspondent Italian and English words.

18.

a) [[dark $]_A$ + ness $]_N$

b) [[oscur(o) $]_A$ + ità $]_N$

c) [[improvis(e) $]_V$ + ation$]_N$

d) [[improvvisa $]_V$ + zione$]_N$

e) [[key $]_N$ # [word $]_N$ $]_N$

f) [[parola $]_N$ # [chiave $]_N$ $]_N$

g) [[wind $]_N$ [mill $]_N$ $]_N$

h) [[mulino $]_N$ [a $]_{Prep}$ [vento $]_N$ $]_N$

In example 18a-d) it is possible to notice that all constructions (in Italian and English) are right headed (for a definition of head, see par. 1.3). It is the suffix that determines

[9] It has also been claimed that generally head position reflects syntactic order. English would be a counterexample being right headed and with a canonical SVO order. The right headedness of English compounds could be explained in diachronic terms: English right-headed compounding is a remnant of an earlier dominant SOV word order of Old English (Scalise & Guevara, 2005).

the lexical categories of the whole word. For example in the word *darkness*, the root *dark* is an adjective, while the output category is a noun. The head cannot be the root and thus should be the suffix -*ness*. This is confirmed by the analysis of other derived words sharing the same suffix. Indeed, all words ending in –*ness* are nouns. The same analysis is true for the Italian words ending in -*ità* (18a-d). But let's consider the last two examples (18.e) and 18.f)) that are *cognate* compound words in English and Italian. The order of the constituents is the reverse. In English 18.e) it "mimics" the structure of a morphologically complex word, with a modifier *key* on the left that projects a property on the head *word* (the rightmost constituent). On the contrary, in 18.f) the Italian word *parola chiave*, the order of words seems to violate the canonical position of the morphological head (at least compared to derived words), still being cogent with productivity in compounding that is recognized being mostly left-headed (Baroni, Pirrelli, Guevara, 2007). The examples in 18.g) and 18.h) further support this difference between English and Italian, the same referent (the wind mill) is a right headed compound in English, whereas in Italian is a *syntactic phrases* (or *prepositional compound*), an "atom" (see par. 1.2), but in which it is present a syntactic element (the preposition).

To summarize, it seems that while *compounding* in German languages belongs more clearly to the domain of morphology, in Romance languages it is a process between morphology and syntax. Even if Italian compounds (especially novel compounds) are generally with the head on the left the presence of the head in the rightmost position is possible (see 1.4.3) thus making the picture even more complicated. This double possibility of head position is also present in other Romance languages. French, for example, has many right headed compounds. A potential explanation of this phenomenon is given by the distinction of **centre** and **periphery** of word formation rules, the subject of the next paragraph.

1.6 Centre and periphery of word formation rules

The concepts of **centre** and **periphery** are very useful to explain the presence of some anomalies in word formation. Virtually in every language it is possible to identify, in word formation rules, a centre of regularities and a periphery of

irregularities. This periphery is made by the residuals from oldest languages or loans from other synchronic languages. Within this theoretical perspective strict compounds (e.g. *gentiluomo*, 'gentlemen') are part of the periphery of Italian compounding, while loose compounds (e.g. *parola chiave,* 'key word', lit. 'word key') are part the centre of regularities (see par. 1.4.3).
The origin of some compounds can thus be traced to language in which Italian has its root, like Latin and Greek

19.

a) manoscritto ('manuscript, lit. hand script')
b) psicologia ('psychology')

The example in 19.a) is a curious anomaly for Italian, since it is a right headed word. The origin of the word *manoscritto* is however from the Latin *manuscriptum*. In Latin compounds are right headed and this explain the anomaly for the Italian *manoscritto* which inherited the structure from the Latin word. The example of 19.b) posits other problems. Its structure is in fact at the crossroad between a prefixed word and a derived word. *Psico* is originated from Greek. Actually, neither *psico* nor *logia* are Italian words. Yet it is possible to understand the meaning of the elements that form the whole word: *psico* means 'mind' and *logia* 'discipline, from Greek *logos* lit. discourse'). So *psicologia* can be interpreted as "discipline that studies the mind". This is true for almost all words ending in *logia*. In Linguistics these elements can be defined as **semiwords**: they actually behave mostly as nouns, their meaning is highly consistent, but they are not truly words since they cannot appear as free form in the language (e.g. **bio, *psico*). Compounds formed by semiwords from Latin or Greek are common in many languages and are called **neoclassical compounds**. Neoclassical compounds are very common as scientific terms and shows a certain degree of recursion (e.g. the recent *psychoneuroendocrinology*).

Other compounds generated in the periphery of compounding are **traces** from other languages. Italian is full of traces from English.

20.

a) scuola bus (‘school bus’)
b) videogioco (‘videogame, lit. video-game’)

So the existence of right headed compounds in languages as Italian or French (see par. 1.5) can be easily explained by making a distinction between the main core word formation rules (called the *centre*) and other, less relevant, word formation rules, (called *periphery*). Right headed Italian (or French) compound words would belong to the *periphery*, while left headed, genereated with the synchronic rule of word formation (Scalise, 1994) would belong to the *centre*.
However, in the next paragraph we will see, that dicothomy *centre* vs *periphery* could be an excessive simplification. By analyzing some characteristics of Italian N + N compounds we will see that they show an interesting behaviour that deserve more attention.

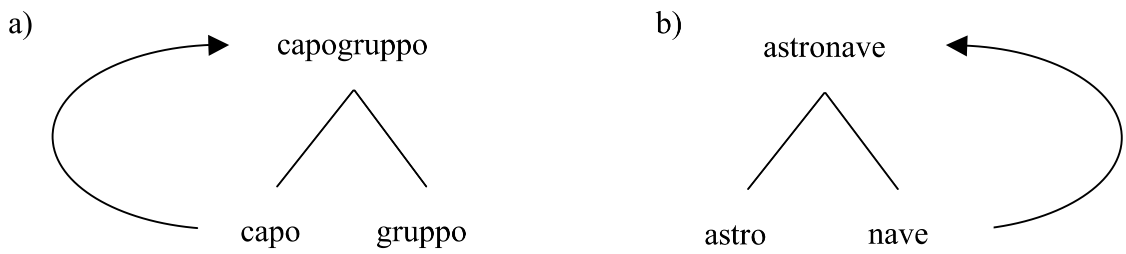
1.7 Italian Noun-Noun compounds and the ambiguity of headedness

We already have described the Noun – Noun compounds (N + N) as the compounds formed by two nouns as *sala stampa* (‘press room, lit. room press). We also have underlined that is well known the preference of this kind of compounds in many languages (see par. 1.4.1). Italian Noun – Noun compounds have been largely under the focus of Linguists since they offer an interesting opportunity for studying how conceptual combinations between two nouns happens. However, under a Linguistic perspective, it has been long neglected the peculiar characteristic of Italian compounds of being both left and right headed (the role of head and position, on the contrary, has been of crucial interest in the field of psycholinguistics, see par 2.4). As said in the previous paragraph, the phenomenon of right headedness of some Italian compounds could be ascribed to the periphery of word formation rules (Scalise, 1994). This periphery, however, could be more important than its label suggests. Let’s consider for example the words in 21.

21.

a) videochiamata ('video call')
b) fangoterapia ('mud therapy')
c) radiocronaca ('radio)

All compounds in 21 are legal Italian compounds with right head and all of them represent quite recent compounds. Italian has several right headed Noun-Noun compounds. Although many right headed compounds are neoclassical compounds (see par1.6 or *strict compounds* 1.4.3) novel (loose) right headed compounds are continuously coined. We can speculate that this phenomenon is due to the increasing influence from English (that is right headed). Many technical terms come from English and, in analogy on word formation rules of English, many Italian compounds are right headed.
So, the situation of headedness in Italian Noun-Noun compounds can be rapresented as in 22.
22.



With structure in 22a) representing the percolation (see par. 1.3) in left headed compounds and in 22b) representing the percolation in right headed compounds.
It could be argued that in Italian some words like *video* could behave like and affixoid. This however doesn't seem possible, since, unlike affixes the word video can occupy the initial or the final position of a word (*videochiamata, stazione video*). The word *video* itself can even appear freely in a phrase *per favore, guarda quel video*, 'please watch that video'. The same is true for the word *radio*.
Some words seem even to prefer right headedness in compounding (e.g. *terapia* in *fangoterapia*, *aromaterapia*, *musicoterapia*).

In reference to the theory of Di Sciullo, Italian right headed compounds represents an interesting anomaly. In a left headed compound a direct analogy between the structure of the compound and the syntax can be seen, e.g *capo*$_1$*banda*$_2$ is in prefect analogy with *capo*$_1$ *della banda*$_2$ ('head of the band'), the same is not true for right headed compounds as *fango*$_1$*terapia*$_2$, whose phrasal correspondent may be hypothesized being something like *terapia*$_2$ *basata sull'uso di fango*$_1$ 'therapy based on the use of mud'. In the latter condition (right headed) constituent order doesn't reflect the canonical order of Italian syntax. Italian right-headed compounds, behaving as compounds of other languages (such as Greek, Latin or English), thus behave more like suffixed words generated by derivation: the first term act as a modifier of the head, the rightmost constituent (almost all Italian suffixed words are right headed, see par. 1.5). According to the theory of Di Sciullo (2009) it is possible to speculate that Italian left-haeded Noun-Noun compounds are generated in the *Syntactic Plane* (DS), while the right headed compounds are generated in *Morphological Plane* (DM). Under a Linguistic perspective, these consideration should be further verified, but this not an aim of this work. Finally it is important to underline that all the considerations given don't want to contradict what argued by Scalise (1992) and Baroni et al. (2007), that Italian compounding productivity is left headed. Italian compounding productivity is beyond doubt (and scientifically proven, Scalise, 1992) mostly left headed.
The examples in 21 suggest that the influences of *periphery* of word formation rules can be more relevant than what is thought and that in some occasions, in Italian, a right headed productivity might be even preferred to the canonical one.

Summary

In this chapter we have first underlined the fundamental role of compounding in the morphology of almost every language (par 1.1). We have given a definition of compound and compounding in formal terms and we have analyzed the criteria that allow to distinguish compounds from non-compound structures (par. 1.2). We have defined what is the "head" of a compound, we have seen that the head may occupy different positions and that in some compounds cannot be identified as one of the constituents (par. 1.3). Several possible classifications of compounds have been provided, according to different characteristics (lexical, semantic and formal) of the compound and of the constituents of compound. (par. 1.4). However, Italian compounding seems to show a strict relation between syntax and morphology (par. 1.5). After defining that in the word formation rule there is a centre of regularities and a periphery of irregularities (par. 1.6) we have further analyzed the important class of Noun-Noun compounds, and we have seen that they may be both right and left-headed. We have finally argued that the peripheral phenomenon of right headed Noun-Noun compounds may indeed play a crucial role in the productivity of Italian compounding.

Do the properties of a compound analyzed in this chapter influence the representation of the compounds in the mind? Do these properties influence the online processing of the compounds by the cognitive system? How compounds are represented and processed by the brain? Giving an answer to these questions is one of the aims of psycholinguistics and Neurolinguistics studies on compounds, the subject of the next chapter.

2. COMPOUNDS IN LINGUISTICS AND NEUROLINGUISTICS

2.1 Compounds under a "user-centered" perspective

In the previous chapter, compounds have been analyzed under a Linguistic perspective. Several crucial concepts have been introduced such as the morphological head of a compound (par 1.3) and the interplay between syntax and morphology in compounds formation (par 1.5). A special attention has been devoted to the special case of Italian Noun-Noun compounds, that in Italian can be right- or left-headed.

The theoretical framework given by Linguistics, focuses is on the language itself and, mostly, on the *competence*, that is the potentiality related to language usage (Chomsky, 1965). A different perspective is that offered by Psycholinguistics and Neurolinguistics. These two disciplines focus on linguistic *performances*, how language users actually use the language (and not their potentiality).

Psycholinguistics investigates how language is organized in the mind, and what cognitive processes are involved in its representation and processing. Differently from linguistics, one of the main concerns of psycholinguistics is determining the sequence (and the timing) of events that characterizes the elaboration of linguistic stimuli. **Neurolinguistics**, that in some of its facets can be seen as a development psycholinguistics, focuses on the biological bases and correlates of language processing and representation, and on the study of the brain/mind/language interface (Fromkin, 2007). Within the field of Neurolinguistics we will focus on Neuropsychology and Neuroimaging studies of language.

Neuropsychology investigates, in general, the cognitive breakdowns that may follow brain damage. *Neuropsychology of language* focuses on linguistics deficits that may follows brain damage. There are two main reasons that make the neurpsychological approach fundamental in the study of language. First of all, it can give use information about the neural substrates that subserve a cognitive process (specifically, that are *necessary* to accomplish a cognitive task). Secondly, and most importantly, the observation of behaviour of disrupted normal flow of cognitive process, allows to study otherwise invisible phenomena and makes achievable getting insights on the "black box" functioning (Semenza & Mondini, 2006). Furthermore, it allows the study of traditionally neglected linguistic task, that in normal people are problematic for scientific inquiry, such as repetition or writing.

The **Neuroimaging studies of language**, make use of techniques that record the activity of the brain (EEG, MEG, PET, fMRI) in order to investigate the neural correlates of language processing. The coregistration of the neural activity during the accomplishment of language tasks can give us fundamental information on two sides: it may suggest different kind of processing associated with different stimuli (as reflected by different pattern of activation, or by different electrophysiological components); it may give information on what brain areas (or network) may subserve a specific language process.

Psycholinguistic and Neurolinguistics (with both Neuropsychological and Neuroimaging branches) shouldn't be seen as two separated approaches, but as discipline in synergic interaction in giving us information on how mind use language, and how brain functioning underlie language use. Summarizing, in contrast to Linguistics, that is mainly focused not on what language users "can do" (*competence*), these approaches are focused on what language users "do" when they use language (*performance*).

2.2 The problem of storage and computation

Psycholinguistic studies on morphology investigates how morphologically complex words are processed by the cognitive system and how they are represented in lexicon, i. e. the mental vocabulary of a person. The psycholinguistic investigation of morphology can be conceived as the study of dynamics of two opposite forces, one that push towards the **computation** of linguistic material, and the other pushing towards the **storage** of the same material (Baayen, 2007). As pointed out by Libben (2006), for both monomorphemic words and sentences, the issue of storage and computation is not much of a question. With few exceptions, simple words are arbitrary associations of sound and meaning. All simple words must then be stored in the lexicon, since their meaning cannot be computed from their structure. For sentences is the opposite: most sentences that one encounters in the normal use of language are unique (and novel) events and, as such, must be computed via the elaboration of their components. For morphologically complex words, the situation is different. Let's consider, for example some Italian words ending with the derivational suffix *–ità.*

23.

a) oscurità → [[oscur-]$_A$+ ità]]$_N$ ('darkness')

b) tranquillità → [[tranquill-]$_A$+ ità]]$_N$ ('tranquillity')

c) felicità → [[felic-]$_A$+ ità]]$_N$ ('happiness')

In all examples in 23 word roots are adjectives the suffix -ità bear the same meaning and is the head of the construction (see par. 1.3 and 1.5). Although all words ending in *–ità* could be stored as whole, the lexicon could reasonable only the rule associated with the suffix *–ità*. Every time a word with this suffix is encountered the rule is applied and the morphologically complex word can be understood. The same could be hypothesized for all affixed words. But this hypothesis clearly doesn't tell the whole story. If we consider, for example, the word *pubblicità* ('commercial') it could be decomposed in the stem *pubblic-* ('lit. public), and the suffix *–ità,* but the meaning is not the one expected by simply the simple application of the rule. So, the word *pubblicità,* together with its peculiar meaning must be in some way stored. It could be argued that the regular forms are computed, while the irregular are stored (with regularity and irregularity referred either to the form and to the meaning). The investigation of the computation/storage issue by the study of derived words, however, presents some important limits. Firstly, it would necessarily require two different sets of stimuli[10] (regular vs irregular) (Baayen, Dijkstra, Schreuder, 1997; Mondini, Kehayia, Gillon, Arcara, Jarema, 2009). Furthermore every languages usually has only a limited set of affixes with an usually highly consistent meaning,

[10] However it's noteworthy underline that psycholinguistic studies suggest that, even within regular derived words, several variables interact in determining the way a word is represented in the lexicon and processed (Baayen, Schreuder, De Jong, Kroot, 2002)

thus studying this words is possible to get insight only on a limited facet of the storage/computation dilemma.

Given these considerations compound words represent an ideal domain of investigation. They share characteristics both with sentences (since they're formed by more than one word), and with words (since they have a unique denotation). We have already seen this in chapter 1 par. 1.5 compounds, especially in some languages as Italian seem at the boundary between morphology (that studies multimorphemic units) and syntax (that studies sentences, made by more than one word). Compounds thus, "[…] offer a unique opportunity to study the interplay between computation and storage in the mind, the manner in which morphological and semantic factors impact the nature of storage the manner in which the computational processes serve the demands of on-line language comprehension and production" (Libben, 2006, pp. 3). On one hand compounds need to be easily segmentable in their constituent morphemes, just as sentences need to be segmentable in their constituent words. This is fundamental in order to understand novel compounds (e.g. *monkey man,* see also par. 1.4.4). On the other hand the compound sequences as a whole must be stored in memory. This is fundamental in order to access to all idiosyncratic traits associated to the words (e.g., the fact that *Spiderman,* is not simply a 'man' with some traits of 'spider', but that is a superhero with powers of spider, that is a character of comics and movies, etc.). Hence, the psycholingusitic study of compounds, beside giving us information on compounds, can give us more general information on how language constraints interact and determine how the mind deal with representation and process of language.

Are compounds listed as whole units or they are accessed via their constituents? What variables influence the way in which a compound is represented in the lexicon and processed? In the next paragraph, we will make a brief review of all theories and studies that have addressed this problem in general for morphologically complex words. In paragraph 2.4 we will make a review of all relevant psycholinguistic studies that have addressed these questions.

2.3 Mental representation and processing of compounds

The storage/computation problem is mirrored in the two main (and opposite) theories on mental representation of morphologically complex words.

On one side, according to **full listing** theories (Butterworth, 1983; Bybee,1995) all morphologically complex words are listed as whole units in the lexicon.

On the other side, according **full parsing** theories (McKinnon, Allen, & Osterhout, 2003; Taft, 2004; Taft & Forster, 1976), all morphologically complex theories are accessed via their constituents. Both these two theories obviously need to be complemented with the possibility of exceptions. A *full listing* theory cannot explain, alone, how the cognitive system understand new morphologically complex words (as novel compounds) and this theory cannot deal with so called “agglutinating” languages (like Turkish or Finnish) in which morphological productivity is very relevant. A *full parsing* theory alone, on the other side, cannot explain how the idiosyncratic meaning of some morphologically complex words can be accessed (as for opaque compounds, see par.1.4.4), since their meaning cannot be derived from the meaning of the constituents.

Actually, studies from different fields (see 2.4, 2.5) converge in suggesting a dual representation of morphologically complex words, with both a whole word representation and the possibility of accessing a word via its constituents. Models that account for both the possibilities can be divided in two major families: *Dual route* models and *late decomposition* models.

In **Dual route** models (Caramazza, Laudanna, & Romani, 1988; Schreuder & Baayen, 1995), every time a morphologically complex word needs to be processed both the routes (decomposition and whole word) are activated in parallel in a sort of “horse race”. Depending on several word variables (The frequency of usage, for example) one route will “win” over the other, determining the way the word would be processed.

In **Late decomposition models** (Giraudo & Graigner, 2000, 2001), instead, the first activation is always on the whole word representation. The complex word is decomposed in its constituents only under some condition, for example if the whole word representation is not available, or if the relation between of whole word and its constituents is transparent (see also par. 2.4.2 for different account of morphological processing locus).

Together with the already described, mainly popular, models of morphologically complex word processing, other models have been suggested.

Connectionist models of morphology have the peculiar characteristic of not assuming assume any abstract morphological structure. “Morphology” would emerge as a consequence of the consistency that morphologically complex words bring to the mapping between orthography and meaning[11]. A connectionist model of compound processing with these very characteristics has been suggested by Plaut & Gonnermann (2000). Within this model the potential differences between languages with different language complexity is included. According to the model, morphological “rich” languages (as Semitic languages) would mostly lead to morphological parsing more often compared to morphological “poor” languages like English.

Further models have been suggested: Kuperman, Bertram & Baayen, (2008) has recently described a **multi-route** interactive model of complex word processing. In their model, Kuperman et al. (2008) suggest that several sources of information interact in word access (see par.**Error! Reference source not found.**). Duñabeita, J.A., Perea, M., Carreiras, M., (2007), suggested that an **activation-verification** could explain how compounds are processed. Within this framework compounds are processed in a serial fashion, with the first constituent that activates several candidates and the second constituent allow the verification of what is the correct candidate among the ones previously activated.

All model discussed so far, mainly focused on compound (or complex words in genral) recognition, especially in reading.

Levelt, Roelofs, & Meyer (1999) developed a theory on word production, that is worth mentioning, since this theory is often claimed, especially in neuropsychological studies on compounds (see par. 2.5). According to Levelt et al. (1999), word production involves three main sequential stages. In the *conceptual preparation*, an activation of meaningful lexical concept occurs. This stage is followed by the *lemma* stage in which a lexical representation that contains syntactic and semantic features of the word occurs. This lemma stage is followed by another stage in which the phonological (or orthographic) form of the word is retrieved (the *word form* stage). According to Levelt et al. (1999) formulation, compound words would be represented as a single lemma that activates multiple word form representations. Some authors,

---

[11] A connectionistic approach of morphology has been contested by Pinker (1999) that argues that pattern regularities cannot explain all word formation phenomena.

however, claimed the existence of a multiple lemma (and multiple word form) representation for compounds (Mondini, Luzzatti, Saletta, Allamano, Semenza, 2002).

.

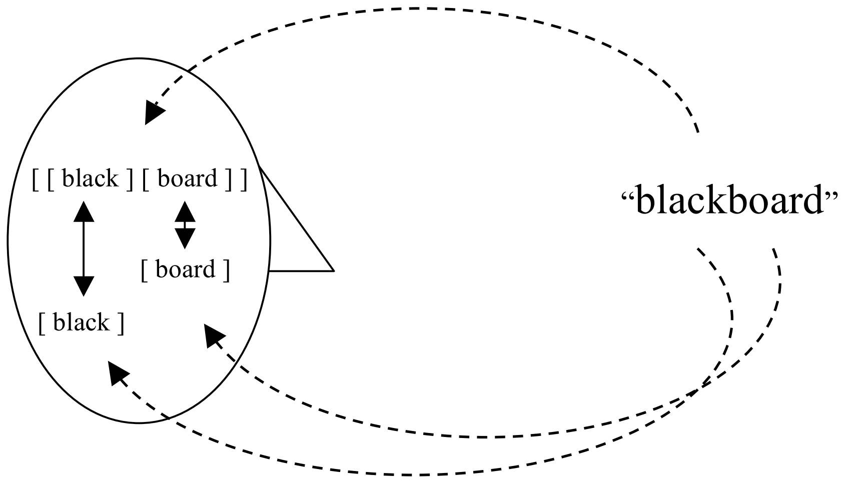


Figure 2.1 Parallel constituent and whole word processing (adapted from Libben, 2006). In the figure a schematic representation of dual route processing is given. The two routes compete and interact for the activation of a word. A structured whole word representation is connected with its constituents.

Although the way compounds (and complex words in general) are represented in the lexicon and processed it's still highly debated, several of the modes introduced claim the existence of more then one available route for morphological processing. This could seem the less economic choice for our cognitive system since neither the storage, nor the computational efficiency is optimized. The advantage in relying on this multiple mechanism is two-fold: firstly it maximizes the *opportunities* for a successful processing, and secondly it saves the cognitive system from making any choice on how processing words, choosing one route or the other (Libben, 2006). To accomplish the demanding task of a fast word recognition of everyday language use, all available information available are activated. For example, according to *dual route* model of morphological processing, if the compound is frequently used, a whole word form will be accessed faster, if the compound is encountered for the first time (or is not frequent) a access via its constituents will be made (see Figure 2.1).

Concluding, within the dominating framework of a multiple route processing, the focus of research on morphology has been shifted: the main issue is no more addressing if morphologically complex words are represented in the lexicon as whole units or as decomposed units, but addressing *when* a representation is preferred to the other (and hence, what variables influences the type of representation), and *when* a kind of processing, i.e. decomposition or whole-word access, is preferred to the other (and hence, what variables influence the type of processing). Specifically for compounds, the interest has been focused also in establishing what information are stored within the representation of the compound, i.e. the order of the constituents, the lexical category of the constituents and the information about the head.

2.4 Psycholinguistic studies on compounds

In the following paragraph the principal studies on compound processing will be described. Studies will be divided according to the methodology employed.

2.4.1 Simple lexical decision

The study by Taft and Forster (1976) was the first to investigate the mental representation of compound words by the using of lexical decision task. In their experiments, it was found that compound nonwords whose first constituent was a word (e.g., *dustworth*, *footmilge*) took longer to classify as nonwords than compound nonwords whose first constituent was not a word (e.g., *trowbreak*, *mowdflisk*). The presence of a word in the second constituent position appeared to be irrelevant. The results obtained were interpreted as evidence of a prelexical effect of morphological structure, and thus consistent with a *full parsing* model of word processing. Furthermore a primacy of the first constituent in compound access was hypothesized.

A subsequent study on compounds by Andrews (1986) manipulated the constituent frequency of constituents in a lexical decision task. Higher frequency of the first or the second constituent correlated negatively with response time for the lexical decision on the whole compound, thus supporting (at least in part) Taft and Forster (1976) full parsing account. The experiments carried by Andrews (1986), however, raised some doubts about a methodologically bias that could have influenced the results. In one experiments of her study, the effects with derived words were significant only in when compound words were included within the stimulus set, leading the author to conclude that decompositional effects, including compound constituent effects, could be controlled rather than automatic, and related to strategic aspects elicited by the task.

Thus, in these first studies on compound proceesing, headedness seemed not to play a crucial role in lexical acces. However in a more recent study by Juhasz, Starr, Inhoff, & Placke (2003, Experiment 1) manipulation of constituent frequency frequency lead to different results, and a prominent role of second constituent was found.

These results were recently confirmed in a study by Duñabeitia, Perea & Carreiras (2007). Differently from the studies on English previously described in this paragraph Duñabeitia et al. (2007) used Spanish and Basque. The authors were interested in study cross-linguistic differences in compounds processing, and specifically in the interaction between position and headedness. Spanish is in fact have mostly right headed compounds while Basque have mostly left-headed compounds. The authors found a frequency effect for the second constituent, both in Spanish and in Basque, and concluded that the lexical access for compounds may happen by a language-independent and blind-to-semantics mechanism.

2.4.2 Studies with priming

Many experiments on compound processing made use of priming paradigms, in order to avoid some intrinsic problems of simple lexical decision (Forster, 1998). In priming paradigm the target word is preceded by another stimulus that can be related (or not) with the follower. A priming effect in the related condition (i.e., a decrease in reaction times) is taken of index of interaction between prime and target, able to shed light on lexical processes of the target. Sandra (1990) carried out an experiment using a *semantic priming paradigm*. Transparent and opaque compounds might be preceded by a word semantically related to one constituent (e.g. *moon-sunset*, or *moon-Sunday*). A significant priming effect was found with transparent, but not with opaque, stimuli. Sandra concluded that automatic access to semantic representation of constituent does not occur with opaque compounds, thus challenging full parsing theories on lexical access. Analogous results, leading to similar conclusions, were found by Zwitserlood (1994), in Dutch.

In the *constituent priming* paradigm, the first or the second constituent of the compounds, are used as primes, in the related condition (e.g., *book-bookstore, store-bookstore*). Studies with constituent priming in Greek and Polish (Kehayia, Jarema, Tsapkini, Perlak, Ralli, Kadzielawa, 1999) suggested an activation of both the constituents of the compounds, but with a primacy of the first constituent. In another cross-linguistic study Jarema, Busson, Nikolova, Tsapkini and Libben (1999) used the constituent priming with French (that have mostly left headed compounds, see par. 1.5) and Bulgarian compounds (only right headed). Their results suggested a complex interaction between headedness, position and transparency, suggesting that some features of compound representation and processing might be language specific. The issue of transparency was investigated under a different perspective from Libben, Gibson, Yoon et al. (2003). In one experiment of their study, they took into account the transparency, classified as a property of the constituents rather then of the whole compounds (the classification is reported in example 14 , par. 1.4.4). Their results showed priming for both constituents and a crucial role of the rightmost constituent, the head for English compounds.

A very recent work with constituent priming of Italian compounds (Marelli, Crepaldi & Luzzatti, 2009a) suggested again a greatest importance of the rightmost head constituent. In two experiments, Marelli et al. (2009b) compared constituent priming with Noun-Noun (left headed and right headed) and V-N compounds (exocentric) (see par.1.5 and 1.7). In V-N and left headed N-N they found equal priming effects for both constituents, while for right headed N-N compounds they found a priming effects when the prime was the second constituent (hence the head). Their results were interpreted claiming a flat representation for V-N and left headed N-N compounds, and a hierarchical representation for right headed N-N compounds. Results described so far, show thus inconsistencies. Some studies (Jarema et al., 1999) suggest an interaction between position and headedness that may be language related. Other suggested language-indipendent mechanism for lexical acess Duñabeitia, Perea, Carreiras, 2007). Even within the same language (English) some studies argue for a most important role of the first constituent (Taft & Forster, 1976; Jarema et al., 1999), while other suggest a priority role of the second constituents (Juahsz et a., 2003; Libben et al., 2003, Marelli et al., 2009a). Almost all studies, however, suggest that, at least for transparent compounds, individual constituent are accessed during lexical processing.

Constituent priming experiments discussed so far, employed a time duration of prime that was always long enough to make possible a conscious perception of the prime.

One of the main critics moved to the studies that employ a visible prime is that results may be influenced by explicit, or implicit, expectation of the subjects. In these paradigms usually the prime duration is quite short, around 100 ms, thus making implausible the explicit use of strategies in the lexical decision (e.g. "when I see a word that is included within a compound, probably the target word will be a compound"). Subjects, however, would inevitably notice the relation between the constituent and the target in the related condition (e.g., *grass – grasshopper*) and they could develop beliefs about the nature of the task and the expectations of the experimenter. These beliefs could influence the way stimuli are processed (for example inducing a decompositional strategies for word processing). Some similar problems have been raised for simple lexical decision task (Andrews, 1986).

The main problems associated with *visible priming* are avoided with the use of *masked priming* paradigms (Forster & Davis, 1984), in which the prime is presented for very short durations (around 50 ms) and is virtually invisible to the subjects. Priming effects obtained in a masked priming experiment cannot be due to strategic influences, since subject don't notice the presence of the prime, thus making the subject blind to the nature of the experiment and its objectives. Usually the effects obtained in these experiments are interpreted as reflecting very early process of word processing. Shoolman and Andrews (2003) carried out an experiment analogous to constituent priming (e.g. *man-mankind*), but with masked primes. They found priming effects for both the first and the second constituent for opaque and transparent compounds, thus suggesting an early morphological segmentation based on morphemic structure, irrespective to the semantic transparency. Duñabeita, Laka, Perea and Carreiras (2009), in an experiment in Basque used compounds as prime and other compounds (that could share, or not a constituent with the prime) as target. They a significant priming effect when prime and target shared a constituent, regardless of position.

Fiorentino & Fund-Reznicek (2008) used a different experimental approach to explore compound decomposition. In their experiment compound words were used as masked primes, while individual constituents were the target (e.g *teacup – tea*). Within this design, any expectation of the subject is avoided, since there is no awarness about the focus of the experiment. Results by Fiorentino & Fund-Reznicek (2008), showed again robust priming effects irrespective of compound transparency and constituent position of compounds. Being the prime presented very shortly, an early decomposition of compounds based on the morphological structure was again supported, in favour to *full parsing* models of lexical access. Results with masked priming of compounds by Fiorentino & Fund-Reznicek (2008) are along the line of the popular theory on *blind decomposition* of morphologically complex words (Longtin, Segui e Hallé, 2003, Longtin & Meunier, 2005). According to the *blind decomposition* framework an automatic parser would act every time morphologically plausible structure is encountered regardless of the real morphological status of the structure and of its semantic. The claim of a blind decomposition mostly comes from masked priming experiments with derived words that is worth mentioning. In a series of studies with English and French robust priming effects were found not only when thee is a real morphological relation between prime and target (e.g. *marker – mark*), but also when the relation is only apparent (*corner-corn*; Rastle, Davis, New, 2004) and if the complex word is not existent but made of a real suffix and a real stem (**sportation-sport*, Longtin and Meunier, 2005). A priming effect, on the contrary, is never found when the overlap is only formal, i.e. when the ending segment of prime is not an existing morphological suffix (e.g. *brothel – broth*). All these conclusions support a ***sublexical morphological processing***: access to morphological structure and decomposition occurs before the activation of the lexical representation of words.

However, these results were not always confirmed and some alternative explanation has been suggested. With the same experimental design of Fiorentino & Fund-Reznicek (2008), Diependaele, Sandra & Grainger (2008) carried out in a series of experiment with Dutch compounds .The authors found significant priming effects with familiar compounds also when the bigram at the morpheme boundary was removed (e.g. *bookshop-book→ boo__hop-book*). This lead them to argue that in visual masked priming lexical influences *only* require a whole-word form representation of the prime to become sufficiently active relative to other lexical form activation. Indeed, some experimental results seem to support ***supralexical morphological processing*** account for morphological masked priming (in contrast to sublexical activation claimed in *blind decomposition*, see above) (Giraudo & Grainger 2001). According to Giraudo and Grainger (2001) morphological priming effects arise after a whole word activation occurs and not before.

In a recent paper Feldman, Connor, Moscoso del Prado Martín (2009) made a review on several papers that used the morphologically complex words as prime and the stem as target (e.g. *marker-mark, brother-broth*) and they ran an experiment with a better control of semantic characteristics of stimuli. Results of their experiment and review of literature, suggest indeed that a semantic effects is always present and that morpho-semantic aspects interact in early stage of lexical processing. Conclusion by Feldman are in line to conclusion by Diependaele, Sandra & Grainger (2009), that suggested an ***hybrid model of lexical processing***, according to which input words are mapped in parallel onto morpho-ortographic and morpho-semantic representations.

2.4.3 Eye-tracking studies

Eye-tracking studies it is a very valuable technique for investigating morphological complexity. Eye-tracking allow the study of *eye movements* during natural reading, while in traditional psycholinguistic studies, words are presented almost always in isolation, a very unusual condition for natural language usage. Eye-tracking studies, furthermore, warrant an high temporal resolution of data recording (just as other techniques as EEG, see par. 2.6) thus offering a fundamental way for studying the online temporal dynamics of language processing. Finally, eye movements are thought to be less influenceable by strategies that could arise during a psycholinguistic task.

Many eye movement studies have been used to investigate how compound words are processed during reading of English and Finnish compounds (Inhoff, Briihl, & Schwartz, 1996; Bertram & Hyönä, 2003; Juhasz et al., 2003; Andrews, Miller, & Rayner, 2004, Pollatsek & Hyönä, 2005, Kuperman, Bertram, Baayen, 2008). Results highlighted an effect of both constituent frequency and of whole-word frequency in reading measure, and influence of transparency, in support of *dual route* theories on lexical access.

A recent study by Kuperman et al. (2008) offered an interesting analysis of how morphological processing of compounds (and other morphologically complex words) could happen in reading. In the PROMISE model developed by the authors, information carried by every morpheme of a compound, and of the whole compounds are considered as a probabilistic source of information that are interactively used in a *multi-route* processing, that maximizes the opportunities for a correct lexical access (Libben, 2006; see par 2.3).

Summarizing, the main results from psycholinguistic studies on compounds suggest the following conclusions. During compound processing, constituents are accessed 2.4.1. All compounds, regardless of their semantic transparency, seems to be early decomposed in their constituents during lexical processing 2.4.2, 2.4.3. Several variables influence compound processing and decomposition in complex interactions. Taken together these results seem to support parallel dual route models of lexical access. However, of the conclusions discussed are still controversial. Some results haven't been replicated, and alternatives explanations for some aspects (e.g. influence of semantic) have been suggested.

2.5 Neuropsychological studies on compound words

The first studies of brain-damaged people on compound processing were carried out exclusively on patients with *aphasia*, the linguistic deficit syndrome that may follow a brain insult. These earlier studies were focused more on describing the performance of aphasic patients rather then on drawing inference about the normal processing from pathological behaviours (Semenza & Mondini, 2006). The actual neuropsychological studies on compounds (and languages in general) are very different. Studies are more often focused on understanding the architecture of language in the mind and in the brain, rather than on the aphasic characteristic. Moreover, studies of linguistic performance are no more confined only to patient with primary linguistic, but also extended to other patients, whose deficits allow to study the interplay of language processing and more peripheral aspects of cognitive functioning such as attention or short term memory (see Experiment 3 and 4).

A really consistent phenomenon reported in neuropsychological literature on compounds is the so-called "compound effect" (Semenza & Mondini, *in press*;

Semenza, Butterworth, Panzeri & Hittmair-Delazer, 1992; Hittmair-Delazer 1994; Blanken, 2000; Badecker, 2001). In picture naming tasks, when an error is made and the target word is a compound, aphasics tend to substitute compounds with other compounds (e.g., German *zuckerdose*, 'sugar jar' instead of *salzstreuer*, 'salt shaker'). The same effect is not found in non-compound monomorphemic words that present an embedded word (e.g., the embedded segment *pen* contatined in *penguin* is not substituted). These results strongly suggest that an implicit knowledge of the morphological structure of compound word is present and that this knowledge is separated from that of the phonological form of compound. These results are hardly explainable within a framework of full listing that make no distinction between monomorphemic and morphologically complex words

Results obtained by Semenza, Luzzatti and Carabelli (1997) suggest strikingly that morphological decomposition of compound does occur during lexical processing. The performance of a group of Italian Broca's aphasic patients in compound processing was investigated by the administration of a picture naming task that included V + N (Verb-Noun) compounds (e.g. *aspirapolvere*, 'vacuum cleaner'). Broca's aphasics usually have a more impaired performance in verbs compared to nouns. However (see par. and 1.2 and 1.4.1) in Italian V + N compounds are almost always nouns. If V + N were processed as nouns (and so just as whole words) no difference compared to other compounds or monomorphemic words would be expected. Broca's aphasics committed more errors in the verb constituents of the V - N compounds (usually omitting them). These results were more recently confirmed in a multiple-case study by Mondini, Luzzatti, Zonca, Pistarini, Semenza (2004). Altogether, these results suggest a decomposition of these types of compounds.

Another important insight come from the study by Blanken (2000), in which Broca's aphasics committed similar errors in opaque compounds, even more than in transparent compounds and suggesting the presence of a decomposed representation even for those compounds traditionally thought they were represented as a whole.

The study on gender agreement by Mondini, Jarema, Luzzatti, Burani, and Semenza (2002) suggest instead that the presence of a whole word representation of compounds might exist. The authors compared performance of two Italian agrammatic patients who showed a severe deficit in the production of the inflectional suffixes that indicating gender and number. The patients were asked to produce the correct inflectional suffix of A - N compounds (e.g., he was presented with *croce ross_* for *croce rossa,* 'red cross, lit. cross red'). Their performance in A - N compounds was compared to that with adjective noun pairs resembling compounds (*croce gialla,* yellow cross, lit. cross yellow' in place of *croce rossa*). An advantage for A - N compounds compared to noncompounds was found, thus supporting a whole word representation for these type of compounds.

A different pattern of results, however, was found with Italian *prepositional compounds* (Mondini et al., 2005) (e.g. *mulino a vento,* see par. 1.2 and 1.5). Even if awareness of compound structure was present, a specific difficulty in retrieval of the linking element (i.e. the preposition) was found, thus suggesting compositional process in compound retrieval combined to the availability of a whole word representation.

According to Levelt et al. (1999) theory of lexical production (par. 2.3) compounds would be represented a single-lemma and multiple word form structure. Several results discussed so far, seem to contradict this hypothesis. The "compound effects", as well as other results found in neuropsychological studies of compounds (e.g. Mondini et al., 2004, Mondini et al., 2005) are better explainable assuming that compounds are represented with a multiple lemma (and multiple word form) representation.

Several neuropsychological studies tried to address the issue on influence of positional and headedness effects in compound processing. Results, however, are not consistent (Semenza & Mondini, *in press*). A recent studies on bilingual aphasics by Jarema, Perlak and Semenza (2009) suggest that this lack of consistencies may depend on the idiosyncrasies of each patient that may shows impairment in different locus of lexical processing, and thus show different pattern of deficits.

Decomposition of compounds has been also claimed in writing. In the study by Badecker, Hillis & Caramazza (1990), patient DH performance was interpreted as a deficit of the *graphemic buffer*, the memory system in which abstract graphical representation are temporarily stored in writing process. DH showed an advantage in writing compounds compared to non compounds. Given these results, the authors argued that words are entered in the graphemic buffer decomposed in their constituent morphemes.

Summarizing, results from neuropsychological studies of compounds, mostly support dual route (or late stage selection) models of lexical access. In addition to the results from psycholinguistic studies they have suggested that compounds are represented in the lexicon (differently from monomorphemic words) as morphologically structured units (see Fig. 2.1) with available information about the lexical categories of constituents.

2.6 Neuroimaging studies on compounds

Within the big family of Neuroimaging studies of language, there are very different techniques with very different characteristics. The common denominator of all techniques is that of giving information about the brain activity (usually online) correlated to a linguistic task. Actually, studies on compounds are mainly interested on defining the temporal dynamics of their processing. This interesting is mirrored in the choice of the neuroimaging processing most suitable to reveal this information, i.e. Electroencephalography (EEG) and Magnetoencephalography (MEG) (Steinhauer, Connolly, 2008).

With both techniques it is possible to identify specific *components* of physiological signals, that are thought to be related with specific features of processing: semantic (Kutas, 2000), morphological (Lavric, Clapp & Rastle, 2007) or morphosyntactic (Friederici, 2004). The high temporal resolution of these techniques also allows making fine discrimination otherwise impossible with reaction time measures.

So far, only few studies investigated specifically, neural correlates of compound processing. Koester, Gunter, Wagner and Friederici (2004) conducted a series of experiments with German compounds auditorily presented to the subjects. In their experiments the gender-agreement between a determiner and the initial (non-head) compound constituent and between a determiner and the rightmost compound constituent (that in German is the head, e.g., *das Presse – amt,* 'the $_{NEUT}$ press$_{FEM}$ office$_{NEUT}$'). Although only the head is morphosyntactically relevant in German, both constituents elicited left anterior negativity (LAN) if the gender was incongruent. This finding, later replicated by Koester et al. (2007), was taken as an indication of automatic morphosyntactic decomposition. In an experiment with MEG (Magnetoencephalography) Fiorentino & Poeppel (2007) investigated the influence of constituents frequencies and of whole word frequency, in lexical access of compounds, by orthogonally manipulating these variables. Both behavioural and physiological results, suggested an early automatic parsing of compound words. In particular the M350 component of MEG (that can be conceived as a subcomponent of the N400 component, Pylkkänen and Marantz, 2003), was affected by the morphological status of target (compounds bs compounds) and taken as index of early morphological parsing.

In a study by Vergara-Martínez, Duñabeita, Laka and Carreiras (2009) with Basque compounds frequency of first and second constituent was orthogonally manipulated in a reading task. ERP showed pattern of activations interepreted by the authors in support to *activation-verification* theory on compound processing. A first negativity, arising around (100-300 ms) would reflect the activation of candidates triggered by the first constituent, while the following negativity (N400), would reflect the integration of information activated by the second constituent with that previously activated by the first one.

The study by El Yagoubi et al. (2008) investigated the role of headedness in Italian Noun-Noun compounds. As extensively discussed in the previous chapter, par 1.7, Italian Noun-Noun compounds may be both left headed (e.g. *capobanda*) and right headed (e.g. *astronave*). El Yagoubi et al. (2008) investigated this issue in a lexical decision task. As experimental stimuli they used compounds and non compounds . Compounds were presented in two conditions: with the constituents in normal order (e.g. $capo_1banda_2$) or in reversed order (e.g. *$banda_2capo_1$). Similar conditions were used with monomorphemic words, that were presented normally (e.g. $cocco_1drillo_2$) or with its halves in reversed position (e.g. *$drillo_2cocco_1$). An increasing negativity interpretable as the LAN, was found in compounds compared to noncompounds, suggesting a morphosyntactic processing of these stimuli. Furthermore, a modulation of P300 component related to headedness was found, with a greater P300 for right-headed compounds. The authors argued that this difference could be explained by the canonicity of left-headedness in Italian compounds compared to right-headedness. That is, the left component is “automatically” recognized as the head. When the second constituent is encountered, the information need to be updated and this would result in a P300 increase.

Studies on neural correlates of compound processing thus confirm an early activation of constituents and an early morphosyntactic parsing. The study on Italian Noun-Noun processing and headedness suggests that headedness may indeed play a role in processing. However, the issue of Noun-Noun processing is still not unquestioned. In the next chapter we will see some open issues that will be addressed in this thesis.

PART 2

3. AIM OF THE RESEARCH

The aim of the present research is the study of the representation and processing of Italian Noun-Noun compounds. In particular, Neuropsychological and Electrophysiological methods have been used to investigate how mind and brain elaborate this linguistic constructions.

Why study Italian Noun-Noun compounds? Mostly, the interest on these linguistic structures is related to the ambiguity of *head* position. In a given language, usually the *head* of a compound (the constituent that determines the lexical category and mostly the semantic traits of the compounds, see par. 1.3) falls always in the same position. This is obviously an advantage in term of interpretability. However, In Italian (as in other Romance languages), the head may be in the first or in the last position. How the cognitive system faces this ambiguity? For some cases this is not an actual problem: for example in compounds made by an adjective and a noun, the head is almost always the noun constituent[12]. In these cases it is possible to rely on a lexical analysis of the compound to identify the head (see par 1.3 criterion 7.a). In the cases of Noun-Noun compounds both constituents are potential candidates for being the head. Sometimes for identifying the head it is possible to rely on morphosyntactic features of the stimulus (see par. For example, in *astronave* the first constituent *astro* is singular masculine, while the second *nave* is singular feminine. The whole compound is singular and feminine, and thus the head must be the second constituent. However this strategy not always is helpful. For example in *capogruppo* both first (*capo*) and second (*gruppo*) constituent are singular masculine. The only way to identify the head in these cases is relying on semantic features (see par. 1.3, criterion 7.c) : a *capogruppo* inherit its semantic traits (through "percolation", see par 1.3), from *capo* that is [+Aninamate], in contrast to *gruppo* that is [-Animante]. Furthermore, the moderate productivity of Italian compounding (compared to that of English) allows investigating how the morphological structure of a language can

[12] Some Adjective - Nouns like *giallo limone* (lit. 'yellow lemon') represent a counterexample, since the adjective is the head. However this case is circumscribed only to colours, and thus represents just an exception easily understandable.

influence its lexical processing (see Plaute & Gonnerman, 2000 for a connectionist account of this effect). In par 1.3 the *head* of a compound has been described (among other features) as the constituents that contributes principally to determine the meaning of the compound.

The study of compound representation and processing, specifically on Italian Noun-Noun compounds, several questions left, that will be addressed in the present work.

Are left headed and right headed Noun-Noun compounds processed differently?

Only a few studies examined headedness in Italian compounds. In a study by Marelli et al. (2009a) different constituent priming effects were found for right-headed compounds compared to left-headed compounds. An interesting analogy was found for left headed Noun-Noun compounds and Verb-Noun compounds leading the authors to conclude that these compounds have a flat structure inherited by syntax, in contrast to right headed Noun-Noun compounds that have a hierarchical structure as truly morphological complex words.

In an ERP study by El Yagoubi et al. (2008), differences related to headedness in left- and Right headed Noun-Noun compounds were found in the P300 components, with an higher P300 for right-headed compounds, that required an higher processing cost. The authors concluded that right headed Noun-Noun compounds may be more difficult to be processed because of their unconventional structure, or because the information about head position must be updated during the lexical processing.

Both these studies thus suggest that left headed and right headed Noun-Noun compounds are processed differently, although different explanation of the phenomenon have been provided.

When exactly information on head occurs?

Study by El Yagoubi et al. found a different modulation of P300 component in left-headed and right headed Noun-Noun compounds. However the task employed did not allow to investigate if an early and automatic decomposition of Noun-Noun compounds occur and if information on headedness is early accessed.

"How" the information about head is encoded?

The *head* of a compound determines mainly the syntactic, lexical and semantic properties of the whole compounds (see par 1.3). Thus, it is of particular interest investigating how this complex information encoded in the lexicon and retrieved.

What's the role of syntax and morphology in compound processing?

The linguistic analysis made in chapter 1, outlined the important relation of syntax and morphology in Italian compounding, and how right-headed Noun-Noun compounds represent an anomaly if compared to left headed Noun-Noun compounds and Verb – Noun compounds. Study by Marelli et al. (2009a) suggested that this characteristic might indeed influence the way stimuli are represented and processed. This issue, however, needs to be further investigated.

How compound processing interacts with more peripheral cognitive process?

Only a few studies investigated how compound processing interact with non-linguistic processes (as attention) and peripheral cognitive structures (as buffers). Two studies of the present work will be specifically focused on this issue.

These are the main question we tried to address in the experiment presented in part 2.

4. STIMULI AND STATISTICAL ANALYSES

4.1 Stimuli description

All experiments reported in this dissertation included Italian left headed Noun-Noun compounds (NN1) and right headed Noun-Noun compounds (NN2). As explained in par. 1.3, there are several criteria that allow to identify the head of a compound. The morphological *Head* was identified by the application of the semantic and syntactic criteria. In few cases in which the two criteria did not converge and the semantic criterion is considered as prevailing. Thus, headedness was always identified on a linguistic basis and not, as seen in some studies (Inhoff et al., 1996, Vergara-Martínez et al., 2009) on a subject rating basis. As explained in par. 1.6 and 1.7, right headed Noun-Noun compounds are generated by the *periphery* of Italian word production rules. Within this periphery it is possible to find compounds inherited from ancient languages (Greek or Latin, e.g., *terremoto*, 'earthquake' comes probably from Latin *terrae moto*, 'movement of heart') or borrowed from modern languages (usually English, as *videogioco* 'videogame'). Right headed compounds used in the experiments of part 2 come from both the two sources. However, irrespectively of the origin, in every compound two existing words are always recognizable, and they are never made by *semiwords* (see par. 1.6) or contained amalgam (see par. 1.4.3). All compounds used in the experiment are transparent (although as pointed in par. 1.4.4 slightly different degree of transparency were recognizable). All compounds are **lexicalized** and written as a single word.

Thee groups of thirty subjects rated **Imageability**, **Age of Acquisition** and **Familiarity** of both whole compounds, constituents and embedded words (see experiments *Stimuli* section of each experiment for further details on stimulus types.). Ratings were given on a seven-point Likert scale, and each subject rated only one dimension (Imageability, Age of Acqusition of Familiarity). Summary Tables of values of psycholinguistic variables are reported in Appendix. Other variable considered were **Length, Frequency and Neighbourhood size.** Stimuli length was always calculated as number of letter. Stimuli Frequencies were collected from a online corpus of written Italian (http://dev.sslmit.unibo.it/corpora). Frequency values were referred to form frequency (ignoring case) and were always logarithmically

transformed before any statistical analysis in order to avoid undesired influences of data skewness.
Stimuli Neighbourhood size was calculated as number of words (in a corpus) sharing all letters but one with the word.

4.2 Psycholinguistic variables and statistical analyses

One of the main concerns of studies on compounds (and of language studies in general) is that of ruling out the influence of psycholinguistic variables that are not of primary interest of the study, but that are related with the variable of interest.
In most of the studies described in the previous paragraphs (2.4.1, 2.4.2) the influence of relevant psycholinguistic variables (e.g. frequency, length, etc.) was controlled by matching the stimuli for the variables or, when the influence of the pycholinguistic variables was the main goal of the study, they were orthogonally manipulated in a factorial design (e.g. Duñabeitia, Perea, Carreiras, 2007). Many studies, however, has recently embraced a difference approach by using regression analysis. In these approach, several psycholinguistic variables are taken into account by including them as predictors in multiple regression models.
The use of regression designs has two main advantages: it allows partialling out the effect of variables that couldn't be matched (Shoolman & Andrews, 2003); it allows to study several variables at the same time and to determine which one has the greatest effect. Furthermore, it allows to study the effect of variables as they actually appear "in nature". As argued by Cohen (1983) some dichotomization (for example high vs low frequency) represents indeed an artificial categorization of a phenomenon that is originally along a continuum.
The use of regression models allowed to discover that many other variables beside the more common length and frequency, may influence recognition time and hence lexical process of compounds: for example morphological family size, constituent family size, positional entropy of the constituent, etc. (De Jong, Feldman, Schreuder, Pastizzo, Baayen, 2002; Bien, Levelt & Baayen, 2002). Contributions of regression modeling are noteworthy: they underlined how, in compound processing, several

variables may interact in complex ways that cannot by captured by traditional analysis and factorial design. For example Bien, Levelt and Baayen (2002), in the analysis of joint data from four experiments, found compound frequency was a significant nonlinear predictor, with facilitation in the low-frequency range and a trend toward inhibition in the high-frequency range. Many of the study with eye movement, described in par.2.4.3 used regressions to evaluate the effect of several variables.

In the experiments of the present dissertation, where possible, data were analyzed through **multiple regression mixed-effects model** (hencefor "mixed models", Baayen, 2004, 2007). Mixed Models are really useful for psycholinguistic studies of language for several reasons, particularly in dealing with the "language as fixed effect fallacy" (Clark, 1973; Raaijmakers, Schrijnemakers, Gremmen, 1999). In order to understand the advantage of mixed effect modeling a brief summary of the issue of "language as fixed effect fallacy" is necessary. Subjects that participate to an experiment represent only a *random* sample of the population of all possible subjects and words used in a given experiment represent a *random* sample from a wider population of words[13]. Results from a traditional ANOVA, with only subject as random effect, broadly tell us if the same effect would be replicated with other subjects. Experiments that deal with language, however, are also interested if the observed effect is expected to be replicated with other linguistic stimuli belonging to the same category.

According to Clark (1976) it is possible to overcome this problem through the utilization of *minF'*. This value is calculated from two separate ANOVA, one with subjects as random effect (F_1) and one with words (or the given linguistic unit) as random effect (F_2). However since the 1980 a simpler approach has been adopted using only the results F_1 and F_2 (and no more calculating minF') and interpreting a result as significant only when results from both F values are significant (Raaijmakers et al., 1999). All these approach, however, has several limits (Baayen, 2004). First of all they deal limitedly with the problem of words as random variable. Moreover they are not compatible with more detailed statistical analyses, such as multiple regressions.

Baayen, in several papers (2002, 2004, 2007) underlined how mixed models represent a better alternative. They can deal with several random variables (subjects and words

[13] Obviously, both subject and words usually are not really *random* sample but, *pseudo-random* samples assumed to be *random*.

first of all, but also nested variables) and with both categorical and continuous variables. The advantage of mixed models in the present work is two-fold.

Firstly they allow partialling out the effect of unmatched psycholinguistic variables. Although stimuli were matched as much as possible, a satisfactory matching was not always possible. Mixed effects models allow to rule out explanation of the results related to imperfect matching of variable. For example, in a model that include both length and the category of words, the effect of category is calculated ruling out the effect of length.

Secondly, a multiple regression approach allows to study the influence of covariates, that may be interesting per se.

All analyses with mixed effect models regression give, as results, a series of coefficients (just as multiple regression) plus the estimated variances associated to the random effects (generally for Word and Subject). For every covariate a single coefficient is estimated, while for every factor one coefficient for every level of the variable is estimated (levels of categorical variables are recodified as dummy variables). If one or more categorical variables are present, one level for every variable is taken as the reference and it represents the *Intercept* of the model. The model is additive and the contribution of every coefficient must be added to obtain the final prediction for a given stimulus.

The final model was selected as follow. An initial model was fit on the data, including all covariates and all categorical variables. Variables whose effect had a $|t| < 1$, were excluded from the model one at time, starting with the variable with the lowest $|t|$. For categorical variables no effect was removed if it belonged to a factor in which at least one level had a $|t| > 1$. Once this procedure was over, the p-value associated to each coefficient was calculated. As pointed out by Baayen, Davidson and Bates (2008) traditional p-value in mixed effect model regression might be anticonservative. A more robust alternative is a p value obtained through Markov chain Monte Carlo (MCMC) simulations (hence pMCMC). The final model included only significant variables.

Mixed models were not used for analyzing ERP data, since no study used this approach in ERP and language, only recently suggested (Davidson, 2009).

All statistical analyses were performed using R, release 2.10.0 (R Development Core Team, 2007).

5. EXPERIMENT 1 - Masked Priming of Noun-Noun Compound Constituents: Neural Correlates of Early Access on Morphological Structure

5.1 Introduction

In chapter 2, it has been largely discussed the prominent role of headedness in psycholonguistic studies on compounds. Theories largely differ about the conclusion about this issue and the results varied depending on the methodology employed and on the language studied.

Results from experiment introduced in chapter 3 yet suggest that head may indeed play a role in morphological processing. However, the methodology employed didn't allow getting information about the timing of head access. Does information access on head occur early, or is it inferred later, after the access to whole compound occurs? As pointed in par. 2.4.1 and 2.4.2, the task employed could largely influence the way in which compounds are processed and in which every constituent contributes to lexical access. In constituent priming experiments an interaction of head position and language has been found (Jarema et al., 1999), with an advantage of the first constituent and of the head constituent, in synergic interaction. Studies that manipulated constituent frequency, on the contrary, seem to suggest a role of the second constituent, regardless of the head position (see par. 2.4.1). However, many studies suffer from the potential influence of other effects. Shoolman & Andrews (2003), for example, manipulated explicitly the "biasness" of task by contrasting different nonword conditions. To investigate strategic influences on morphological decomposition, a condition in which nonwords consisted of combinations of unrelated words (e.g. *toadwife*) and nonwords (unbiased context), was compared with one including a high proportion of word-word nonwords constructed from highly associated words (e.g., *fastslow*) or by reversing the constituents of real compound words (e.g. *budrose*) (biased context). Nonword stimuli influenced the performance: with the second condition (biased context) leading to greater reliance on decompositional processing. Shoolman & Andrews (e.g. thus concluded that some effects in Reaction Times could be influenced by decision-related processes rather than from lexical retrieval per se.

A way to avoid the problem of "triggered decomposition" is to use masked priming paradigm with compounds as primes and individual constituents as target. In this way, subjects are not aware on the nature of the task and the use of explicit or implicit strategies is ruled out. Furthermore, the short presentation of compound warrants us to study the lexical influences related to early stage of processing. The very same approach has been adopted by Fiorentino & Fund-Reznicek (2008) with English compounds. They found a priming effect when compounds primed either of their constituents (*teacup-tea; teacup-cup*), but not when the prime has just an orthographic overlap with the target (e.g. *penguin, pen*). Both opaque and transparent compounds primed equally their constituents, thus leading to the conclusion that semantics play no role in this early phase of processing, that is supposed be truly morphological. Similar results have been found by Diependaele et al. (2008), which found a significant priming effect of bimorphemic Ducth compounds, on both the first and the second constituent. Analogous effects were found both for familiar and unfamiliar compounds. These results support the dominant framework of *blind decomposition* in morphological processing (Longtin et al., 2003; Rastle et al., 2004; Gold & Rastle, 2007; see also par. 2.4.2). Blind decomposition has been theorized mainly on results from masked priming experiments that found a priming effect if a cogent morphological structure is encountered (even if only apparent). So both *player* and *corner* would be decomposed (even if *corner* is only apparently a morphologically complex word composed by the stem corn and the suffix *–er*). Behavioural results on blind decomposition have been recently confirmed in an ERP study by Lavric, Clapp & Rastle (2007). They replicated the study by Rastle et al., (2004) with the addition of ERP recordings. Three masked priming conditions were included: transparent (*player-play*), opaque (*corner-corn*), and formal (*brothel-broth*). Behavioural results of Rastle et al. (2004) were confirmed and a biphasic modulation of N400 was found. In an early windows, was modulated by all three priming condition, but transparent and opaque condition primed equally for all N400 duration. Results have been interpreted as supporting an early morphological parsing of words.

Blind decomposition theory, however, has been challenged. In a review of several studies Feldman et al. (2009) outlined how a "latent" semantic effect seems to be present in all experiments that claimed blind decomposition, thus suggesting that also semantics may play an important role in early visual word processing. According to this alternative point of view morphological effects could arise through the interplay of sublexical (morpho-orthographic) and supralexical (morpho-semantic) representations. In line with this theory, the experiments of Diependaele et al. (2008), for example, showed a significant priming effect with familiar compounds also when the bigram at the morpheme boundary was removed (e.g. *bookshop-book→ boo__hop-book*), suggesting the importance of whole-word representation.

Summarizing, results of masked priming studies of compounds have suggested that both an activation of single constituent and of whole word happen very early in lexical processing. The study of Italian Noun-Noun compounds can further shed light on this issue, given the relatively smaller compound productivity of Italian, compared to the other language, suggesting less decomposition (Plaut & Gonnermann, 2000). Studying the early effect of compound headedness can also shed further light on the role of morpho-orthographic and morpho-semantic in visual word processing. Following the same approach by Lavric et al. (2007), ERP can give fundamental information on how this early lexical processing unfold in time, and if neural correlates can suggest a semantic independent early morphological decomposition.

5.2 Materials and Procedure

5.2.1 Partecipants

Thirty students from the University of Padova, all right-handed native speakers of Italian, participated to the experiment. All subjects were right handed, they had no neurological pathologies, normal or corrected to normal vision and were paid for their participation. Three participants' data were excluded from the analysis because of excessive artifacts in the electroencephalogram (EEG). One partecipant was excluded because of the excessive number of errors in the task (18%).

The remaining 26 participants had a mean age of 25.15 years (range 19-32); 15 of them were women and 11 were men.

5.2.2 Materials

Experimental *word list* was composed by 336 prime-target pairs, with 168 pairs for the related condition, and 168 pairs for the unrelated condition. In the related condition, primes were the same of experiment by El Yagoubi et al. (2008) and of the experiment presented in chapter 3, with just few replacements, that did not affect overall matching yet observed between Noun-Noun compounds (see par. 7.2.2).

Compounds in the related condition belonged to two different categories: 28 left headed Noun-Noun compounds (**NN1**, e.g. *pescespada,* 'swordfish, lit. fishsword'), 28 right headed Noun-Noun compounds (**NN2**, e.g. *astronave,* 'spaceship, lit. starhip'). Two non-compound categories were included as orthographic control: non-compound with left embedded segment (**NC1**, e.g. *coccodrillo, '*crocodile', where *cocco,* 'coconut', is an Italian word, with no morphological neither semantic relation with *coccodrillo*) and non-compound with right embedded segment (**NC2**, e.g. *tartaruga,* 'tortoise', where *ruga,* 'wrinkle', is an Italian word, with neither morphological nor semantic relation with *tartaruga*). In related pairs, each compound could be followed with both its two constituents (*astronave-astro,* lit. 'starship-star'; *astronave-nave*, , lit. 'starship-ship), while each non-compound could be followed by its embedded word (*coccodrillo-cocco,* lit. 'crocodile-coconut'; *tartaruga-ruga,* lit. 'tortoise-wrinkle'). The other segments of NC1 and NC2 (**drillo* for *coccodrillo, *tarta* for *tartaruga*) were not included as target nonwords, in order to circumvent the possibility that Subjects would understand the nature of the experiment.

Each constituent of compounds and each embedded word was used to form the 168 unrelated pairs. Within this pairs unrelated Noun-Noun compounds (e.g. *nullatenente-astro*, lit. 'nothingholder-star') and unrelated word with an embedded segment primed respectively compound constituents and embedded word of noncompounds (*alabarda-cocco,* lit. 'halberd-coconut', with embedded segment *ala* 'wing'). Prime words in the unrelated condition were matched for length and frequency to words in the related condition. Fourty filler pairs, composed by a derived word followed by an unrelated word (e.g. *confederazione-gesto*, lit. ' confederation-gesture'), were added in order to reduce the overall prime-target relatedness to 0.20. The final number of pairs that required a *word* response in the lexical decision task was 336. In order to avoid the influence of any repetition effect, the 336 words of the *word list* were split in two subsets. In every subset every target word was presented only once. For example, the related pair *astronave-astro* and the unrelated pair *ossobuco-astro* were in different subsets. 168 nonwords pairs were made up by a compound or a derived word followed by a legal nonword (e.g. *segnalibro-*ritinie*).

Nonwords were created by changing with two or three letters of real words taken with length and frequency matched to that of word. All nonwords respected phonotactic rules of Italian. Each subject saw only one of the *word list* subsets, all fillers and the nonwords, for a total of 168 word targets and 168 nonword targets. Primes were real words in all conditions.

Several psycholinguistic variables were taken into account: Length, Frequency, Familiarity, Imageability, Age of Acquisition of both whole words and constituent or embedded segments. In overall analysis of whole words, stimuli differed in Age of Acquisition [$F(3,108)=3.11$, $p=0.03$], with NN2 acquired later in comparison to other categories. Length was also different, with compounds longer than noncompounds, but with no difference within the same morphological category [$F(3,108)=10.03$, $p<0.001$]. No difference were found for familiarity, imageability and frequency. Constituents and embedded word in first position were matched for all psycholinguistic variables. Constituent and embedded words in second position differed for Age of Acquisition [$F(2,76)=5.43$, $p=0.006$], Imageability [$F(2,76)=3.97$, $0=0.02$] and Length [$F(2,76)= 11.614$, $p<0.001$]. Right constituents of NN2 resulted acquired later than first constituent NN1 and embedded word NC1. Both NN1 and NN2 resulted longer than NC1 embedded word. The effect found for Imageability was due again by the differences, already pointed, of NN1 and NN2 second constituent that did not differed from embedded constituent of NC2.

Further analyses were carried to investigate inter-categories differences between head and modifier in N-N compounds. Planned contrasts showed that in NN2 head was longer, less imageable and acquired later compared to head of NN1 compounds. Length of NN1 modifier was also longer than Length of NN2 modifier.

Compounds (NN1 and NN2) were slightly longer then noncompounds [$t(110)=3.08$]

Table 5.1 displays the stimuli pairs and the assigned category names.

EXPERIMENTAL CONDITIONS				
Condition	**Prime**	**prime type**	**Target type**	**e.g.**
rel_NC1_1	related	**NC1**	left embedded word	*coccodrillo*-COCCO
unrel_NC1_1	unrelated	**NC1**	left embedded word	*alabarda*-COCCO
rel_NC2_2	related	**NC2**	right embedded word	*tartaruga*-RUGA
unrel_NC2_2	unrelated	**NC2**	right embedded word	*requisito*-RUGA
rel_NN1_1	related	**NN1**	1° constituent	*capobanda*-CAPO
unrel_NN1_1	unrelated	**NN1**	1° constituent	*malelingue*-CAPO
rel_NN1_2	related	**NN1**	2° constituent	*capobanda*-BANDA
unrel_NN1_2	unrelated	**NN1**	2° constituent	*astronave*-NAVE
rel_NN2_1	related	**NN2**	1° constituent	*astronave*-ASTRO
unrel_NN2_1	unrelated	**NN2**	1° constituent	*carovita*-ASTRO
rel_NN2_2	related	**NN2**	2° constituent	*astronave*-NAVE
unrel_NN2_2	unrelated	**NN2**	2° constituent	*cavolfiore*-NAVE
unrel_filler	unrelated	**derived word**	filler	*confederazione*-GESTO
unrel_nonword	unrelated	**derived word or V-N**	nonword	*affermazione*-*SDREIA *contagocce*-*GUSPODE

Tab. 5.1 Stimuli pairs

5.2.3 Procedure

The participants were tested individually in a dimly lit, quiet room. They were administered a series of letter strings presented one at a time in the center of the screen, and they were asked to decide as quickly and accurately as possible whether or not each string was a word. No mention of the existence of the prime stimulus was made. A fixation point was presented for 1000 ms and followed by a forward mask of hashes with the same number of letter of the prime was presented for for 500 msec. The hashes were followed by the prime, presented in Courier lowercase for 46 msec. Prime was followed immediately by the target, displayed in Courier uppercase. Target remained on the screen until an answer was given and nor more than 1500 msec. No feedback on the correctness of answer was given.

Trial sequence of event is depicted in Fig. 5.1. Targets were presented in a different random order for each participant, and the participants were given 5 practice trials before the beginning of the experiment. Stimuli presentation and response recording was made with E-prime software.

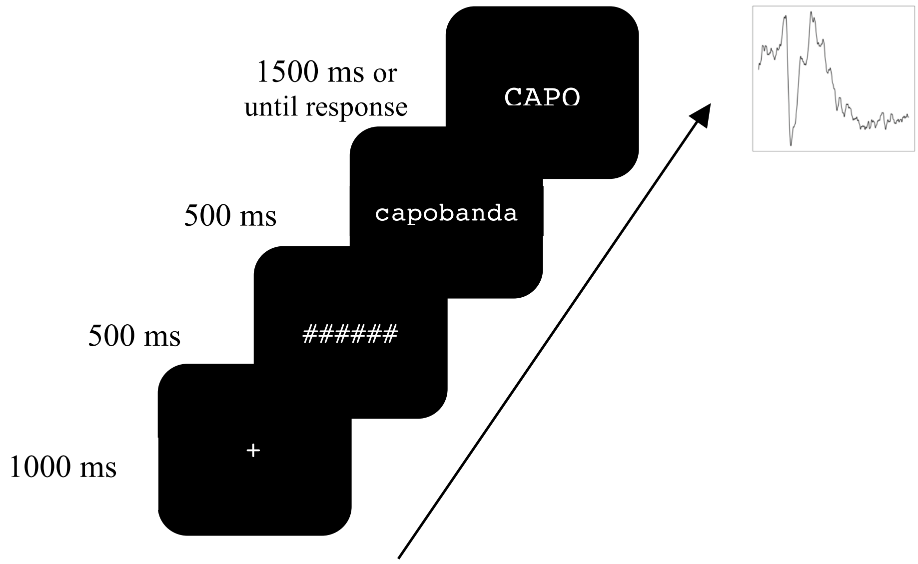


Figure 5.1 Trial structure of the experiment. ERP recording was time-locked with target presentation

Behavioural data were analyzed through mixed effect modeling (Baayen, 2004; 2007). Subject and Items were considered as random effect. Several psycholinguistic variables of were considered as fixed effect covariates: Frequency, Length, and Neighborhood size of primes (noncompounds and compounds); Frequency, Length, Neighborhood size, Age of Acquisition, Familiarity and Imageability of target (compound constituents and embedded words in noncompounds). Prime relatedness and target type were inserted as fixed effect factors.

EEG collection and analysis

EEG was recorded from 28 scalp electrodes mounted on an elastic cuff and located at standard leftand right-hemisphere positions over frontal, central, parietal, occipital, and temporal areas (International 10/20 System, at Fz, FCz, Cz, CPz, Pz, Oz, Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, Ft7, Ft8, Fc3, Fc4, Cp3, Cp4, Tp7, Tp8). These recording sites plus an electrode placed over the right mastoid were referenced to the left mastoid electrode. The data were recorded continuously by a SynAmps amplifier and NeuroScan 4.3 software. Each electrode was rereferenced offline to the algebraic average of the left and right mastoids. Impedances of these electrodes never exceeded 8 kV. The horizontal electro-oculogram (HEOG) was recorded from a bipolar montage with electrodes placed 1 cm to the left and right of the external canthi. The vertical electro-oculogram (VEOG) was recorded from a bipolar montage with electrodes placed above and below the right eye. The EEG was amplified by a Synamp's amplifier digitized at a rate of 500 Hz and filtered during the offline analysis with a band pass of 0.01–30 Hz. Data were reduced by using Edit 4.3 software, and ocular artifact were reduced with built-in function based on the method by Semlitsch, Anderer, Schuster, and Presslich (1986). ERP data were analysed only for correct responses. The dependent variable considered was the mean amplitude in the given intervals. The global window considered lasted 1,400 ms, starting from the onset of the target stimuli. The preceding 100-ms period before target onset was taken as prestimulus baseline. Epochs in which amplitude exceeded -70μV - +70μV range were excluded.

Four Region Of Interests (ROI) made up by four electrodes were thus identified: Left Anterior, LA, (F7, F3, FT7, FC3), Right Anterior, RA (F4, F8, FC4, FT8), Left Posterior, LP (TP7, CP3, P7, P3), Right Posterior, RP (CP4, TP8, P4, P8). Amplitude value for each ROI was obtained by algebraic mean of single electrodes amplitude.

Data were analyzed through repeated measure ANOVA. Mauchly's test was used to check sphericity assumptions and Greenhouse–Geisser correction for sphericity departures was applied when necessary (Geisser & Grenhouse,1959). Together with Condition (12 levels, see tab 6.1), two topographical variables (reflecting ROI distinction) were included in the analysis: Caudality (anterior vs posterior) and Laterality (left vs right hemisphere).

Results

Behavioural data.

Error rate was really low (2%-6%) and thus errors were not further analyzed. RT was log transformed before entered in the analysis, in order to reduce skewness. Table 5.2 shows mean RT for condition.

Observed Reaction Times			
category	**mean (sd)**	**e.g.**	**translation**
rel_NC1_1	672.52 (177.85)	*coccodrillo-COCCO*	crocodile - COCONUT
unrel_NC1_1	706.75 (168.55)	*alabarda-COCCO*	halberd- COCONUT
rel_NC2_2	718.29 (155.64)	*tartaruga-RUGA*	tortoise - WRINKLE
unrel_NC2_2	716.25 (157.03)	*requisito-RUGA*	requiriment - WRINKLE
rel_NN1_1	632.27 (126.90)	*capobanda-CAPO*	band eader (lit.leader band) - LEADER
unrel_NN1_1	670.45 (131.04)	*malelingue-CAPO*	insults (lit. evil tongues) - LEADER
rel_NN1_2	679.43 (146.05)	*capobanda-BANDA*	band leader (lit.leader band) - BAND
unrel_NN1_2	660.09 (142.76)	*centrotavola-BANDA*	centre-piece (lit. center table) - BAND
rel_NN2_1	669.94 (137.49)	*astronave-ASTRO*	spaceship (lit. star ship) - STAR
unrel_NN2_1	659.51 (144.37)	*carovita-ASTRO*	high cost of living (lit. cost life) - STAR
rel_NN2_2	679.42 (139.38)	*astronave-NAVE*	spaceship (lit. star ship) - NAVE
unrel_NN2_2	719.11 (159.31)	*cavolfiore-NAVE*	*cauliflower (lit.* cabbage flower) - NAVE
unrel_filler	709.17 (160.59)	*confederazione-GESTO*	*confederation - GESTURE*
unrel_nonword	780.13 (161.71)	*affermazione-*SDREIA* *contagocce-*GUSPODE*	*affirmation* *dropper*

Table 5.3. Observed Reaction Times

Model 1 (priming effect)

Final mixed effect regression model was selected by backward elimination of not significant variables. Initial models included: factor Condition with ten levels (rel_NC1_1, unrel_NC1_1, etc.,) and all psycholinguistic variables listed in par 5.2.2. Filler and nonwords were excluded from this analysis and no interactions were included. Group was included as random effect, with Condition as nested variable. Subject, word and Subject group were added as further random effects.

Regression coefficients of final model are listed in Table 5.4. Random effects are listed in 5.5.

MODEL 1 - Mixed Model fixed effects				
Variables	**Coefficients**	**Standard Error**	**t-value**	**pMCMC**
rel_NC1_1 (Intercept)	6.56	0.04	164	0.00001**
unrel_NC1_1	0.05	0.04	1.25	0.08
rel_NC2_2	0.06	0.04	1.5	0.07
unrel_NC2_2	0.06	0.09	0.67	0.45
rel_NN1_1	-0.02	0.01	-2.1	0.01*
unrel_NN1_1	0.05	0.04	1.25	0.11
rel_NN1_2	0.05	0.04	1.26	0.13
unrel_NN1_2	0.03	0.01	1.95	0.04*
rel_NN2_1	0.04	0.03	1.33	0.27
unrel_NN2_1	0.03	0.04	0.75	0.31
rel_NN2_2	0.006	0.01	0.1	0.64
unrel_NN2_2	0.07	0.02	1.75	0.04*
Target Age of Acquisition	0.02	0.002	10.18	0.0001**
Target Frequency	-0.01	0.002	-5.57	0.0001**

Table 5.3. Fixed effects. Table shows the significant predictor of final mixed effects regression. Dependent variable is predicted Reaction Time after logarithmic transformation.

MODEL 1 - Mixed model random effects	
Variable	**variance**
Word	0.001
Subject	0.01

Table 5.4. Random effects. Table shows the significant random variables of final mixed model regression.

Data in table 5.2 must be interpreted as follows: First row (rel_NC1_1**)** indicates the intercept: i.e. the level taken as reference level. For the theoretical predicted RT associated to this reference level, all covariates (target length, frequency and age of acquisition, in this model) are assumed to be equal to 0. Coefficients for levels of categorical variables other than the one taken as Intercept (so unrel_NC1_1, rel_NC2_2, etc.) must be added to the Intercept in order to obtain the predicted value for the intercept. The H_0 of statistical test reported in table 5.2 is that the considered coefficient is equal to 0, and thus that there is no difference between the reference level and the considered level. For example the coefficient for unrel_NC1_1 is 0.05. The statistical test is not significant: H_0 cannot be rejected and the coefficient cannot be considered different from 0. Predicted value for unrel_NC1_1 is Intercept plus the coefficient for unrel_NC1_1, this would mean that predicted value for unrel_NC1_1 is virtually 6.56 + 0 and so the same as rel_NC1_1. Coefficients for covariates must be interpreted differently. A positive coefficient indicates that as the variable increases, the RT increase (e.g. target Age of Acquisition = 0.02) and a negative coefficient indicates that as the variable decreases, RT decrease (target Frequency= -0.01). All p values listed in table 5.2 are not traditional p values. As pointed out by Baayen, Davidson and Bates (2008), within mixed effects model p value estimates may easily lead to type I errors. A more robust alternative is represented by estimated p values, obtained with Markov chain Monte Carlo (MCMC) simulations, thus **pMCMC** (that can be interpreted just as traditional p value see Baayen, Davidson and Bates, 2008 for some statistical detail on MCMC simulations). All pMCMC listed in tab. 5.3, are not so informative, because the statistical test carried tell if the coefficient is different from the reference level. To investigate the presence of differential priming effects, pMCMC factorial contrasts were carried out comparing related and unrelated condition for each stimulus Type. For the first constituent a difference was found for NN1 [pMCMC = 0.01] but not for NN2 [pMCMC = 0.56], and NC1 [pMCMC = 0.08]. For the second constituent a difference was found only for NN2 [pMCMC = 0.04] but not for NN1 [pMCMC = 0.31]. . These pMCMC values are calculated within a mixed effects model, and thus with the effects of covariates and random variables ruled out

Thus, a priming effect was found only for Noun-Noun compounds when the constituent was the head. Figure 5.2, 5.3 and 5.4 shows the effect of the significant variable of Model 1

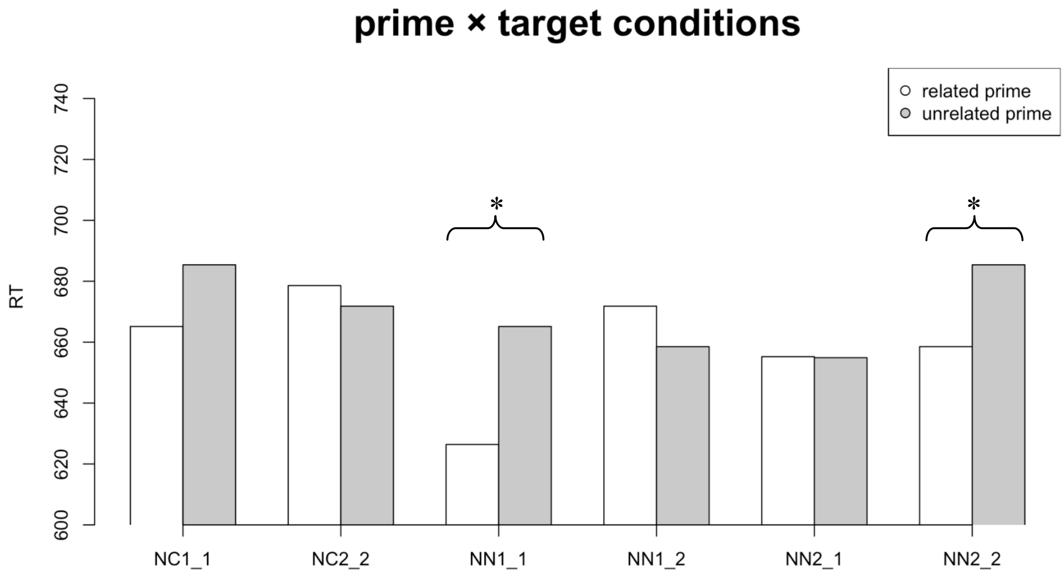


Figure 5.2 Prime × target condition effect.
The plot shows the effect of Types on RT across the different conditions. White bars indicate the related prime condition, while Gray bars indicate the unrelated prime condition.

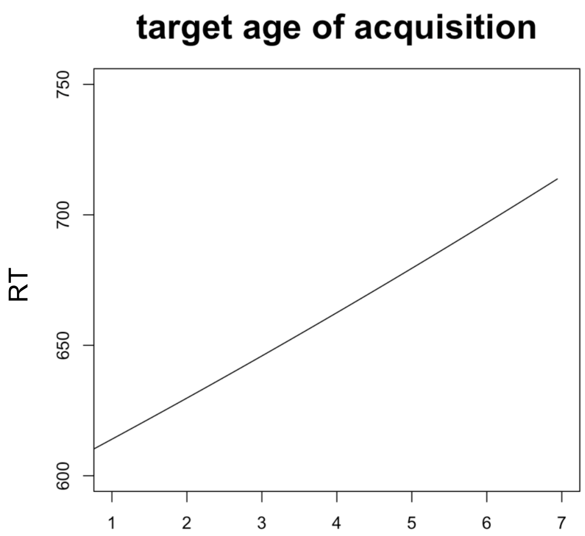


Figure 5.3 Target age of acquisition effect
The plot shows the linear effect of age of acquisition

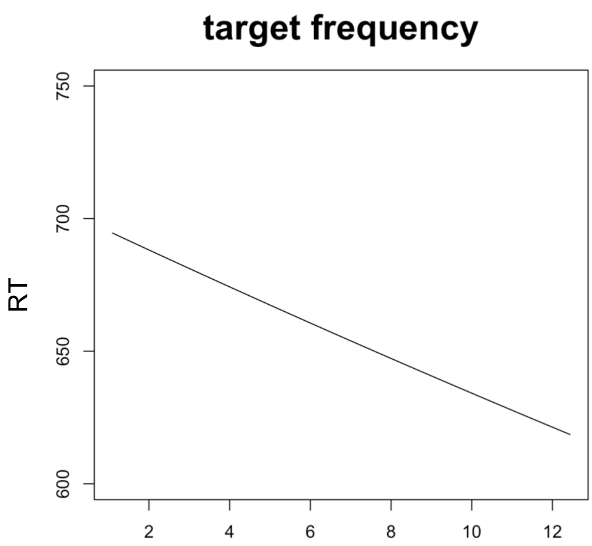


Figure 5.4 target frequency effect
The plot shows the linear effect of target frequency

MODEL 2 - Priming amount

In order to investigate the differences of priming amount, another model was fit, using the RT difference between related and unrelated condition (prime = RT related – RT unrelated) as dependent variable. Initial models included factor *type* with six levels (NC1_1, NC2_2, NN1_1, NN1_2, NN2_1, NN2_2) and all psycholinguistic variables listed in par 5.2.2. No interactions were included. *Group* was included as random effect, with *condition* as nested variable. *Subject* and *word* were added as further random effects.

Regression coefficients of final model are listed in Table 5.25 Random effects are listed in Table 5.6.

MODEL 2 - Mixed Model fixed effects				
Variables	**Coefficients**	**Standard Error**	**t-value**	**pMCMC**
NC1_1 (Intercept)	-28.28	32.40	-0.87	0.52
NC2_2	32.01	21.26	1.5	0.41
NN1_1	-62.09	21.21	-2.95	0.01*
NN1_2	40.15	22.24	1.8	0.11
NN2_1	29.22	21.37	1.36	0.33
NN2_2	-50.18	22.54	-2.22	0.02*
Target Frequency	-11.74	2.86	-4.10	< 0.001**

Table 5.5. Fixed effects. Table shows the significant variables of final mixed effects regression. Dependent variable is Reaction Time difference between related and unrelated condition.

MODEL 2 - Mixed model random effects	
Variable	**variance**
Word	62.7
Subject	10.6

Table 5.6. Random effects. Table shows the significant random variables of final mixed model regression.

Coefficient for the reference level (NC1_1) was not statistical significant. This means that the difference of RT in the related condition minus RT in unrelated conditions for NC1_1 is not different from 0 and hence that there is not a priming effect for this type. As expected, a significant difference was found for NN1_1 and NN2_2, indicating that the RT difference is lower in this condition, and then that the priming magnitude is higher. A contrast between RT difference for NN1_1 and NN2_2 was carried out in order to investigate the difference in the amount of priming between these two conditions. The contrast was significantly different (pMCMC = 0.01*), indicating that priming effect for NN1 was higher than that of NN2. Figure 5.7 shows the different priming amounts associated with target category, while figure 5.8 show the effect of frequency found in Model 2.

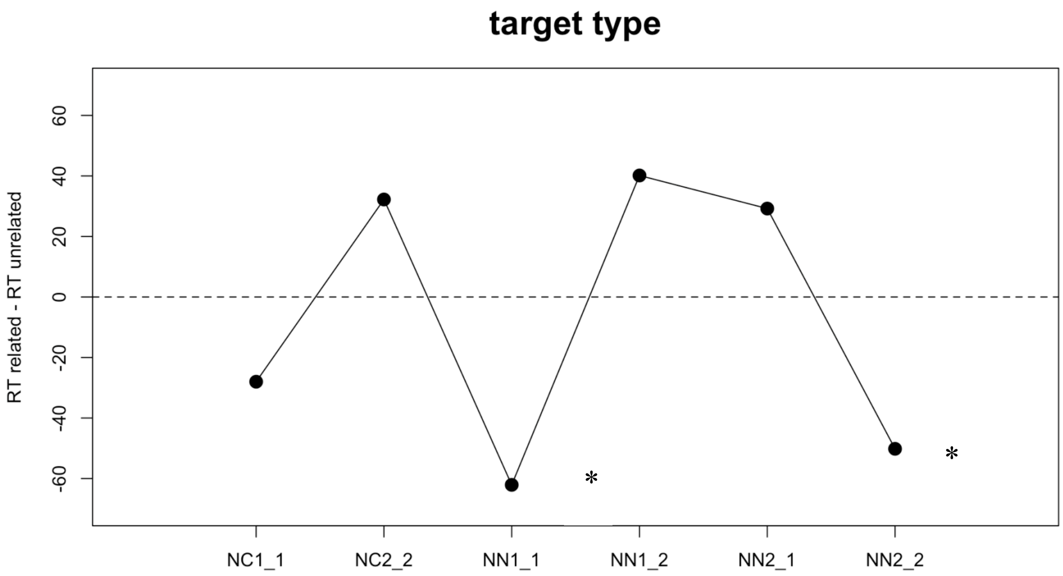


Figure 5.7. Prime amount
The plot shows the priming amount (RT related – RT unrelated) across the different conditions.

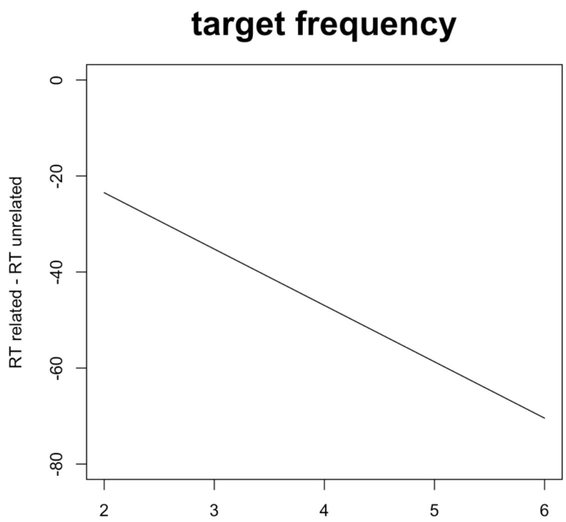


Figure 5.8 Target age of acquisition effect
The plot shows the effect of age of acquisition

ERP Data

The traces presented in Figure 6.9 and 6.10 show the grand average potentials recorded at the scalp (figure 6.10) shows enlarged images of electrodes Fz, Cz, and Pz). Target stimuli elicited the N1-P2 complexes followed positive shift (P300) appeared followed. Then a negetive shift started at about 300 ms and lasted until 600 ms poststimulus approximately (N400). Analysis on ERP data were confined to seven time windows, based from visual inspection and from latencies defined in Lavric et al. (2007). Four windows spanned N400, in 2 latencies windows of 100 msec each through 370–and 470–570. One more window included all N400: 370-570. After visual inspection of data, Two more windows 180-210, 220-260, reflecting the early positive shift peak around 200 ms (P200) and the following negative shift. Trials with erroneous responses (4.5%) were excluded from the analyses as well trials with excessive artifacts (16.5%). Thus 79% of the trials entered the averaging processing.

A first repeated measure 12 × 2 × 2 ANOVA was carried out on all data, with factors "condition" with ten levels (rel_NC1_1, unrel_NC1_1, etc., see table 5.1), Laterality with two levels (left hemisphere vs right hemisphere), and Caudality with two levels (anterior vs posterior). Data from 26 subjects were merged in 13 *supersubjects*. Each subject of one subset was paired to one subject of the other subset, with pairs randomly generated. In this way, in every supersubject all experimental conditions were covered. Before merging, data were z-transformed (a prelimiray check confirmed the normality of data distribution for all subjects, thus legitimating the data transformation).

Interval 180-210 (P200)

In first interval, no significant differences were found.

Interval 250-350 (P300)

In the second interval considered, only variable Caudality resulted significant with lower amplitudes for anterior electrodes compared to posterior electrodes [$F(1,12)=84,37$, $p<0.001$].

Interval 370-470 (N400 window 1)

In first N400 window, Laterality and Caudality × Laterality were significantly different. Amplitude in the left hemisphere was more negative than that of the right hemisphere [F(1,12)=19.69, p<0.001]. Caudality × Laterality was also significant. [F(1,12)=7.67, p=0.02], but no contrast highlighted significant differences: LP and LA were no different, but had significantly lower amplitude compared to RP and RA [t(12)=-4.44, p<0.001]; as LP and LA, also RP and RA weren't significantly different.

Interval 470–570 (N400 window 2)

In second N400 window, the same results of N400 windows 2 were found, with Laterality [F(1,12)=21.14, p<0.001]. Caudality × Laterality was significant [F(1,12)=8.31, p=0.01]. Contrast showed a statistically significant difference between LP and LA [t(12)=3.88, p=0.002] with more positive amplitudes for LP compared to LA, but no diffrences between RP and LP.

Interval 370–570 (N400)

When the overall N400 window was considered, Caudality [F(1,12)=29.17 , p < 0.001] Caudality × Laterality [[F(1,12)=9.10, p=0.01] , and Type [F(11, 12)=3.57, p = 0.02]. Within caudality, a more positive amplitude for posterior sites was found. Caudality × Laterality were significantly different, with highest amplitudes for LP compared to LA (t(12)=4.1 and no differences between RP and LP. An effect of Conditions × Caudality emerged. PlaN-Ned contrasts were carried out, comparing each related condition with its unrelated homologous, with separate comparisons for anterior and posterior clusters. A significant difference was found only for NN1 in the posterior clusters compounds with an higher amplitude for unrel_NN1_1 compared to rel_NN1_1 [t(12)=2.18, p = 0.03]. When comparing overall priming effects of compounds versus noncompounds no significance effect was found. Figures 5.11 and 5.12 show the averaged potentials for NN1_1 contrasting related and unrelated condition. Figure 5.13 and 5.14 show, for comparison purposes, the averaged potentials for NN2_2.

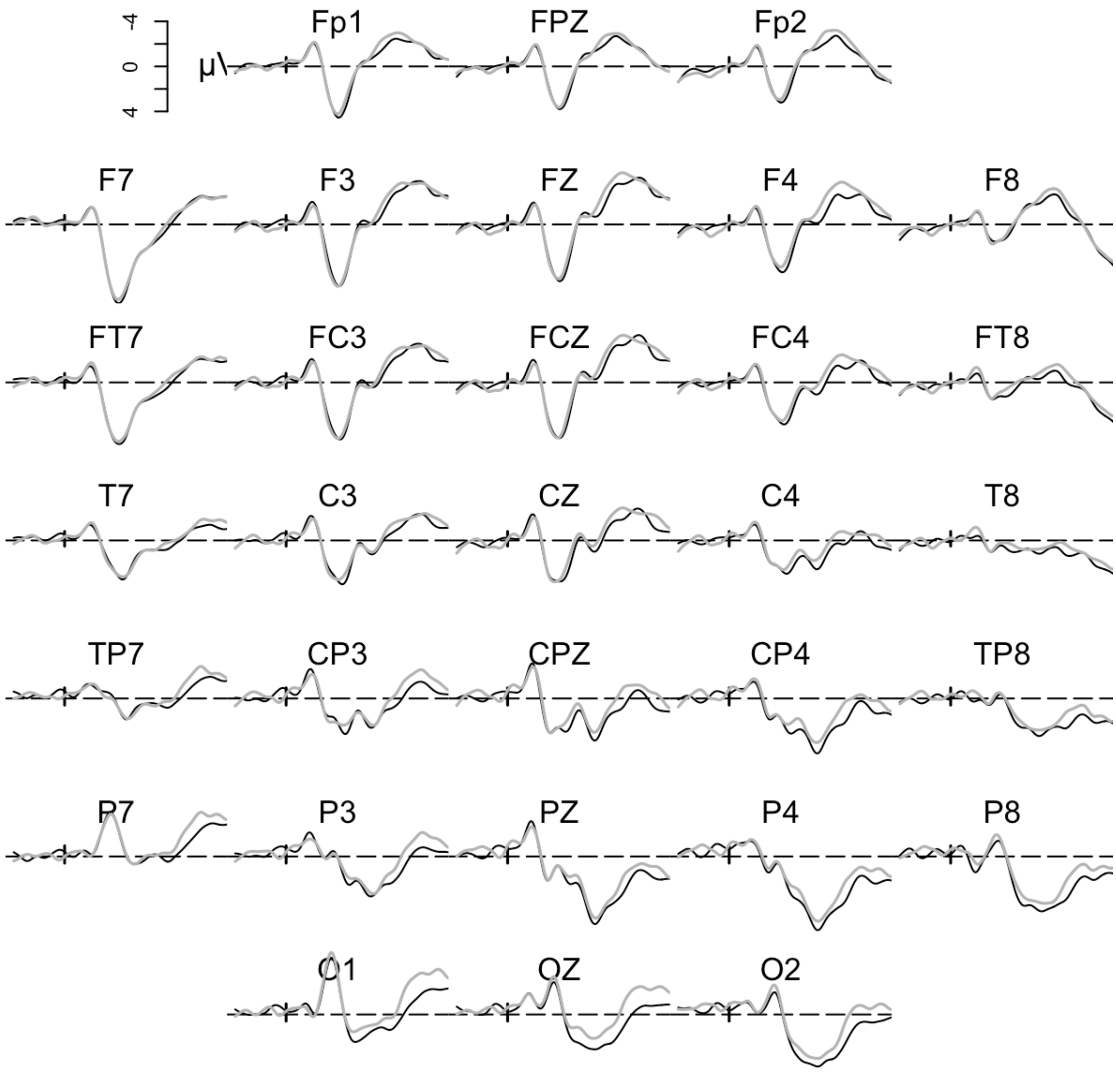


Figure 5.9. Related vs Unrelated. Grand average event-related potentials (ERPs) obtained for related (black) vs unrelated (gray) condition. No overall differences were found between the two conditions.

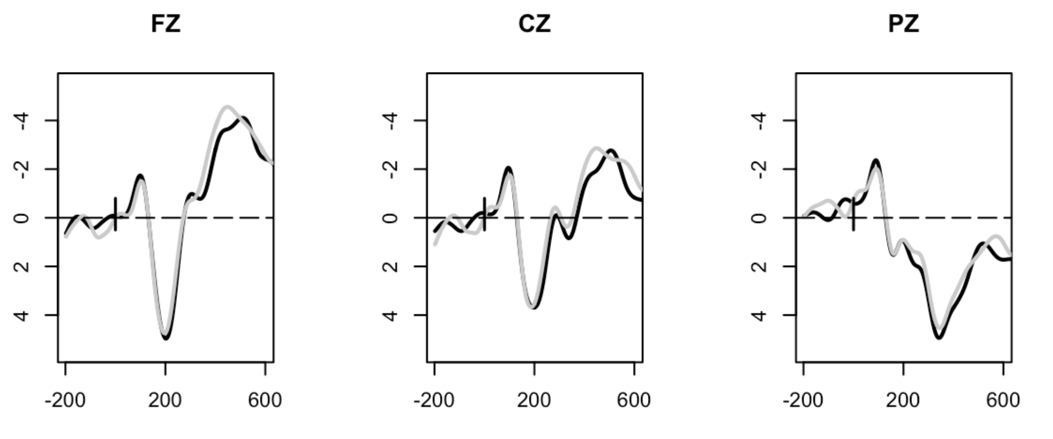


Figure 5.10. Related vs Unrelated. Grand average event-related potentials (ERPs) obtained for related (black) vs unrelated (gray) condition. Enlarged representation for midline electrodes. No overall differences were found between the two conditions.

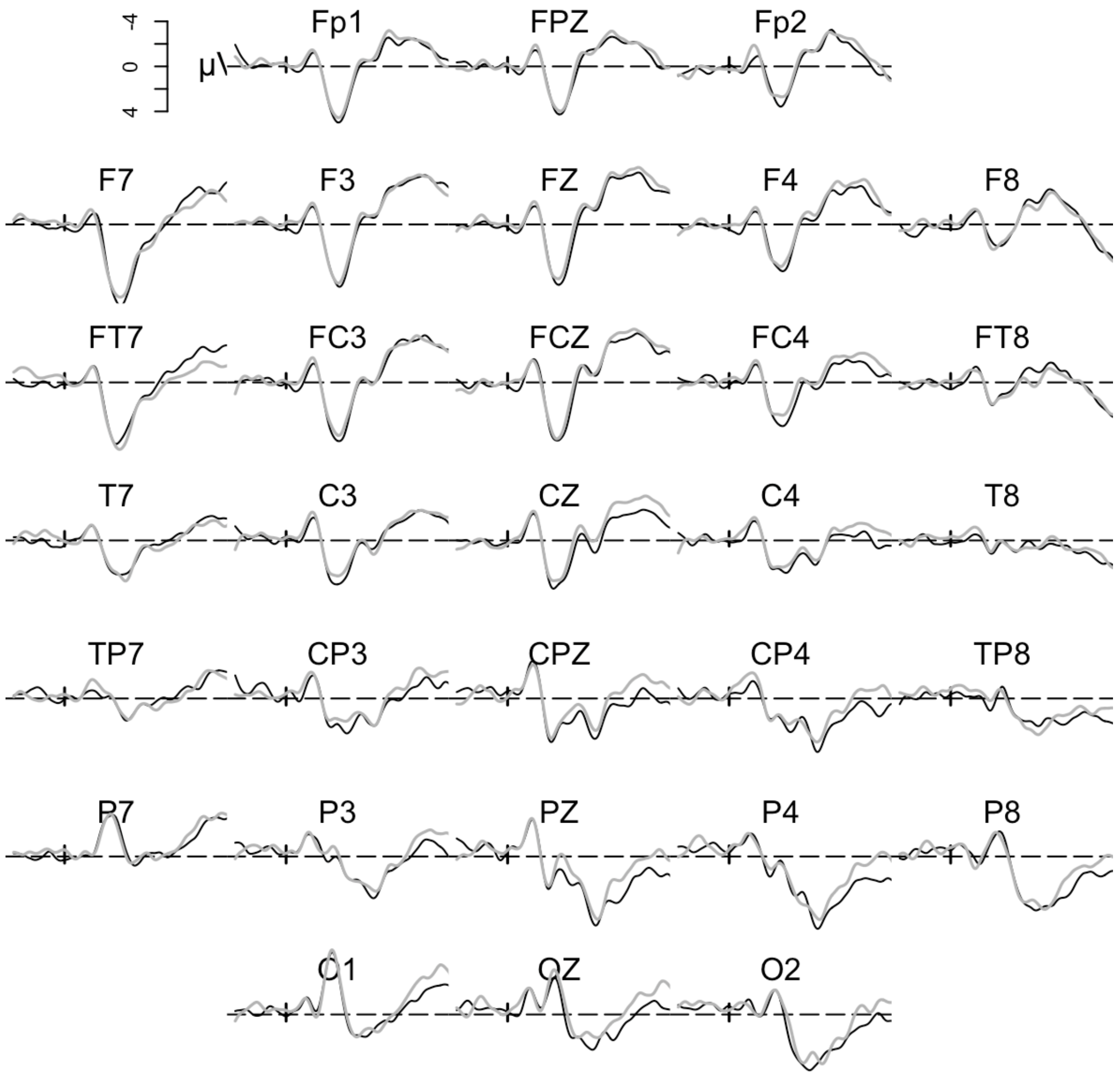


Figure 5.11. NN1_1 Related vs Unrelated. Grand average event-related potentials (ERPs) obtained for related (black) vs unrelated (gray) condition. A significant difference was found for posterior sites. with lower amplitude in the overall N400 span for unrelated, compared to related condition.

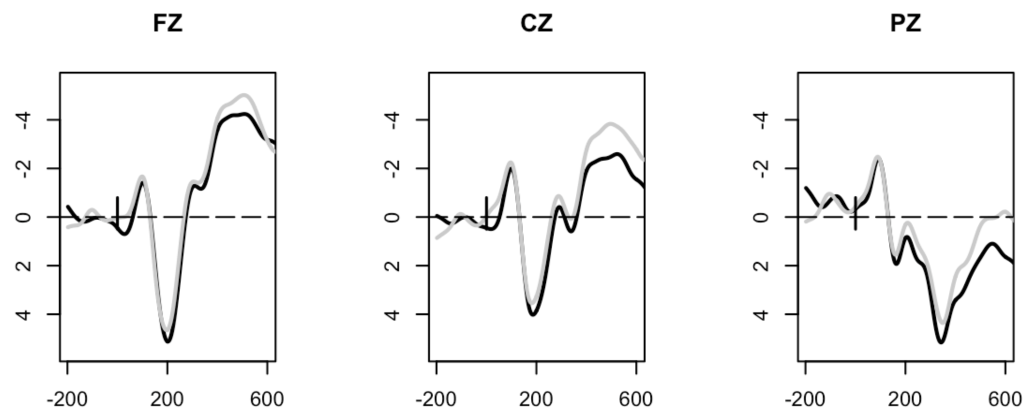


Figure 5.12. NN1_1 Related vs Unrelated Grand average event-related potentials (ERPs) obtained for related (black) vs unrelated (gray) condition. A significant difference was found for posterior sites, with lower amplitude, in the overall N400 span, for unrelated compared to related condition.

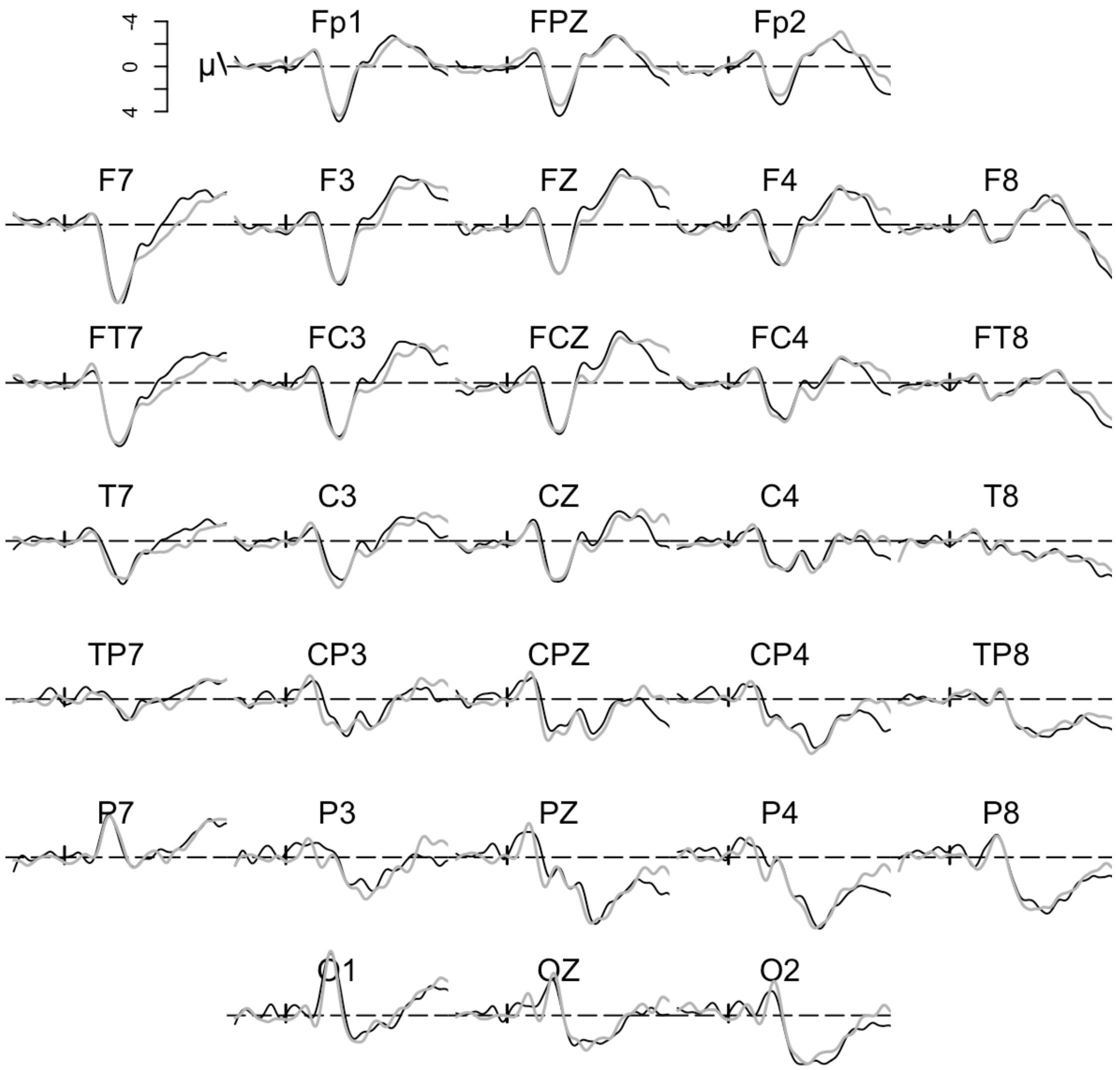


Figure 6.13. NN2_2 Related vs Unrelated Grand average event-related potentials (ERPs) obtained for related (black) vs unrelated (gray) condition. No differences were found between the two conditions.

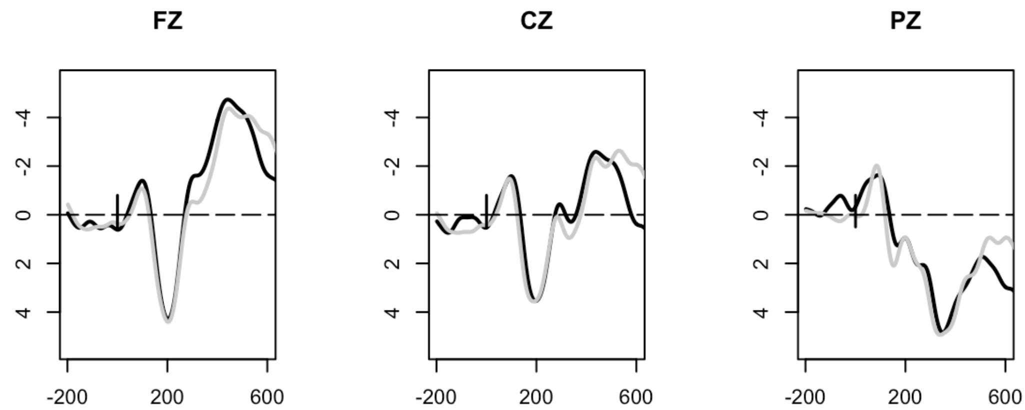


Figure 6.14. NN2_2 Related vs Unrelated The grand average event-related potentials (ERPs) obtained for related (black) vs unrelated (gray) condition. No differences were found between the two conditions.

5.3 Discussion

Effects of covariates observed where those expected according to literature: as Frequency of the target increases, Reaction Times decrease, while as Age of Acquisition increases (hence, the later the target is acquired), Reaction Times increase (Harley, 2008). The most important result is a priming effect found only on compounds the head constituent of NN1, the leftmost (*e.g. capobanda-capo*) and on the head of NN2 compounds, the rightmost constituent (*e.g. astronave-nave*). No priming effect was found for the orthographic conditions with noncompounds followed by the embedded constituents (*e.g. coccodrillo-cocco*, lit.'crocodile-coconut'; *tartaruga-ruga,* lit.'tortoise wrinkle').

Fiorentino & Fund-Reznicek (2008), using the same experimental design with English compounds, found an equivalent priming effect on both first and second constituents, but did not control for head position. Similar results were found (in part) by Diependaele et al. (2008), that found a significant priming effect on both constituents from Dutch familiar and unfamiliar compounds. These results could be considered as in support to full-parsing theories, and are in line with *blind-decomposition* framework of morphological processing. However, other recent results suggest that the picture is more complicated than what claimed by supporters of *blind decomposition* (Feldman et al. 2009). Diependaele et al. (2008) in their experiment found, with familiar compounds, a priming effect also when the bigram at the morpheme boundary was removed (e.g. *bookshop-book*→ *boo__hop-book*). This result thus suggests that also a whole-word representation may play a role in early processing and not only the morphemic structure.

How the results of the present study can be explained? The most parsimonious explanation is that of a semantic priming effect. Although sometimes contested and less robust compared to other types of priming, masked semantic priming is found (Forster, 1998). As previous stated (see par. 1.3) one of the main characteristics of the *head* of a compound, is that it determines, mostly, the semantic feature of the whole compound (at least for transparent compounds). Thus we may suppose that there is a stronger semantic connection between the compound and its head, compared to the connection with the compounds and the non-head constituents (as represented in Fig. 2.1, in which the shorter arrow between board and blackboard, indicates a stricter relationship between the two words). The differences from results of Fiorentino & Fund-Reznicek (2008) and Diependaele et al. (2009), could be language related. Compounding in Italian is moderately productive compared to English and Dutch and thus, the proportion of Italian compounds is smaller compared to that of the other languages. It has been suggested that morphological family size plays a role in triggering morphological decomposition (as claimed for example by Plaut & Gonnermann, 2000), with languages rich in morphology leading to parsing more than morphological poor languages. Even if Italian, in general, is a language with rich morphology the relatively limited number of compounds could favourite a whole-word first processing of this kind of words. This would explain why a truly morphological effect (with equivalent priming for first and second constituent, despite headedness) has not been found in the present experiment. A whole-word and semantic interpretation of results could also fit the results found with N400. The N400 component, traditionally associated with detection of a semantic incongruency (Friederici, 2004), can be also be modulated by unconsciously perceived masked words semantically related to target (Holcomb, 1993; Kiefer, 2002, Holcomb, Reder, Misra, Grainger, 2005). The N400, found in the present study is really similar to the traditional semantic N400, with highest amplitude for posterior sites

Why a neural priming (a modulation of N400) has been found only for NN1? Postulating an advantage of first position (as claimed in some cases for lexical access, Forster & Davis, 1976, Jarema et al., 1999) seems not the case, since it is difficult to conciliate an advantage of first position with the whole-word activation that seems to underlie the priming effect observed.

As already outlined in chapter 1, Italian Noun-Noun compounds are mostly left headed, and according to the core compound formation rule, novel compounds are mostly left headed. We also have seen how in Italian there are also right headed Noun-Noun compounds, and that there is also a certain degree of productivity of right headed compounds, which can be placed in the *periphery* of word formation rules (see par. 1.7). This could make less evident the “constituency” (the evidence that the word is made by constituents) of the compound. In other words, the connection between the compound and its constituent can be less strong, compared to NN2. These considerations could explain the results of the present experiment if placed within the *hybrid model of morphological proceesing* purposed by Diependaele et al. (2009) (Figure 6.15). According to this model morphological structure is represented at two distinct levels: 1) the *morpho-orthographic level* in which are mapped cluster of sublexical represntation of word forms on base of frequency. 2) the *morpho-sematic level* in which are mapped the regularities of word form onto semantics. In this model both morpho-orhographic (via constituents) and morpho-semantic (via whole words) interact in determining how morphologically complex words are processed.

The model is capable to explain both the findings of purely morpho-orthographic effects (Longtin and Meunier, 2005), the findings of semantic effects and the synergy of both effect (Giraudo & Grainger, 2001; Feldman et al., 2009).

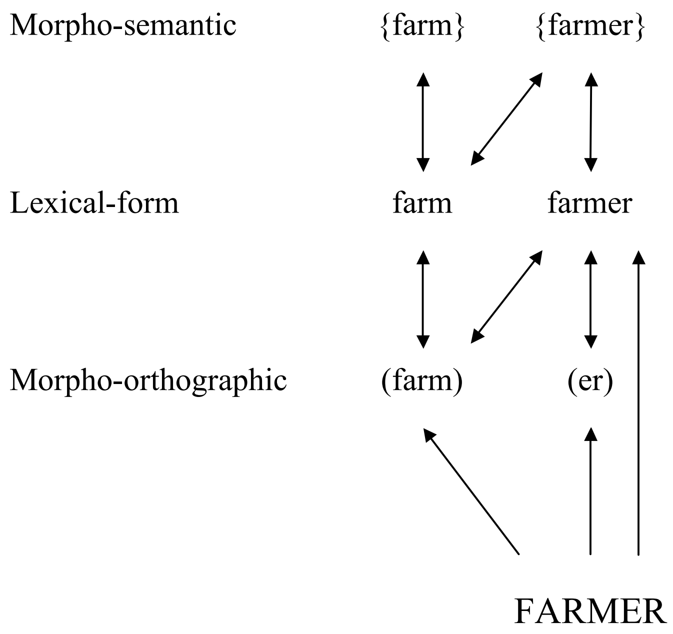


Figure 6.15 The hybrid model of morphological processing.
The hybrid model of morphological processing depicted from the perspective of a hierarchical interactive-activation account of word recognition. The input is mapped in parallel onto morpho-orthographic and morpho-semantic representations, via whole-word form representations in the latter case. Both online and offline interactions between morpho-orthographic and morpho-semantic representations are possible through feedback connections (taken from Diependaele et al., 2009).

Figure 6.16a and 6.16b show, respectively, the hypothetical lexical processing of NN1 and NN2. The strength of the connections is expressed by the thickness of the arrows. For both NN1 and NN2, the tickest line is the one that goes, through a whole word representation and that activate mostly the first constituent for NN1 and the second for NN2, the head constituent for both case.

In NN1 once the whole word activation occurs, feedback connections on morpho-orthographic level activate both constituents. In the case of the second constituent the activation is not sufficient to lead to a priming effect, while for head constituent (the first) magnify the effect already present.

In NN1 the only relevant way is the one that goes through the morpho-semantic level to the lexical form and no synergic interaction with the morpho-orthographic levels occurs.

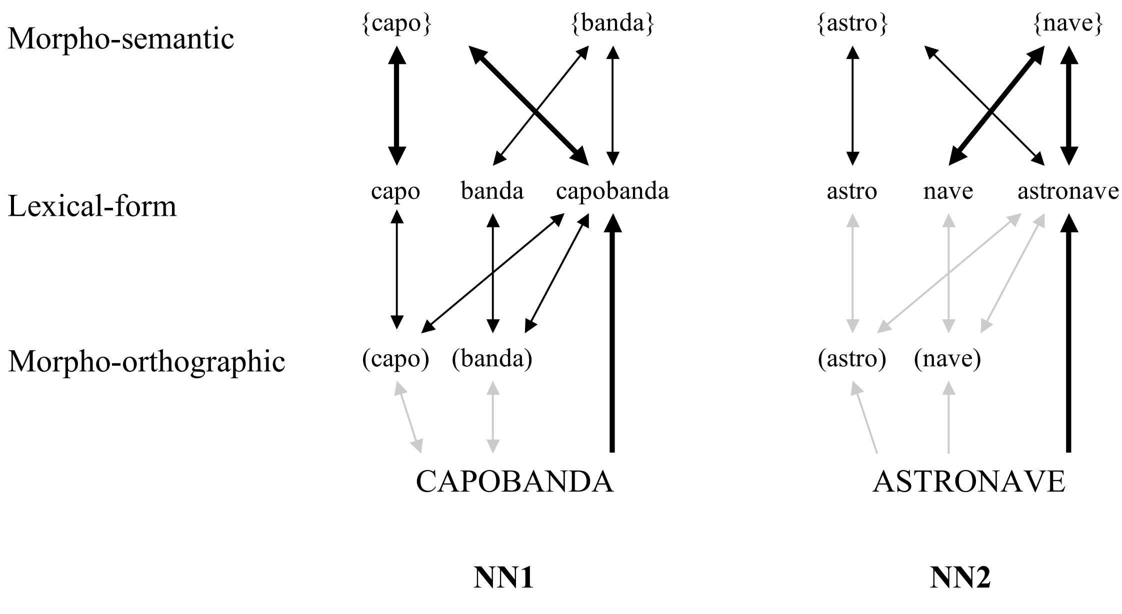


Figure 6.10 NN1 and NN2 within the framework of hybrid of morphological processing.
In both NN1 and NN2 constituent activation goes through the activation of whole word
In NN1 compounds a relevant activation of constituents come from morpho-semantic level and morpho-orthographic level. In NN2 compounds a relevant activation of constituents comes only from morpho-semantic level.

Summarizing, results of the present experiment suggest that head information of Italian Noun-Noun compounds is early accessed via an activation of a whole-word representation and its connection with morpho-semantic representations. An advantage of left headed N-N compounds has been found. This advantage could be related to the more evident morphemic structure of NN1, leading also to an activation of its constituents[14].

However, according to these results (and the result of the previos experiments) NN2 structure could be not only morphologically expected, but even not expected at all. In other words NN2 could be "frozen" compound structures, in which the morphological constituency doesn't play any role in online processing, and that could be processed as non-compound words. Linguistic evidences presented in chapter 1 (par 1.7), suggest that this is not the case, since a certain degree of right-headed productivity is present in Italian.

This issue will be further explored in the next experiment.

[14] An alternative explanation of the results can be found in the imperfect matching of psycholinguistic variables. While in behavioural data it has been possible could rule out the effect of covariates, in ERP analysis was not. These differences could influence the modulation of N400. Head constituents of NN2, in particular, were less imageable than head constituent of NN1. In a study by Nittono, Suehiro and Hori (2002) stimuli with high Imageability showed higher N400 than stimuli with low Imageability. Thus, the neural priming found only in NN1 could be related on the characteristics of the stimuli, that elicited an observable neural effect only in NN1. However, The imperfect matching cannot explain the presence of a behavioural priming for NN2 head.

6. EXPERIMENT 2 - Broken Compounds: "Constituency" and Morphological Representation of Compounds.

6.1 Introduction

Results from the previous experiment suggested that, in lexicalized Italian Noun-Noun compounds, information on head is early accessed, through the activation of whole word representation (chapter 5). Importantly, left headed Noun-Noun compounds (NN1) showed an advantage (in terms of a neural priming) compared to right headed compounds (NN2). These effects seem not to be merely semantic, but rather morpho-semantic: it has been hypothesized that NN2 have a less evident morphological structure compared to NN1. The combined effect of less evident constituency and of morpho-semantic activation could explain the results found and how information about head occurs.

In order to get better insight on how and why NN1 and NN2 are differently processed, Verb Noun compounds have been included a further control will be included in the present experiment. As pointed out in par. 1.7 both NN1 and V-N compounds are generated in accordance to the main word formation rules of Italian compounding, with the constituent order that reflects the canonical order of Italian syntax. As potential explanation of the results, NN2 could be processed only as a whole. Study that will be presented in Experiment 4, as well as other neuropsychological studies on Italian compounds yet suggests that this is not the case, since constituency effects are found also for NN2. An early effect of Headedness in Italian Noun-Noun compounds was also found by El Yagoubi et al. (2008), with right headed compounds showing a bigger P300 then left headed compounds, suggesting an higher processing cost. However, the experimental design employed and the stimuli used could have triggered a morphological decomposition (Shoolman and Andrews, 2003; Andrews, 1986) that possibly, doesn't occur in normal processing. So, although access to constituents is possible in NN2 processing, it is not clear if this occurs automatically as for NN1. Linguistically, a whole-word processing of NN2 won't be surprising, since they are less frequent morphological structures. Obviously this won't explain how novel right-headed compounds are processed, but could explain the results found with the stimuli used in the previous experiment, since they're all lexicalised compounds.

The aim of this experiment is to further explore this issue, by employing the experimental paradigm of "broken compounds" (Libben et al, 2003), particularly suited for investigating structural representation in the lexicon. In this paradigm, compounds and control stimuli are presented in two condition: as whole word (e.g. *capobanda*) with constituents separated by two spaces (e.g., *capo banda*)

The rationale behind the "broken compound" paradigm is straightforward. Morphological decomposition of compounds could be conceptualized as the separation of the compound into its constituents. If this is the case, it would be reasonable to assume that compounds that are actually decomposed into their constituents would be less affected by an alteration of the stimulus into two actual words than those that are not decomposed (Libben et al., 2003). This paradigm was originally used to study the differences between transparent and opaque compound processing, but it is possible to extend the same line of reasoning to the issues raised so far on Italian compound structure. If V-N and NN1 compounds has a morphological structure that keeps its constituent as "more separated" (i.e. the morphological structure is more evident), than a smaller cost in term of RT increase for broken compounds would be expected for these categories in comparison to NN2. Furthermore, the comparison with monomorphemic controls, whose structure resemble that of a compound could disentangle if NN2 indeed behave as just as structure represented as a whole.

In order to obtain a better insight on early and automatic processing of words in this context ERP data were recorded during task administration. The modulation of two components was expected. The N400, index of lexico-semantic and morphological integration (Kutas & Federmeier, 2005; Steinhauer & Connolly, 2008) was expected to be higher with NN2 compared to NN1 and VN. A different modulation of P600 was also expected. P600 is associated with syntactic reanalysis of material (Hahne, Friederici, 2006) and it has been found also in compound processing (Koester et al. 2006). A modulation was also expected for the earlier components, specifically for early negativities (200-250 ms latencies) sometimes interpreted as LAN (Hahne, Friederici, 2006), and already found in compound processing (Koester et al., 2007; El Yagoubi et al. 2008, Vergara-Martínez et al., 2008).

6.2 Materials and procedure

6.2.1 Partecipants

Twenty-five students from the University of Padova, all right-handed native speakers of Italian, participated to the experiment. All subjects were right handed, they had no neurological pathologies, normal or corrected to normal vision and received course credits for their participation. Two participants' data were excluded from the analysis because of excessive artifacts in the electroencephalogram (EEG).

The remaining 23 participants had a mean age of 23.08 years (range 20-28); 15 of them were women and 9 were men.

6.2.2 Materials

Experimental material consisted in 240 stimuli, with 120 words and 120 nonwords. *Word list* was composed by 24 **NN1** compounds, 24 **NN2** compounds, 24 V-N compounds, 24 **NC1**, noncompounds with left embedded word, and 24 **NC2**, noncompounds with right embedded word. Stimuli were the same of precedent experiments with some replacements. Within the N-N list, four adjective nouns were included two right headed (A-N) and two left headed (N-A) in order to have right and left headedness condition as matched as possible for psycholinguistic variables.

All compounds of the different cateogories were chosen such as the single word form (e.g. *motosega*), was highly prevailing, in terms of frequency compared to the separate form (e.g. *moto sega*). Stimuli of *nonword list* were built mirroring the experimental condition of word list: 48 **nonNN2**, stimuli made by two semantically unrelated existing words that never appears in lexicalized compounds (e.g. **tortacinghia***,** **'***cakestrap'); 24 **nonV-N** made up by a verb and an unrelated word. Verb was inflected in according to normal Verb-Noun compounding (e.g. *chiudicurva*, '*closecurve')

24 **nonNC1** stimuli made by the combination of a word (in first position) and of a real existing word segment in the second position (e.g. *<u>*uovo*</u>*taria*, with *uovo* 'egg' and **taria* that is an ending segment that occurs in real Italian words, as in *monetaria,* monetary); 24 **nonNC2**, stimuli made by the combination of a real existing word

segment in the first position ad of a real word in the second position (e.g. **consepera*, with **conse* that is a beginning segment that occurs in real Italian words as in *consegna* 'delivery' and *pera* 'pear', that is a real Italian word. Stimuli were presented in two condition: as whole words (e.g. *capobanda, astronave, coccodrillo*) or "broken" in two separated letter strings (e.g. *capo banda, astro nave, cocco drillo*). In the "broken" condition words were splitted with these criteria: compounds and nonwords resembling compounds were splitted according to morphemic structure; word and nonword with an embedded word were splitted such that the embedded segment was separated from the other portion of the whole stimulus (e.g. *cocco drillo, tarta ruga, conse pera*). Stimuli of nonword conditions were built in order to prevent the influence of unwanted strategies (implicit or explicit) in the task, especially for the "broken" condition. Without a nonword condition that resembled that of compounds, a subject could have afforded the task by simply analyzing one of the segment of the word and using the euristic: "if one of the segment is a word, than the whole stimulus is a word". The nonword list composition, instead, forced the subject to elaborate both the segments before making a choice in the lexical decision.

All Stimuli were divided in two blocks, with each stimulus appearing only once in each block (either in normal or in broken condition). Blocks were presented in counterbalanced order to each subject such that each stimulus was seen by half of the subjects first in normal and then in broken condition and by the other half in the opposite order. Experimental stimuli are listed in Table 6.1.

Planned t-test were carried out to investigates differences in psycholinguistic variables across categories. In case of high departure of normality, t-test were substituted by two-sample Wilcoxon tests.

Stimuli were matched as much as possible, but V-N compounds resulted quite different from the others because of the intrinsic differences with the other stimuli. Noun-Noun compounds were matched for frequency for length of both constituents and whole words. Embedded words were matched for frequency of whole word and for frequency of the embedded segment. Noun-Noun compounds were longer than noncompounds [$t(94)=3.06$, $p=0.001$]. Verb-Noun compounds were longer the noncompounds [$t(115)=5.49$, $p<0.001$] and then Noun-Noun compounds [$t(115)=2.35$, $p=0.007$]. Furthermore, first constituent of Verb-Noun compounds was longer the constituent of compounds [$t(115)=2.35$, $p=0.007$] and less frequent [t(115)=-5.92, p<0.001]. Neighbourhood size was matched within compounds and within noncompounds. However a difference was present between these categories. NC2 has an higher neighborhood size of the right embedded word compared to NN2 rightmost constituent of NN2 [W = 265, p = 0.03]. Neighborhood size of first word segment was almost matched, with only first segment of NC1 compounds with an higher neighborhood size compared to the Verb constituent of V-N [W = 154, p = 0.005]. Neighbourhood size of the rightmost segment was almost matched, with only NC2 rightmost segment, with an higher neighborhood size than the rightmost constituent of V-N [W = 390, p = 0.03]

EXPERIMENTAL STIMULI			
category	**whole**	**broken**	**e.g.**
NN1	*capobanda*	*capo banda*	'band leader' (lit. 'leader band)
NN2	*astronave*	*astro nave*	'spaceship' (lit. 'starship'
V-N	*prendisole*	*prendi sole*	'sundress' (lit. 'take sun')
NC1	*coccodrillo*	*cocco drillo*	'crocodile' (*cocco*, 'coconut')
NC2	*tartaruga*	*tarta ruga*	'tortoise' (*ruga*, 'wrinkle')
nonN-N	**tortacinghia*	**torta cinghia*	*'cakestrap'
nonV-N	**chiudicurva*	**chiudi curva*	*'closecurve'
nonNC1	**uovotaria*	**uovo taria*	*uovo* 'egg'
nonNC2	**consepera*	**conse pera*	*pera* 'pear'

Table 6.1 Experimental Stimuli. Table shows experimental stimuli divided for types.

6.2.3 Procedure

The participants were tested individually in a dimly lit, quiet room. They were advised that they would be seeing a series of letter strings presented one at a time in the center of the screen, and that they would be required to decide as quickly and accurately as possible whether or not each string was a word. They were advised that the target word could appear as a whole string or as two separated strings. They were explicitly asked to ignore the space between the strings and to answer as if they were connected. They were also instructed to make as few eye movements as possible.

A fixation of point was presented for 1500 ms, followed by the target presented in Courier uppercase. Target remained on the screen for 2500 ms or until a response was given. Subjects received no feedback on correctness of the answers.

Trial sequence of event is depicted in Figure 6.1. Targets were presented in a different random order for each participant, and the participants were given 5 practice trials before the beginning of the experiment. Stimuli presentation and response recording was made with E-prime software.

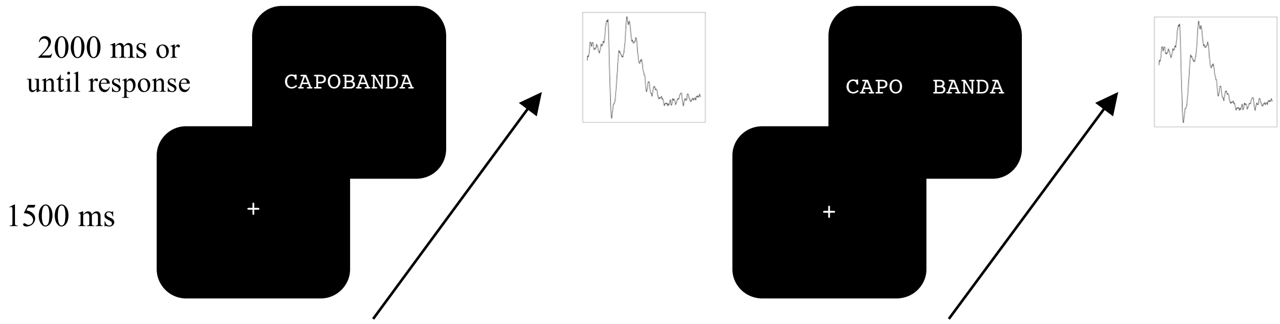


Figure 6.1 Trial structure of the experiment. ERP recording was time-locked with target presentation

Behavioural data were analyzed through mixed effect modeling (Baayen, 2004; 2007). Subject and Items were considered as random effect. Several psycholinguistic variables of were considered as fixed effect covariates: Frequency, Length, and Neighborhood size of target. Stimulus Category (NN1, NN2, V-N, etc.) and Status (normal vs broken) were included as fixed effect.

EEG recording and analysis

EEG was recorded from 28 scalp electrodes mounted on an elastic cuff and located at standard left and right-hemisphere positions over frontal, central, parietal, occipital, and temporal areas (International 10/20 System, at Fz, FCz, Cz, CPz, Pz, Oz, Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, Ft7, Ft8, Fc3, Fc4, Cp3, Cp4, Tp7, Tp8). These recording sites plus an electrode placed over the right mastoid were referenced to the left mastoid electrode. The data were recorded continuously by a SynAmps amplifier and NeuroScan 4.3 software. Each electrode was rereferenced offline to the algebraic average of the left and right mastoids. Impedances of these electrodes never exceeded 8 kV. The horizontal electro-oculogram (HEOG) was recorded from a bipolar montage with electrodes placed 1 cm to the left and right of the external canthi. The vertical electro-oculogram (VEOG) was recorded from a bipolar montage with electrodes placed above and below the right eye. The EEG was amplified by a Synamp's amplifier digitized at a rate of 500 Hz and filtered during the offline analysis with a band pass of 0.01–30 Hz. Data were reduced by using Edit 4.3 software, and ocular artifact were reduced with built-in function based on the method by Semlitsch, Anderer, Schuster, and Presslich (1986). ERP data were analysed only for correct responses. The dependent variable considered was the mean amplitude in the given intervals. The global window considered lasted 1400 ms, starting from the onset of the target stimuli. The preceding 100 ms period before target onset was taken as prestimulus baseline. Epochs in which amplitude exceeded -70μV - +70μV range were excluded.

Four Region Of Interests (ROI) made up by four electrodes were thus identified: Left Anterior, LA, (F7, F3, FT7, FC3), Right Anterior, RA (F4, F8, FC4, FT8), Left Posterior, LP (TP7, CP3, P7, P3), Right Posterior, RP (CP4, TP8, P4, P8). Amplitude value for each ROI was obtained by algebraic mean of single electrodes amplitude.

Data were analysed through repeated measure ANOVA. Mauchly's test was used to check sphericity assumptions and Greenhouse–Geisser correction for sphericity departures was applied when necessary (Geisser & Grenhouse,1959). Together with target *Type* (NN1, NN2, V-N, NC1 and NC2) and *Status* (whole vs broken) two topographical variables (reflecting ROI distinction) were included in the analysis: *Caudality* (anterior vs posterior) and *Laterality* (left vs right hemisphere).

6.3 Results

Behavioural Data

Only 7% of overall responses were incorrect ant thus accuracy was not further analyzed. Behavioural data were analyzed in two steps. In a preliminary analysis difference associated with Lexical status (words vs nonwords) and Status (broken vs whole) was investigated. Successively a more detailed analysis confined to words was made, by investigating differences among Types (NN1, NN2, V-N, NC1 and NC2). All analyses with behavioural data were carried with Mixed Effects Model (Baayen, 2007). Subject and words were included in analysis as random effects. Since stimuli were presented twice (in normal and broken condition) the trial position of items was included as random continuous variable). Although stimuli in each condition were presented in counterbalanced order, this variable was added as further control to rule out the influence of a practice effect.

Model 1 – Reaction Times (Words[15])

Initial model included variable Category as factor and Length and frequency of whole word and of constituents as covariate. Variables with the lowest | t | were removed in successive steps, until the model included only significant effects (see also par. 4.2). Observed reaction times are listed in table 6.2, while final model variables are listed in table 6.3. Table 6.4 shows random effects.

	Word– Mean Reaction Times (sd)	
Type	**Normal**	**Broken**
NC1	929.52(304.14)	1114.91(314.25)
NC2	876.42(279.96)	1019.64(294.34)
NN1	954.13(292.16)	967.36(269.96)
NN2	923.34(288.18)	1014.11(289.69)
VN	967.83(298.85)	1009.84(273.95)

Table 6.2 Mean Reaction Times. The table shows the observed Mean Reaction Times divided for Type and Status (enclosed in parenthesis the standard deviations).

[15] An initial mixed effect regression model was fit only for compounds, in order to investigate differences among categories correcting also for frequency of constituents. A correction of frequency of constituents won't be possible in a model including also NC1 and NC2, since they don't have, respectively, frequency for the second constituent and for the first constituents, but only the frequency of the embedded words. However, since neither of these two frequencies had a significant effect only models including all *types* are discussed.
A separate model was also fit excluding A-N and N-A compounds, in order to investigate if the observed effects could be strongly influenced by the presence of with these stimuli. No differences were found and thus all the analyses are referred to the whole set of stimuli.

Model 1 - Mixed Model fixed effects				
Variables	**Coefficients**	**Standard Error**	**t-value**	**pMCMC**
NC1_broken (Intercept)	6.94	0.09	78.27	< 0.001*
NC1_normal	-0.2	0.01	-15.83	< 0.001*
NC2_broken	-0.07	0.03	-2.07	< 0.02*
NC2_normal	-0.24	0.03	-7.34	< 0.001*
NN1_broken	-0.2	0.03	-6.17	< 0.001*
NN1_normal	-0.23	0.03	-7.02	< 0.001*
NN2_broken	-0.12	0.03	-3.85	< 0.001*
NN2_normal	-0.23	0.03	-7.06	< 0.001*
VN_broken	-0.19	0.03	-5.57	< 0.001*
VN_normal	-0.23	0.03	-7.00	< 0.001*
Frequency whole word	-0.03	0.005	-5.15	< 0.001*
Length whole word	0.003	0.007	4.49	< 0.001*
Trial number	-0.0004	0.00002	-19.66	< 0.001*

Table 6.3 Model 1 fixed effects. The table shows all significant fixed of Model 1. RT were logarithmically transformed before entered in the analysis

MODEL 1 - Mixed model random effects	
Variable	**variance**
Word	0.01
Subject	0.04

Table 6.4 Model 1 random effects. The table shows the significant random effects of model 1.

Effect of category on RT, as defined by the model, are showed in Figure 6.2.

Data in table 6.3 must be interpreted as follows: First row (NC1_broken**)** indicates the Intercept: i.e. the predicted value taken as reference level. In this model the reference is given by the expected reaction times when the lexical status is Nonwords and the status is Broken (e.g. *cocco drillo*). For the predicted RT associated to this reference level, all covariates (target length and Ntrial) are assumed to be equal to 0. Coefficients for levels of categorical variables other than the one taken as Intercept (so NC1_normal, NC2_broken, NC2_normal, NN1_broken, etc.) must be added to the Intercept in order to obtain the predicted value for the intercept. The H_0 of statistical test reported in table 7.2 is that the considered coefficient is equal to 0, and thus that there is no difference in the estimate between the reference level and the considered level. Coefficient for covariates (Frequency of Whole Word, Length of Whole Word and Number of Trial) are interpreted differently. Positive coefficient (Length of whole word in the model for, example) indicates that as the value assumed by continuous variable increases the predicted RT increase. Negative coefficients (Frequency and Number of Trial in the model) indicate that as the value assumed by the continuous variable increases the RT decrease. All p values listed in table 6.3 are not traditional p values. As pointed out by Baayen, Davidson and Bates (2008), within mixed effects model p value estimates may easily lead to type I errors. A more robust alternative is represented by estimated p values, obtained with Markov chain Monte Carlo (MCMC) simulations, thus **pMCMC** (that can be interpreted just as traditional p value see Baayen, Davidson and Bates, 2008 for some statistical detail on MCMC simulations). All pMCMC values listed in tab. 5.2, are not so informative, because the statistical test carried tell if the estimated RT for a given level (e.g. NN1_broken) are different from the reference level (the Intercept, NC1_broken). To investigate the presence of differential effects of Status (Broken vs Normal), pMCMC factorial contrasts were carried out. Contrasts showed that a reliable *Split Cost* was present in all categories [pMCMC < 0.001]: the RT for stimuli with Status broken were always higher than RT with Status normal.

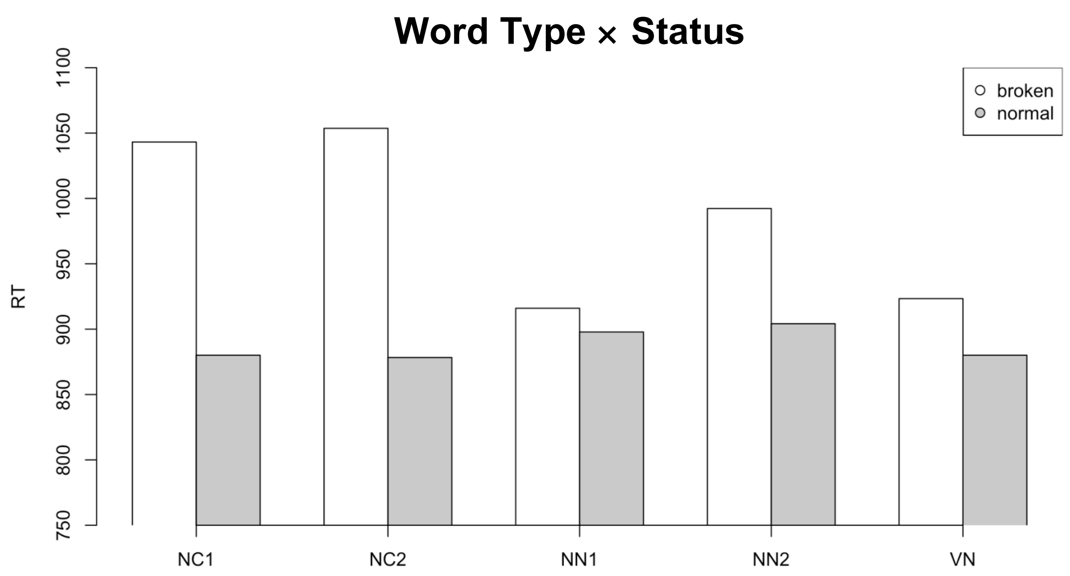


Figure 6.2 Word Type x Status.
The plot shows the Reaction Times across the different Type and status condition.
White bars indicate RT for Status Broken, Gray bars indicate RT for Status Normal

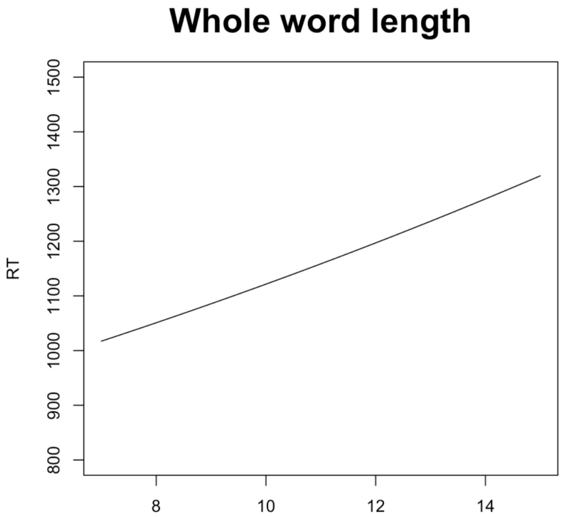


Figure 6.3 Word length effect
The plot shows the Whole Word length effect.

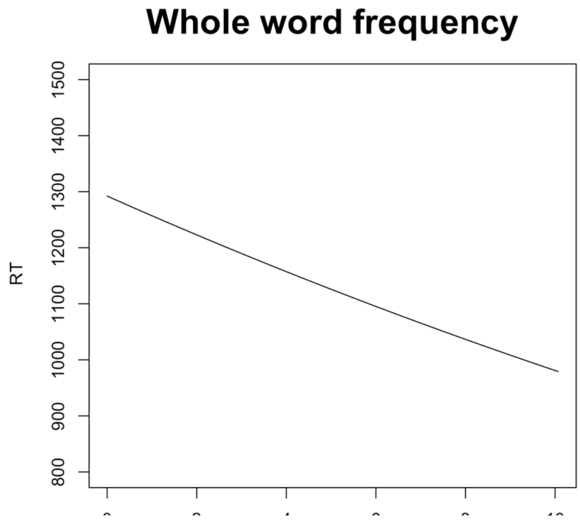


Figure 6.4 Word frequency effect
The plot shows the Whole Word frequency effect.

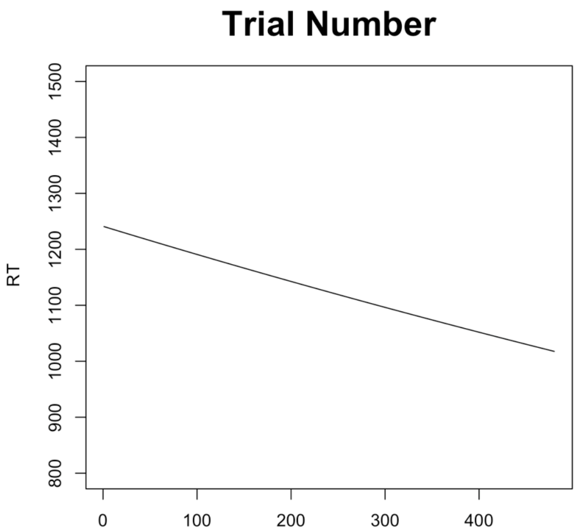


Figure 6.5 Trial number effect
The plot shows the Trial number linear effect.

Summarizing, results from model 1 were the following: A Split Cost was obtained for all categories, with Status broken, associated with Higher RT compared to Status Normal. Other psycholinguitic variables showed significant effect: Word length was correlated with RT (the longest the word, the longest the RT), while both frequency and Trial Number had a facilitatory effect. As the frequency increased RT decreased and as the Trial number increased RT decreased.

Model 2 - Compound Split Cost (only words)

Results from Model 1 indicate that the Broken condition was associated with higher RT for all categories, but don't tell if the RT increase was different across categories. In a second model the *Split Cost*, i.e. the increase of RT when the condition was Broken compared to the RT when the condition was normal. The *Split Cost* was calculated as the difference between RT when the *Status* was *broken* and RT then the *Status* was *normal* (RT Broken – RT Normal). Length, neighbourhood size, of both whole words and constituents and frequency of whole words were included as fixed effect covariates. Subject, words and trial number were included as random effects. Model was selected with the same procedure of Model 1 (see also chapter 4). Logarithm transformation of dependent variable was not necessary since the distribution of difference was normal.

Data in Table 6.5 and are thus expressed as raw RT differences. Significant fixed effect are listed in Table 6.6. No difference was found between the reference level (NN1) and 0, and no difference was found between NN1 and VN. Significant differences were found between NN1 and the other stimulus types (NN2, NC1, NN2). Further contrasts were carried out, in order to investigate differences among categories and obtain a complete picture of differences between Types in Split Cost. VN resulted significantly different fron NN2 [pMCMC = 0.01], from NC1 [pMCMC < .001] and from NC2 [pMCMC <.001]. NN2 were also differed significantly from NC2 [pMCMC <.001] and NC1 [pMCMC <.001]. Finally, NC1 had an significantly higher *Split Cost* than NC2 [p < .001]. A linear effect for frequency was found, with higher split costs as the frequency of whole compounds increases.

MODEL 2 – Mixed model fixed effects				
Variables	**Coefficients**	**Standard Error**	**t-value**	**pMCMC**
NN1 (Intercept)	-25.67	23.95	-1.07	0.28
NN2	66.78	22.23	3	0.003**
V-N	24.75	21.83	1.13	0.26
NC1	163.87	23.27	7.12	< 0.001**
NC2	127.31	22.97	5.54	< 0.001**
Frequency	9.02	3.87	3.39	0.02*

Table 6.5 Model 2 fixed effects. The table shows all significant fixed of Model 1. RT were logarithmically transformed before entered in the analysis

MODEL 2 - Mixed model random effects	
Variable	**variance**
Word	1908.3

Table 6.6 Model 2 random effects. The table the only random effect for Model 2.

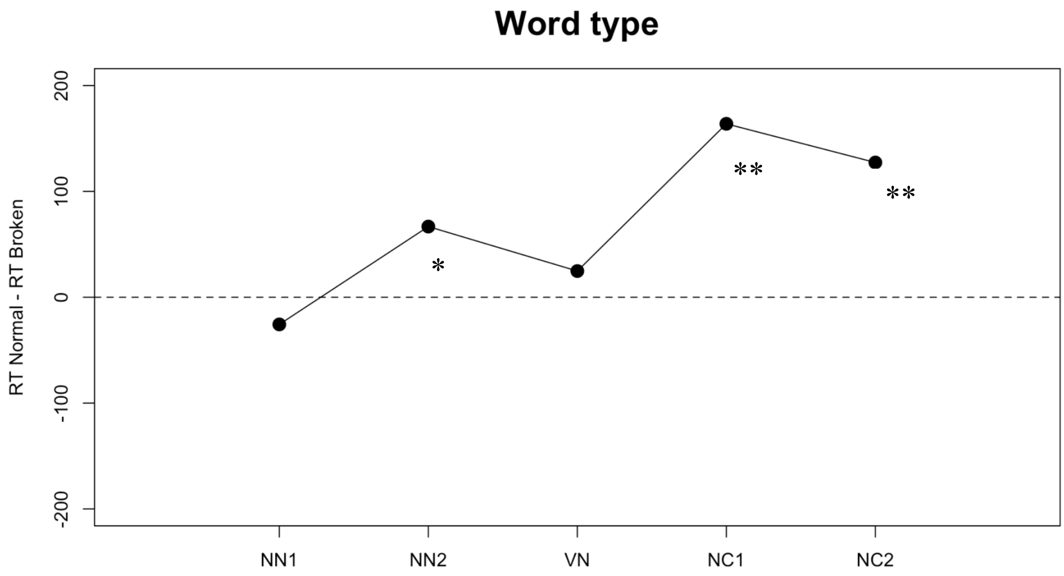


Figure 6.5 Word Type Split Cost
The plot shows the mean Split Cost for every Stimulus type, as defined by Model 2. Stars indicate significant difference from 0.

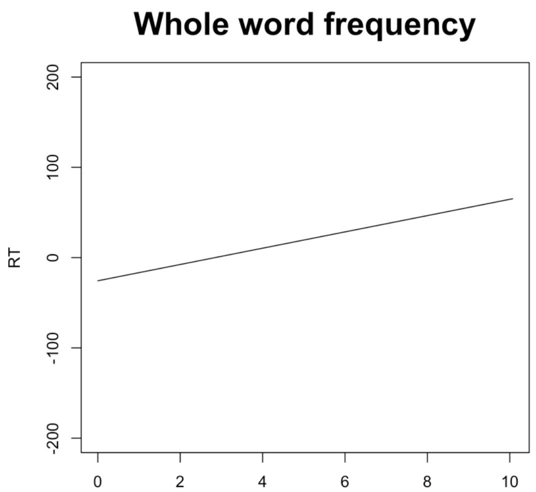


Figure 6.6 Whole word frequency effect
The plot shows the Whole Word frequency effect

Summarizing, results from Model 2 indicate that the effect of Split is somewhat negligible in NN1 (the expected value from NN1, i.e. the Intercept, is not significantly different from 0). NN1 and VN did not differ one from the other: *Split Cost* for these categories were smaller than NN2, NC1 and NC2. Furthermore NN2 had a smaller split cost of NC1 and NC2. NC1 had the highest split cost, compared to all other categories.

ERP data

The traces presented in Figure 6.7 and 6.8 show the grand average potentials recorded at the scalp (figure 6.8 shows enlarged images of electrodes Fz, Cz, and Pz). Target stimuli elicited the N1-P2 complexes followed immediately by a negative shift that appeared wider in the right hemisphere. Then an ulterior negative shift was started at about 400 ms and lasted until 600 ms poststimulus approximately (N400). These negative variation were followed by a positive shift (P600).

Analysis on ERP data were confined to five time windows, based from visual inspection and from Literature. The following time windows were thus identified: 0-130 ms for N100, 130-270 ms for P200, 270-310 for N2, 330-480 for N400 and 600-800 for P600. Trials with erroneous responses (7%) were excluded from the analyses as well trials with excessive artifacts (16%). Thus 77% of the trials entered the averaging processing.

N100 window (0-130)

In the first window the interaction Status × Caudality × Laterality was significant [$F(1,22) = 5.05$, $p = 0.03$], with higher amplitude for stimuli with Status normal, compared to Status broken in the Right Anterior Sites.

P200 window (130-270)

In the second windows the following effects were significant: Type [$F(4,88) = 2.68$, $p = 0.046$], Status [$F(4,70) = 0.04$], Caudality [$F(1,22) = 28.94$, $p < 0.001$], Type × Laterality [$F(4,88) = 6.01$, $p = 0.002$], Status × Laterality [$F(1,22) = 34,28$, $p < 0.001$], Caudality × Laterality [$F(1,22) = 74.58$, $p < 0.001$].

Contrasts showed more positive amplitude for NC2 and NN1 compared to NC1, NN2 and VN, but only in the right hemisphere. Status broken was also associated with lower amplitude compared to normal, but only in the right hemisphere. Overall amplitude was higher in the Left Anterior ROI.

N2 window (270-310)

In the third window the following effects were significant: Type [$F(2,88) = 4.48$, $p = 0.002$], Status × Caudality [$F(1,22) = 7.37$, $p = 0.01$], Type × Laterality [$F(4,88) = 4.033$, $p = 0.005$], Status × Laterality [$F(1,22) = 20.17$, $p = < 0.001$], Caudality × Laterality [$F(1,22) = 86, 51$, $p < .001$]. Status × Caudality × Laterality [$F(1,22) = 20.52$, $p < 0.001$].

The same trend observed in the previous window was found, with higher amplitudes for NC2 and NN1 compared to NC1, NN2 and VN in the right hemishphere. Furthermore Status Broken, compared to Status Normal, was associated with lower amplitudes in Left Anterior ROI.

N400 window (330-600)

The following effects were significant: Type [$F(4,88)=3.8$, $p = 0.01$], Caudalitity [$F(1,22) = 18.33$, $p < .001$], Laterality [$F(1,22) = 10.9$, $p = 0.003$], Type × Laterality [$F(4,88) = 5.25$, $p =0 .003$], Caudality × Laterality [$F(1,22) = 16.22$, $p < 0.001$], Type × Caudality × Laterality [$F(4,88)=3.60$, $p = 0.03$].

Contrasts showed higher amplitudes for NN1 and NC2 in the right anterior sites compared to NC1, NN2 and VN. Overall amplitude was higher in the left hemisphere, and the highest amplitude was recorded at right posterior sites.

P600 window (600-800)

The following effects were significant: Type [$F(4,88) = 7.97$, $p < 0.001$], Status, [$F(1,22) = 6.95$, $p = 0.02$], Caudality [$F(1,22) = 4.70$, $p = 0.04$], Laterality [$F(1,22) = 42.8$, $p < 0.001$], Status × Caudality [$F(1,22) = 6.7$, $p = 0.02$, Type × Laterality [$F(4,88) = 3.33$, $p = 0.03$], Status × Laterality [$F(1,22) = 23.7$, $p < 0.001$], Caudality × Laterality [$F(1,22) = 8.89$, $p = 0.007$], Status × Caudality × Laterality [$F(1,22) = 69.35$, $p < 0.001$].

In the left hemisphere, NC2 had higher amplitude compared to all other categories. In right hemisphere NC2 and NN1 had highest amplitudes compared to NC1, NN2 and VN. Each of these latter categories did not differ from the others.

A higher amplitude was observed in Right hemisphere compared to the Left

Hemisphere. In Right Hemisphere, only Status showed a significant difference, with higher amplitudes in posterior sites compared to anterior sites. In Left Hemisphere a different pattern emerged, with highest amplitudes for whole condition compared to broken condition and with higher amplitudes in posterior sites compared to anterior sites.

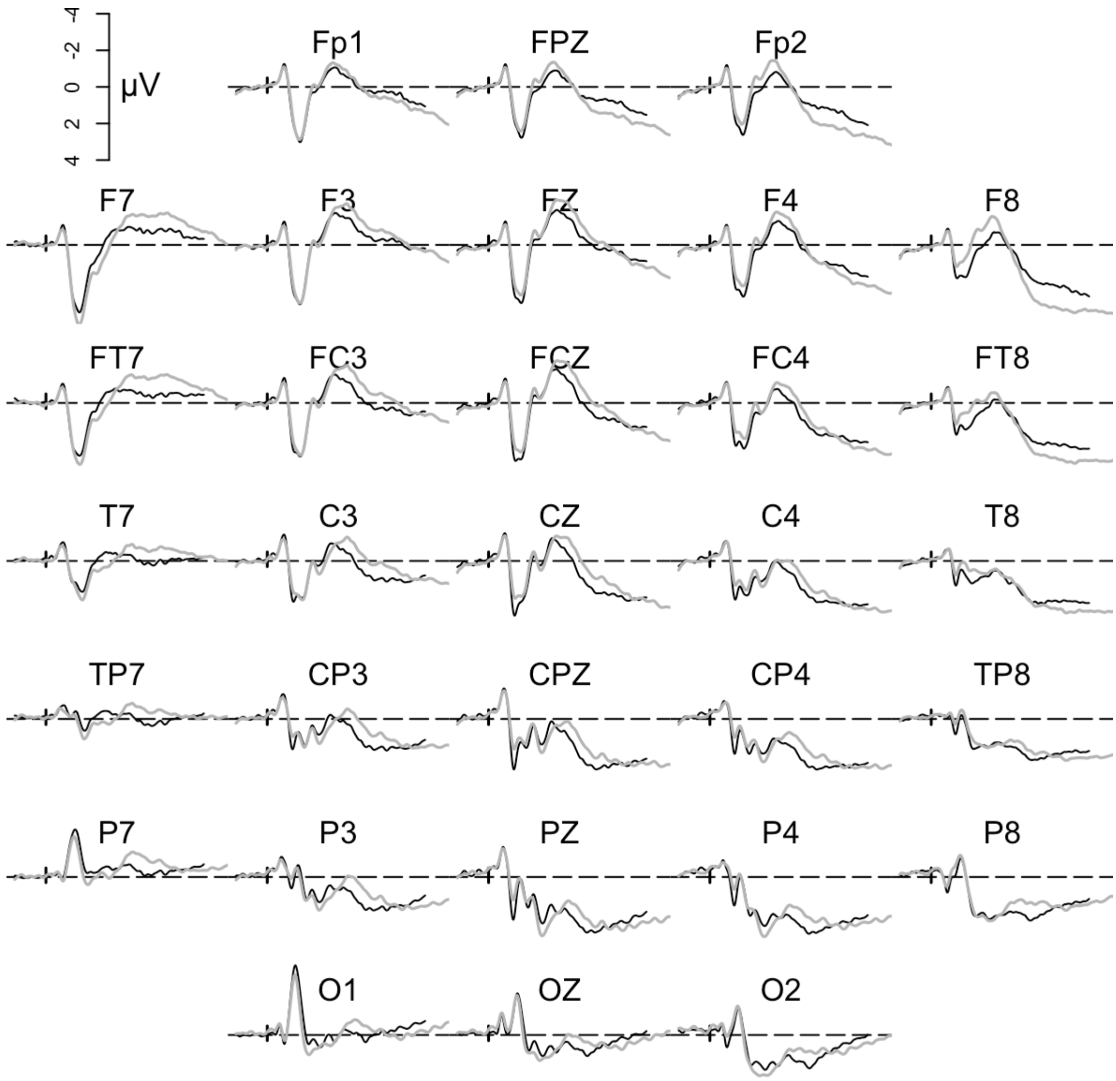


Figure 6.7 Normal vs Broken Grand average event-related potentials (ERPs) obtained for Normal (black) vs Broken (gray) status condition. Differences were found,starting from approximately 200 ms especially in Right Anterior (RA) ROI.

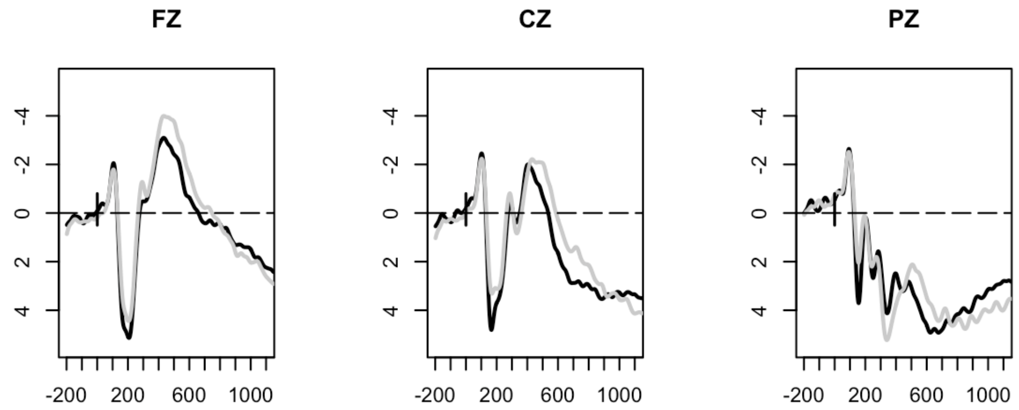


Figure 6.8. Normal vs Broken Grand average event-related potentials (ERPs) obtained for Normal (black) vs Broken (gray) Status. Highlights of midline electrodes.

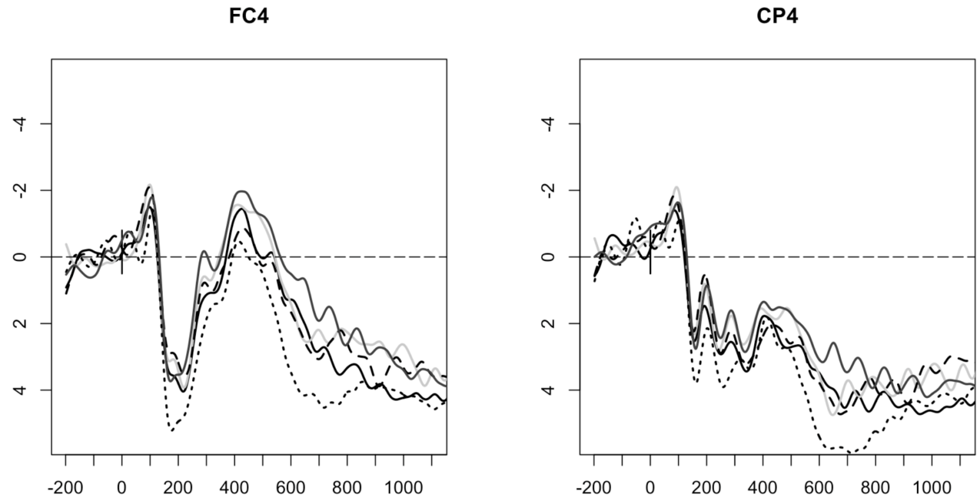


Figure 6.9. Stimulus type The plots show event-related potentials (ERPs) in FC4 and CP4 electrodes. Solid lines indicate compounds. Broken lines indicate noncompounds. Black for NN1; light gray for NN2; Dark gray for VN; Dashed line for NC1; Dotted line for NC2. Figures 6.11, 6.12 show separated plots for compounds and noncompounds

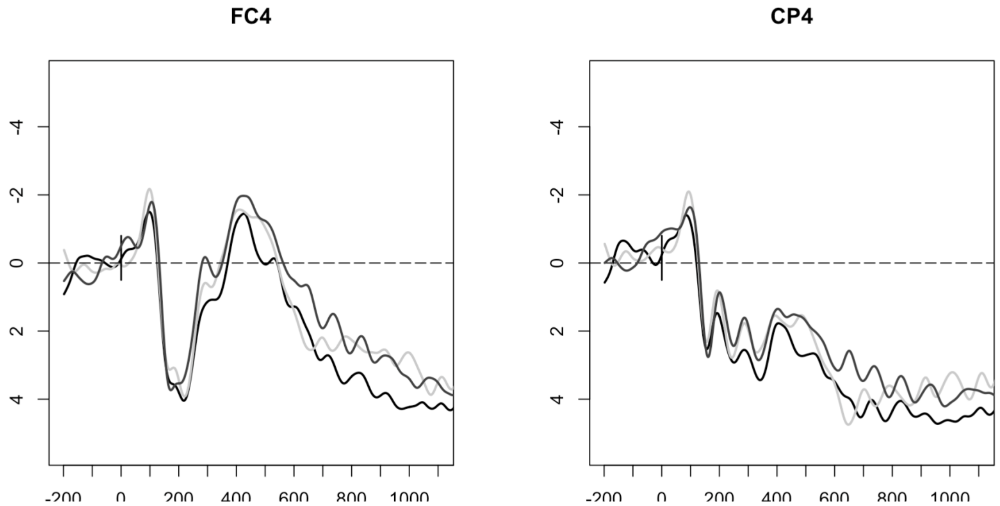


Figure 6.10. Compounds The plots show event-related potentials (ERPs) in FC4 and CP4 electrodes, as representative of right anterior ROI. Black for NN1; light gray for NN2; Dark gray for VN; Dashed line for NC1; Dotted line for NC2.

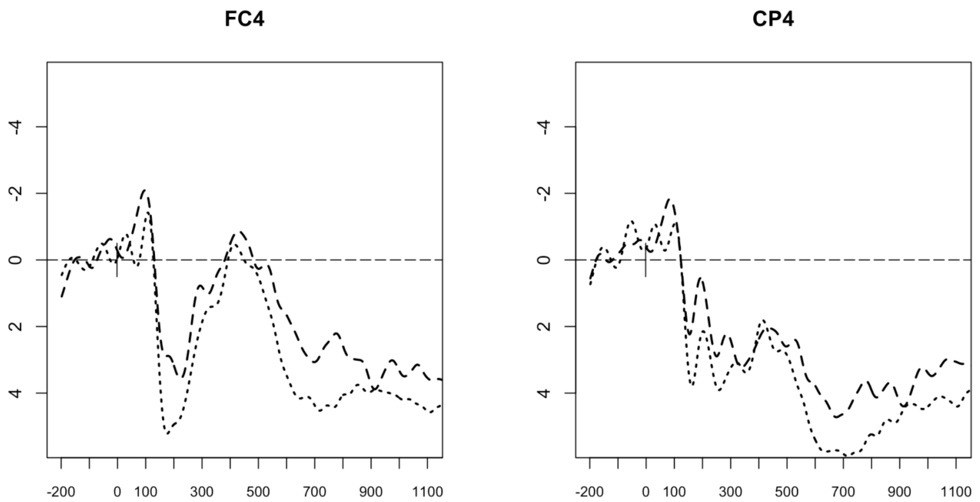


Figure 6.11. Noncompounds The plots show event-related potentials (ERPs) in FC4 and CP4 electrodes, as representative of right anterior ROI. Dashed line for NC1; Dotted line for NC2.

Results from ERP data can be summarized as follows: early processing of the stimuli was influenced by status. The greatest negativity for compounds with Status Broken in the N100 interval was presumably related to spatial attention deployment (Luck, 2005). From the second window considered (P200) a difference emerged across stimuli and remained almost the same across all the windows successively considered. Although the effect observed arose very early (in the time windows ascribed to P200) more probably it reflects it is related to the following negativity (N2). A visual inspection of brain potential shows indeed as these two peaks (P200 and N2) are often overlapped. The window for analysis of P2 and N2 (defined on the basis of visual inspection of grand average and on comparison with literature), for some electrodes captured partly both components. In other words, the smaller positivity of P200 is thus probably due to an higher negativity of N2 that, partially overlapping with P200 and that, in the averaging process, lead to a smaller positivity. A higher positivity was found for NC2 (noncompound words with a right embedded word, e.g. *tartaruga*) and NN1 (left headed Noun-Noun compounds) compared to other stimuli. In the last window considered (P600), NC1 (noncompound words with a right embedded word, e.g. *coccodrillo*). In almost all windows (except the one for N400) status Broken was associated with lower amplitudes, especially in the right hemisphere. No interaction was found between Type and Status, and thus the Status effect was additive: when the Status was Broken a higher negativity was found, with an increase that was irrespective of the Type of stimuli (NC1, NC2, NN1, NN2 or VN).

6.4 Discussion

The present experiment investigated, through the paradigm introduced by Libben et al. (2003) the mental representation and processing of different type of Italian compound words: left headed Noun-Noun compounds (NN1), right headed Noun-Noun compounds (NN2) and Verb-Noun compounds (VN). Stimuli were presented in two conditions: normal (written as a single word, e.g. *capobanda*) or broken (written as two words, e.g. *capo banda*). The assumption underlying this paradigm is the following: lexical decision in broken condition would be influenced by the way words are processed and represented in the lexicon. If words are normally decomposed during their lexical processing (or if they're represented in decomposed form), then a smaller *Split Cost* (the increase in RT when the status was broken) is expected.

NN1 and VN are generated according to the main word formation rules of Italian compounding with order of constituent that reflect order of Italian syntax. NN2 compounds are, on the contrary, produced according to an alternative rule (see par. 1.6). We expected an higher Split Cost for NN2 since their morphological structure may show a lower degree of "constituency" (the evidence of being a compound composed by two constituent), given that they represent an exception (compared to NN1), and thus to retrieve the correct information on lexical and semantic features an activation of whole word representation could be preferred.

The hypotheses were almost perfectly confirmed on the behavioural results: a smaller *Split Cost* was found for NN1 and VN, in comparison to NN2. Specifically, the *Split Cost* associated to NN1 and VN was negligible, with almost no difference in RT when the word was presented as whole or when the word was presented as broken. This suggests, as expected, that the morphological representation of NN1 and NN2 in which "constituency" (the evidence of being composed by constituents) is more evident. NN2, however, did not behave like noncompounds and their Split Cost was smaller to the noncompounds stimuli included in the experiment. This suggest that NN2 (right headed compounds) although are an exception for Italian compounding, and thus being good candidates to be listed in the lexicon as whole word, show a degree of morphological "constituency". This "constituency" is less evident in comparison to NN1 and VN, compounds formed in accordance to main Italian compounding rules. Interesting, a different pattern was observed in comparison to results from Libben et al (2003). We can notice qualitatively that in their experiment an appreciable *Split Cost* effect was found for all compounds while in the present study the effect with NN1 and VN was almost negligible. This results confirm what argued by Marelli et al. (2009a) that suggested that VN and NN1 are structures "imported" from Syntax: they me be conceived more like a juxtaposition of two words rather then truly morphologically complex words.

A final interesting insight comes from the linear effect found in Model 2: as the frequency of Whole Word increase, the split cost increase. This result is in line with *Dual Route models* (or multiple Route models) that account for a whole word representation (Baayen et al., 1997). As the frequency of a compound increase the more likely it would be accessed through the whole word representation rather than through decomposition: this would explain why, as the frequency of whole word increase the higher is the split cost, because an access through its constituent would be

less natural.
ERP results showed a different picture from Behavioural data. Compound status influenced very early brain potential (form the first windows) probably influencing the deployment of spatial attention (Luck, 2005). Status Broken was associated with overall lower amplitudes compared to status normal, especially in right anterior sites. In very early window (starting form 170 ms) a Type effect emerged especially in right Anterior sites. Interestingly, no interaction was found (in every window) between Status and stimulus Type: this means that the effect of Status was the same across all categories. Thus, brain potentials cannot be taken as indexes of morphological composition (since in NC1 and NC2 there are not compounds) but rather as indexes of activation of orthographic representations and of a general cognitive effort in accomplishing the task. Results showed more positive amplitudes for NC2 (*e.g. tartaruga* 'tortoise', where *ruga* 'wrinkle' is a word while tarta is not a word) and NN1 (left headed Noun-Noun compounds, e.g. *capogruppo,* 'group leader') compared to other stimuli (NC1, NN2, VN). In the following time windows the same effect found in the early negativity remained almost constant, thus suggesting that the differences observed in later windows are more carry over consequences of the different negativities observed in early time windows. What is the meaning of this early negativity?

In the same time window a greater negativity was found for words with high neighborhood size compared to word with smaller neighbourhood size (Holcomb, Grainger & O'Rourke, 2002). Holcomb et al. (2002) suggested that the highest negativity in the windows could represent higher lexical activation. In a reading task of Basque compounds, Vergara-Martínez et al. (2008) found a greater negativity in a similar time window (150-300) in compound with high frequency first constituent. They interpreted the results as suggesting a greater lexical activation triggered by the first constituent. Further insights on the potential meaning of this early negativity come from another experiment, by Dell'Acqua, Pesciarelli, Jolicoœur, Eimer and Peressotti (2007) a target word and a distractor in different colours appeared at the same time in the screen, one at the left and one at the right of a fixation point (thus similarly to the condition with Status Broken of the present experiement). The subjects had to perform a lexical decision only on the stimulus of a given color, ignoring the distractor in the other color. A modulation of spatial component N2pc was found when the distractor and the target were semantically related, with smallest N2pc when the two words were related. Thus, very early ERP component may be influenced by semantic relation between words, even when this relation is irrelevant to the task.

Along the same line of reasoning it is possible to hypothesize that the early N2 found in the present experiment reflect an early semantic activation of word segments.

However the effects observed for NC2 and NN1 could be related to different reasons to this early semantic activation. The experimental design is likely to have triggered a left-to-right analysis of the stimulus strings. Some subjects, indeed, spontaneously reported that a left to parsing reading was inevitable.

A first possible explanation reside on the different probability of occurrence of NC2 first segment. In NC2 the leftmost segment was a real Italian segment but, by itself, is not a real word (e.g., in NC2 stimulus *tartaruga*, the segment *tarta* is not a word). In all other stimuli (NN1, NN2, VN, NC1) the begining stimuli was always a word. Thus only 20% of trials begin with a segment that was not a word. Early negativities are seen for NC2 could reflect the fact that NC2 first segment are *deviant* stimuli that, as such, elicited higher N2 (Luck, 2005).

Results for NN1, instead, probably reflect the semantic activation of stimuli. According to results by Dell'Acqua et al. (2007), early components can be modulated by the semantic relation between word segments. It is important to notice that the topography component found by Dell'Acqua et al. (2007) was slightly different from the one observed in the present study (The N2pc is observed on posterir sites, while the N2 in this study is anterior), and that there are some remarkable differences in the task. In the present study no deployment of spatial attention resources on one side despite the other was required (in terms of target selection), and an explicit integration of word segments was required. Thus, Probably the N2 observed in this study reflect a earlier aspect of processing and the differences in topographical distribution. Dell'Acqua et al. (2007) found a smaller negativity for words when the distractor was a related word. Given that both word segments are activated, the greatest "constituency" of NN1 and the consequent strongest semantic connection could be the reason smaller N2 observed. The results same line of reasining could be extended to NC2: since there is no semantic activation of the first segment of an NC2 (that it is not a word), no semantic activation occurs and thus, no competition between the first and the second word segment.

However, the most important comparison is thus that of NN1 and NN2 compounds. According to the surface structure they have no difference: both are composed by two nouns and they have matched psycholinguistic variables. A left-to-right parsing mechanism could have been influenced by the probabilities of guessing the word after the first segment read (given the first segment, the probability that the second segment is the one presented, see par. 7.2.2, for further explanation on how these probabilities are calculated). A check on probability showed no significant differences between NN1 and NN2 [Wilcoxon W = 222, p = 0.25]. The differences observed in the early negativity N2 are thus probably related on the different morphological structure of NN1 and NN2, with NN1 that have a greatest semantic connection between constituent, thus supporting the hypothesis of a greatest "constituency" for this stimuli. The source of the absence of the same effect in VN compounds is difficult to be traced. VN compounds were unmatched in some variables to other stimuli, and a control for the effect of these variables was not possible in ERP analyses. Moreover verb may show different pattern of activation that may further complicate the picture (see for example Nobre & McCarthy, 1994). All these differences could have determined the greatest cognitive effort in early VN processing in this experiment.

Summarizing results of this experiment suggest that NN1 and VN have a different representation from NN2. This different representation is related to the different morphological origin of the words (according to the main compound formation rules for NN1 and VN; according to a peripheral compound formation rules for NN2, see par. 1.6). The different morphological origin influences the way in which words are represented in the lexicon. Specifically in NN1 and VN a higher degree of "constituency" is present, with stronger connection between the compounds and the single words that constitute the compounds (see fig. 6.10a and 6.10b). In NN2 a stronger connection with whole word is present, although a certain degree of constituency is present. This aspect can be seen particulary in the comparison between NN1 and NN2: in the formers an early activation of both constituents suggest a strongest relation between constituent compared to NN2.

7. EXPERIMENT 3 - Reading Compounds In Neglect Dyslexia: The Role Of Headedness

7.1 Introduction

Visuospatial neglect is a neuropsychological deficit that is usually a consequence of right hemisphere lesions. This deficit is characterized by attentional deficit in the controlesional side of the visual space.

Neglect dyslexia, is a deficit that may accompany (but also dissociate) visuospatial neglect (Bisiach, Vallar, Perani, Papagno and Berti, 1986; Bisiach, Meregalli and Berti, 1990). Neglect dyslexia is classified generally as a *peripheral dyslexia*, since the deficit on reading is considered not as consequence of a primary linguistic deficit, but as consequence of the impairment of other cognitive processes involved in reading, i.e. the visuospatial attention.

When neglect dyslexic patients are asked to read single words presented in the centre of their visual space they usually commit omissions, substitutions or insertions of the leftmost word portion. Interestingly, in some cases errors may be modulated by lexical variables of words (Behrmann, Moskovitch, Black & Mozer 1990; Arduino, Burani and Vallar, 2002; Cubelli and Beschin, 2005). Many studies found that words are read better than nonwords (Siéreff, Pollatsek and Posner, 1988; Behrmann et al. 1990). Within nonwords, apparent morphological structure may influence performance: neglect dyslexic reading performance is better when nonwords are made by the combination of a real word and a real suffix (Arduino et al., 2002). Moreover, also neighbourhood size seems to play an important role: words with more orthographic neighbours are more difficult to read than words with few neighbours (Riddoch, Humphreys, Cleton and Fery, 1990). In another study Arduino, Burani & Vallar (2003) found a dissociation between reading aloud and lexical decision performance in patient with Neglect dyslexia, with spared lexical decision performance that was influenced by morpho-lexical variables as for normal subjects.

These results generally support theories that assume "late selection" of attentional processing (Deutch and Deutch, 1963; Umiltà, 2001). If lexical variables may influence reading, then visuospatial attentional components might operate also at a later stage of processing, after information that fall in the "neglected" area have been processed and have undergone higher-level analysis such as lexical and semantic processing.

So far, only a few studies investigated compound reading in neglect dyslexia. Most of these studies were focused on neglect itself, rather than on the insights that neglect performance can give on lexical representation and processing. Patient E.S studied by Vallar, Guariglia, Nico and Tabossi (1996) showed a severe left neglect dyslexia when requested to read aloud compound words (e.g. *camposanto* 'cemetery', lit. 'field holy'), but was able to produce appropriate semantic associations to the compound as a whole (e.g. coffin), suggesting an activation of the information in the neglected area even without awareness. In the study Behrmann et al. (1990), the authors found an advantage in reading real compounds compared to false compounds made up by existing words, and thus suggesting the activation of a whole word representation in reading. The issue of headedness in Italian Noun-Noun compound reading has been recently in a study by Marelli, Aggujaro, Molteni, Luzzatti (2009b) with nine Neglect dislexic patients. Patients were asked to read existing compound words (*pescespada)* and pseudocompounds built from real in which the first constituent of a real compound was substituted with an orthographic neighbour (**pestespada*, lit. '*plaguesword', in place of *pescespada* 'swordfish', lit. 'fish word'). Stimuli were presented for 700 ms in the centre of the screen. Results showed two mains effects: compounds were read better than pseudocompounds and left headed compounds were read better than right headed compounds. These results were interpreted as suggesting that compounds have a structured whole word representation.

The aim of this experiment was to investigate the issue of compound headedness in neglect dyslexia reading, trying to replicate the effect found by Marelli et al. (2009b), i.e. the advantage of left headed compound compared to right headed compounds. Further psycholinguistic variables were introduced in all the analysis in order to rule out alternative explanations. Furthermore a list of V - N compounds (that are exocentric compounds, see par 1.3, 1.4) was included as different control condition.

7.2 Materials and methods

7.2.1 Participants

18 Italian-speaking participants (12 males and 6 females), suffering from vascular injuries confined to the right hemisphere, and affected by left-sided neglect took part in this study. They were all-right headed and free from linguistic disorders. Their mean age was 66.27 years, ranging from 50 to 89. Their mean education level was 8.72 years, ranging from 5 to 13. Neglect was diagnosed via administration of the Bells test, BIT Conventional (score range: 29-118/146) and BIT behavioural (score range: 9-57/81). On the basis of clinical findings and formal tests (line bisection and copy of drawings) their neglect was classified of the egocentric type in all cases. Neglect dyslexia was assessed via the administration of a preliminary list of single words. Table 7.1 shows lesion sites of all Subjects.

Participant	Age	Gender	Education (years)	Site of lesion
1	63	M	5	Temporal
2	63	M	10	Fronto-temporo-parietal
3	89	M	8	Parieto-occipital
4	78	F	8	Temporo-parietal
5	51	F	8	Deep parietal
6	60	M	8	Parietal
7	50	F	8	Temporo-parietal
8	69	M	5	Temporo-parietal
9	64	M	5	Deep parietal
10	53	M	13	Temporo-occipital, thalamic
11	68	M	12	Fronto-temporo-parietal
12	74	M	8	Parietal
13	81	F	13	Fronto parietal
14	86	M	5	Parieto-temporal
15	66	M	13	Fronto-parietal
16	63	M	8	Temporo-parietal
17	59	F	12	Fronto-temporo-parietal
18	56	F	8	Deep parietal

Table 7.1. Partecipants of the experiment.

7.2.2 Materials

The experimental items consisted of 88 Italian compound words: 28 left headed Noun-Noun compounds (**NN1**), 28 right headed Noun-Noun compounds (**NN2**) and 32 V-N compounds (**VN**). NN1 and NN2 were the same stimuli of El Yagoubi et al. (2008) (see APPENDIX 2). Stimuli types are listed in table 7.2

For all three categories the following variables of whole words were considered: Familiarity, Frequency, Age of Acquisition, Imageability, Length and Neighbourhood

size of Whole words[16]. For Noun-Noun all the same variables were considered also for constituents. However for VN constituents, only Frequency, Length and Neighborhood size were available.
Considering all three categories, stimuli differed in frequency [F (2,85) = 3.8; p = 0.02] with whole compound frequency of VN significantly higher than NN1. The Age of Acquisition was also different, with NN2 compounds acquired later than NN1 and VN [F(2,85) = 6.24; p = 0.003]. In relation to constituent variables, only Length of first constituent was different, with VN first constituents longer then [F (2,85) = 3.53; p = 0.03]. All other variables were statistically matched.
Because of the relevance for our study: differences on psycholinguistic variables in NN1 and NN2 constituents were further explored. The first constituent of compounds resulted matched for all these psycholinguistic variables, while second constituent of Left Headed was more imageable [F(1,54)=7.029, p<0.05] and acquired earlier [F(1,54)=10.725, p<0.05] than the second constituent of Right headed compounds.
A particular attention was devoted to control some problems intrinsic to neglect patient studies. In order to exclude that the results could be explained in term of probability of "guessing" the word, given the information of the second constituent several probability measures were computed.

Specifically, the possibility to guess the first component given the second component was also assessed in two different ways. First, the conditional probability of encountering in Italian (as represented in the used word corpus) a specific compound (e.g., *astronave*) given the second constituent (-*nave*) was computed. This probability was computed, according to the method described in Kuperman et al. (2008), as the ratio of two probabilities: the probability to encounter a given compound, estimated by the relative frequency of the compound, and the probability of encountering any compound ending with a given constituent. In the –*nave* example this last probability is represented by the sum of all relative frequencies of all compounds (e.g., *astronave*, *motonave*, etc.) ending with that constituent. These conditional probabilities were thus compared among the three categories (VN NN2, NN1). No difference was found among categories [Kruskal-Wallis χ^2 (2) = 3.46, p = 0.18]. Because of the relevance for the study a comparison between NN2 and NN1

[16] Psycholinguistic variable values for NN2 and NN1 come from El Yagoubi et al. (2008) (see charter 4). Data for V-N were obtained from three groups of ten subject each, with age and scholarity similar to those that made the ratings of N-N.

was made and no significant difference was found [Wilcoxon W =342, p = 0.41]. Second, a number of control subjects (N = 10) was given the list of the second constituents, and knowing that these items were part of a compound word, they were asked to guess the first constituent. There was no difference the number of correct guessing among categories [Kruskal-Wallis χ^2 (2) = 1.01, p = 0.6], and no difference was found between Left Headed and Right headed compounds [Wilcoxon W=325, p = 0.26] .

EXPERIMENTAL STIMULI		
Type	**e.g**	**translation**
NN1	*capobanda*	'band leader' (lit. 'leader band)
NN2	*astronave*	'spaceship' (lit. 'starship'
VN	*prendisole*	'sundress' (lit. 'take sun')

7.2.3 Procedure

Words were displayed on a 19 inches computer screen. All stimuli appeared in black on a white blackground with size 44, Arial typeface. Stimuli were presented to each participant in a different random order. Participants were tested individually. They were asked to read aloud the words present without any time limit. Once the word was read the examiner pressed a key to switch to the following word. When possible, each participant read the stimuli more the one time (in separated session and with different stimulus order).

Data were analyzed through mixed effects modeling. Subject and words were included as random variables, psycholinguistic variables as covariates and variable "Category" (NN1, NN2, VN) as fixed effects (see par. 4.2).
All psycholinguistic variables of stimuli (par. 7.2.2) were included as fixed effect covariates. Since some subjects saw the stimuli more than one time, the number of lists presented was also inserted in the model, as covariate, in order to evaluate the presence of a practice effect.
Accuracy on whole-word reading (expressed binomially) was the dependent variable.

7.3 Results

Overall performance of patients was good, with 13% of errors on total number of considered stimuli (4294). Only errors committed exclusively on the left component of Noun-Noun compounds were considered for data analysis. This lead to the exclusion of 68 errors, that either concerned the whole word (in the large majority) or the rightmost component. The final analysis concerned a total of 558 errors. Errors were classified as follows: Omission errors, Substitution errors (divided into phonological and lexiacl errors) and Non-classifiable errors. Omission errors consisted of complete or partial omission of the leftmost constituent (*audiofrequenza* → _____*frequenza*; *terremoto* → ___*remoto*; *affittacamere*→ _*fittacamere*). Lexical errors were substitutions of the whole left constituent, or of part of the left constituent, with another existing word, with a "semi word" or with some letters, so that the whole word is an existing word or the first constituent is a real world (e.g., *roccaforte* → *cassaforte*; *fangoterapia* → *fisioterapia*; *madrepatria* → **padrepatria*). Phonological errors: the substitution of a phoneme of the left constituent with another phoneme or the deletion or the insertion of a phoneme (e.g., *fangoterapia* → *fauloterapia*; *zootecnica* → *botecnica*). Not classifiable errors: other errors that did not easily fit into the previous categories (e.g. *bordovasca* → **lavasca*; *ceralacca*, → *malacca*).

Table 7.2 shows the percentage of errors across categories while Table 7.3 shows distribution of error types. All errors were merged for the statistical analyses, in order tho have a satisfactory number of stimuli for mixed models.

lexical category	Error percentages
NN1	8%
NN2	11%
VN	6%

Table 7.2 Error Percentages

lexical category	Error Types			
	lexical	omission	phonological	other
NN1	67	58	31	14
NN2	103	81	23	17
VN	62	70	12	20

Table 7.3. Error types across categories

Model 1. All categories included (NN1, NN2, VN)

Initial model included variable Category and all covariates. Variables with a | t | < 1 were removed in successive steps, until the model included only significant effects (see par 4.2). Table 7.4 summarizes the significant fixed effects of the final model. Table 7.5 shows the randm effects.

MODEL 1 - Mixed Model fixed effects				
Variables	**Coefficients**	**Standard Error**	**z-value**	**p**
VN (Intercept)	3.78	0.8	4.1	<0.001*
NN1	-0.15	0.21	-2.749	0.45
NN2	-0.53	0.19	-0.74	0.005*
Length whole	-0.18	0.06	-2.83	0.004 *
log Freq whole	0.15	0.05	3.2	0.001 *

Table 7.4. Fixed effects. Table shows the significant predictor of final mixed effects logistic regression. Dependent variable is expected probability of correct reading expressed in logit (All categories).

MODEL 1 -Mixed model random effects	
Variable	**variance**
Subject	1.88
Word	0.27

Table 7.5. Random effects. Table shows the significant random variables of final mixed effects logistic regression (All categories).

Lexical category (NN1, NN2, VN) was marginally significant (p = 0.06). An AIC (Akaike Information Criterion) comparisons of models with and without lexical categories, however suggested that the variables contributed significantly in prediction (χ^2= 7.24, p = 0.02). Both Indexes on goodness of fit considered were satisfying, indicating a good predicting value of the model [C = 0.84, Somers' D = 0.64].

Model is additive, i.e. every term must be added to calculate predictions. Importantly, any significant variable express a prediction in which the role of other variables has been ruled out (see chapter 4)

Results in table 7.4 must be interpreted as follows: Coefficient associated with VN is taken as reference level. Coefficient associated with VN (3.78) thus indicate, expressed as log odds, the probability of a correct answer when the category is VN.

Positive value indicate a probability higher than 0.5, while negative variable a probability of less then 0.5. For the theoretical probability associated to this reference level, all covariates (length and log frequency in the model) are assumed to be equal to 0. Coefficients associated to other factors (NN1 = -0.15, NN2 = 0.53), indicate the modulation in probability associated to this variables, compared to reference variable (in the model NN2).

For example, to obtain the probability of reading correctly NN2 the procedure is to add the probability reference level (VN=3.78), to the "modulation" of probability of the category (NN2=-0.53). Thus the probability of a correct answer for NN1 is (3.78 - 0.53 = 3.25), assuming all covariate values to be equal to 0. The same could be made for NN1. Significance, indicated by p-value, signals if the coefficient is significantly different from 0, and thus if it is different from the reference value.

With respect to factors, the negative coefficient for NN1 and NN2 indicates, that, respect to the reference variable VN, both showed a lower probability of correct reading, although only NN2 showed a significant lower probability of correct reading.[17].

Covariates, as continuous variables, must be interpreted in a slightly different way. A positive coefficient indicates that as the variable increase, the probability of a correct answer increases as well (e.g. log frequency of whole words = 0.15) and a negative coefficient indicates that as the variable decreases, the probability of a correct answer decreases (length of whole word = -0.18). To calculate the prediction for a given word, the coefficient must be multiplied by the value of the variable for that word. Thus, for example, to a word with length 7 will be associated a probability of 7 × (-0.18) = 1.26. Data in table 7.5 indicate the variance associated to every random variables that has been taken into account within the model.

The model is additive and, as such, the contribution of all variables must be summed to obtain the final prediction. The effect found can be summarized as follows:

1) an effect of compound category was found, with VN and NN1 are read with a higher accuracy than NN2 (see figure 7.1).

[17] When model was refit with NN2 as reference model, a mirroring pattern was found, with a trend on significance in the comparisons between NN2 and NN1 (p =0.068), and a significant difference between V-N and NN2 (p=0.005).

2) as the length of compound increases, the accuracy decrease (see Figure 7.2). 3) as the frequency of the compound increases, the accuracy increases. (see Figure 7.3). All other variables (age of acquisition, familiarity, imageability of whole compound, and of single constituents, length and frequency of single constituent length age of single constituents, conditional probability of guessing the second constituent given the first) had a negligible effect and were not included the final model.

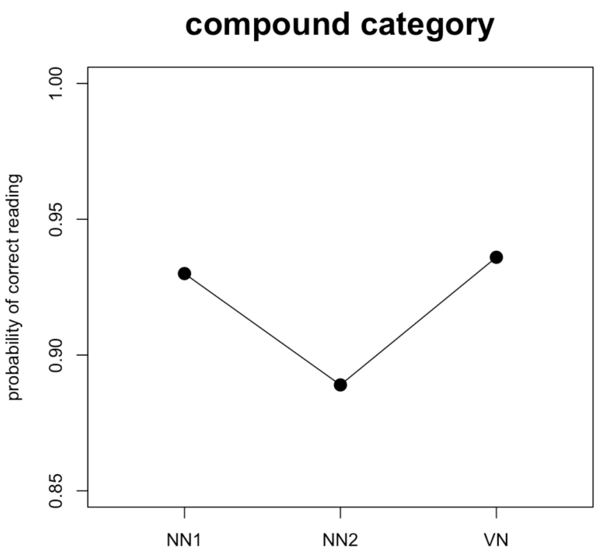


Figure 7.1 Compound category effect
The plot shows the probability of correct reading for different stimuli categories.

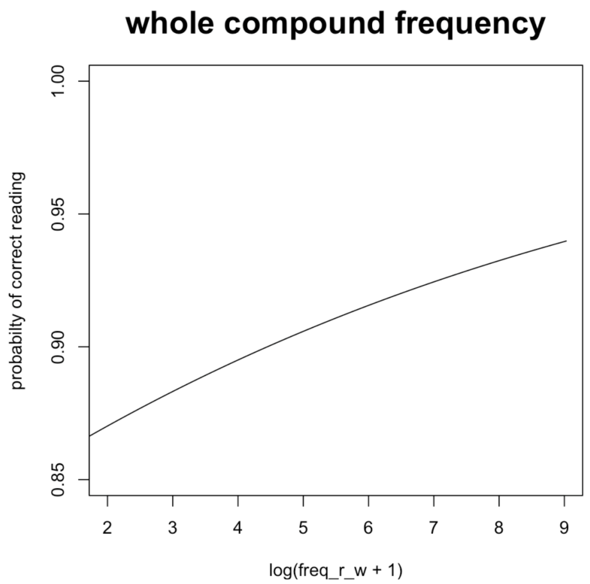


Figure 7.2 Whole compound frequency effect.
The plot shows the probability of correct reading for different stimuli lengths

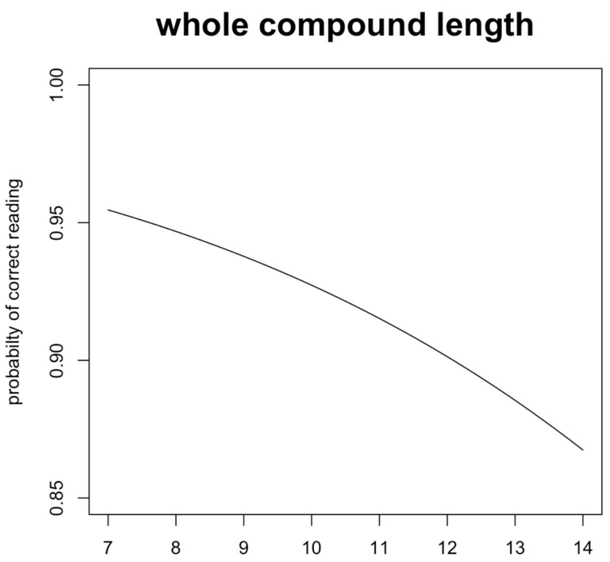


Figure 7.3 Whole compound length effect.
The plot shows the probability of correct reading for different stimuli lengths

Model 2. Only Noun-Noun compounds (NN1, NN2, VN)

Since not every psycholinguistic variable could be inserted in the model with VN, another mixed model in which only N-N compounds (NN1, NN2) were included was fit, in order to have a finest control of psycholinguistic variables.

Initial model included variable Category and all covariates. Variables with | t | < 1 were removed in successive steps, until the model included only significant effects (see par 4.2). Significant fixed effects are listed in table 7.6, while random effects are listed in table 7.7.

MODEL 2 - Mixed Model fixed effects				
Variables	**Coefficients**	**Standard Error**	**z-value**	**p**
NN1 (Intercept)	0.41	0.76	0.54	0.58
NN2	-0.57	0.21	-2.64	0.008 **
Freq whole	0.17	0.06	2.79	0.005 **
Fam first const	0.25	0.11	2.21	0.03*

Table 7.6. Fixed effects. Table shows the significant predictor of final mixed effects logistic regression. Dependent variable is expected probability of correct reading expressed in logit (Only Noun-Noun compounds).

MODEL 2 - Mixed model random effects	
Variable	**variance**
Word	0.29
Subject	2.01

Table 7.7. Random effects. Table shows the significant random variables of final mixed effects logistic regression (Only Noun-Noun compounds).

The significant predictors of the final model were group, Frequency of the whole compounds and the familiarity of the first constituent.

The difference between NN1 and NN2 is confirmed, with a lower predicted probability of correct answer for NN2 compared to NN1.

The effect of whole frequency in model 2 mirrored that obtained in the model. In model 2, an effect of the familiarity of first constituent emerged: as the familiarity of the first constituent increases the predicted accuracy increases. Again, the indexes of goodness of fit indicate a satisfactory model [C = 0.84, Somers' D = 0.64].

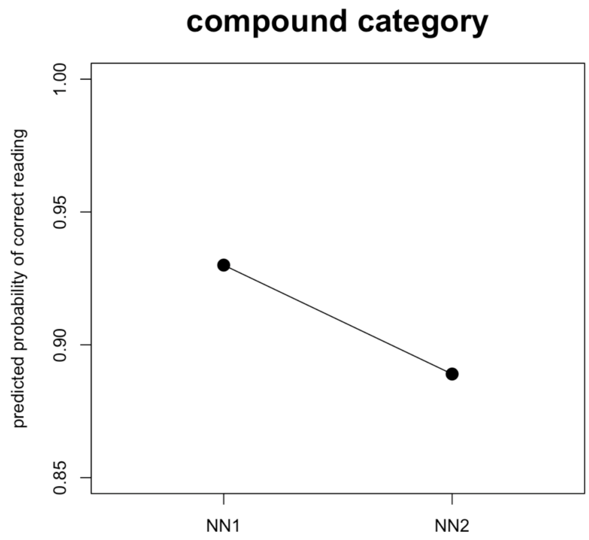


Figure 5.4 Whole compound category.
The plot shows the predicted probabilities for different stimuli categories

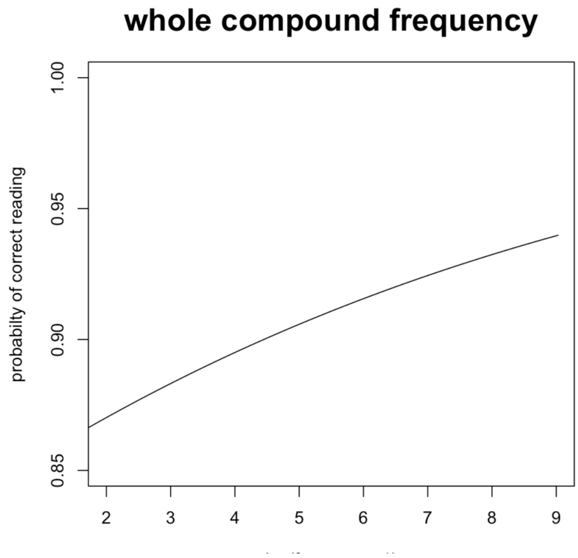


Figure 5.5 Whole compound frequency effect.
The plot shows the predicted probabilities for different stimuli frequencies

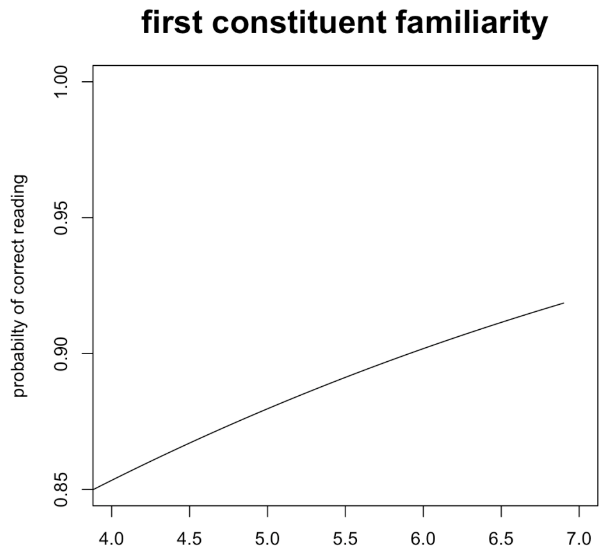


Figure 5.6 First constituent familiarity effect.
The plot shows the predicted probabilities for different first constituent familiarity

7.4 Discussion

Analysis on error types, although only qualitatively, is able to give some preliminary important insights (see table 3). The majority of errors committed by participants were lexical errors or omissions of the leftmost constituent rather then phonological errors, as expected from literature (see par. 7.3). An explanation for the presence this odd distribution can be found in relation to the distintion between *addressed phonology* and *assembled phonology* (Làdavas, 1988). With the term *addressed phonology* is indicated the phonological form retrieval through the routine reserved to already known words while the term *assembled phonology* indicate the phonological assembling through sublexical mechanisms like grapheme-to-phoneme conversion. It has been suggested that in neglect dyslexia, *addressed phonology* is less affected than *assembled phonology* (Làdavas, 1998). Lowest number of phonological errors seems to reflect this preference for lexical routines, whereas sublexical routines were recruited only when the first failed.

Results from mixed model give the most important information. In model 1 a length effect was found: as length of stimuli increase the accuracy decreases. This effect is not surprising, since stimuli were centred in the screen and increase in length meant more letters on the neglected side of the visual space. These can be particularly true for the longer VN. Both in model 1 (all compounds) and in model 2 (only Noun-Noun compounds), an effect of whole word frequency emerged: as the frequency of whole word frequency increase the overall accuracy increases. These results are in line with previous results found in literature and strongly support that lexical factors may influence reading in neglect patients (Riddoch et al., 1990). Results discussed so far could be easily interpreted as suggesting *only* a whole word activation of compounds that may compensate for the attentional deficit. This however seems not to be the case. First of all, in model 2 (when only N-N compounds were included) the familiarity of the first constituent was a significant predictor. These strongly suggest that words are decomposed, and that the access to the first constituent occurs. Moreover a lexical category effect was found in both model: NN2 words (e.g. *astronave*, 'spaceship') were read with a lower accuracy compared to NN1 (e.g. *capobanda*, 'bandleader, lit.leader band') and VN (*aspirapolvere,* 'vacuum cleaner'), while NN1 and VN were read with a similar accuracy. One could argue that the effect found is negligible in terms of effect dimension (a predicted probability of .88 for NN2, against a predicted probability of .93 and .94, respectively for N-N and VN). These findings however, converge with a very recent experiment by Marelli et al. (2009b), who showed a headedness effect in six neglect patients in a timed condition. Their results for NN1 and NN2, even in terms of accuracy. Different results associated with different lexical categories suggest the following conclusion: an access to the structural information of the compounds might have occurred. So, a whole-word access (within a *full listing* theory), although able to explain the length and frequency effects of whole word, is not capable to explain the difference between categories and effect of first constituent familiarity. Hence the results suggest both a decomposition of words and an access on structural information on whole compounds. This is compatible with *dual route* or *late selection models*. Difference between NN2 and NN1 supports the presence of a difference in processing, related to head position, that seem to modulate attentional resources allocation. Results by El Yagoubi et al. (2008) suggest that Italian NN2 may require a greatest amount of attentional resources (Kok, 2001), or, within the framework of context-updating theory of Donchin & Coles (1988), an update of information when an unusual position of head is encountered. Interepretation of El Yagoubi et al. (2008), focuses on the canonicity of left headedness in Noun-Noun Italian compounds (as seen in par. 1.6 the main productive rule of Italian Noun-Noun compounding is left headed, and in par. 1.7 we have defined that this productivity rule belongs to the *centre* of word formation rules, in contrast to right headed compounding, that belongs to the *periphery* of word formation rules). The difference found between NN2 and NN1 however offers several interpretations.

The hihest number of error in the leftmost constituent of NN2 compounds compared to NN1 compounds could be interpreted either as a bigger "attentional saliency" of the rightmost constituent in NN2 compounds (that lead neglecting the first one), or as a bigger "attentional saliency" of the leftmost constituent of NN1 (thus compensating in part the neglect effect), or as a combination of both effects.

One could in part disentangle the issue by the comparison with VN compounds. The analogy in performance between VN compounds and NN1, couldn't be interpreted by the analysis of the lexical categories of the leftmost constituents that is a Verb in the first case and a Noun in the second. NN1 and VN analogy can be explained in two ways. As first explanation, it's possible to hypothesize the presence two different mechanisms acting for different categories. The higher accuracy for NN1 could be related to an allocation of attentional resources towards the head constituent, while the higher accuracy for VN could be related to an automatic allocation towards the Verb constituent. A more parsimonious explanation for this similar behaviour can be found hypothesizing a strict analogy between VN and NN1. As seen in par.1.5, according to Di Sciullo (2009) VN and NN1 compounds in romance language share the characteristics of being generated in the *syntactic plane* (DS) and then transferred to the *morphological plane* (DM). This feature is motivated by the analogy between compound word constituent order and syntactic order (together with other evidence). Moreover an analogy between VN and NN1 has been already claimed in Marelli et al. (2009a), that suggested that these compounds, being "imported" from syntax would have a flat structure, while NN2, truly morphological structure would have a hierarchical structure of elements and thus asymmetry between head and modifier.

Both the results of Marelli et al. (2009b) and of the present experiment however are not able to the tell the whole story. They suggest respectively, that there is a difference between NN2 and NN1, and that there is difference between NN2 compared to NN1 and VN, with a lower accuracy for NN2 in both the experiments. Thus another possible explanations could be found in the fact that NN2 represent just an anomaly compared to other compounds in general. Firstly, N-N compound could be crystallized structure, in which the morphological constituency is not as clear as other compounds. This however seems not to be the case, since often patients respected boundaries between constituents, thus indicating a recognition of word structure. As seen in par. 1.6, Italian NN2 compounds can be considered as belonging to the *periphery* of Italian word formation rules, and as such represent a less common linguistic structure. The reason of the difference in NN2 and NN1 could be found in this difference. Why this differences could lead to a different attention to neglected side. As already pointed in Experiment 1 and Experiment 2 a possible explanation reside on the different "constituency" associated to right headed compounds. Compared to left headed compounds, the less evident morphological structure could make more difficult to access to the representation of the word. This interpretation is in line to what suggeste by Arduino et al. (2002) that suggest that morpholexical information could facilitate reading in neglect dyslexia.

8. EXPERIMENT 4 - Lexical And Buffer Effects In Reading And In Writing Noun-Noun Compound Nouns.

8.1 Introduction

Neuropsychological studies on compounds are mainly on picture naming (see par. 2.5), and only a few studies are available on the production in reading aloud or writing on dictation. Thus, little is known on compound processing in such tasks. Most of these studies concerned patients affected by phonological dyslexia/dysgraphia, who produce omissions, insertions or substitutions of affixes (e.g., Badecker & Caramazza, 1987; Luzzatti, Mondini, & Semenza, 2001; Hamilton & Coslett, 2007). This deficit is interpreted as reflecting the predominant use of the lexical routine as opposed to sublexical processing (e.g. phoneme-to-grapheme conversion). Processing complex words via the lexical routine implies however two alternatives. The first alternative (Taft and Forster, 1976) is that morphologically complex words are de-composed into morphemes before lexical access; the "full-listing" alternative (Butterworth, 1983) implies, instead, that all lexical entries are listed in whole-word form. However, hybrid models have been proposed whereby both types of lexical representation, whole-word and de-composed forms are activated in parallel: distributional properties may then determine who wins the race (Baayen, Dijkstra, Schreuder, 1997). Insofar omission and substitution errors on single components characterize their reading and writing of compounds, phonological dyslexic/dysgraphic patients have provided overwhelming evidence in favour of decomposition. However little is still known about how de-composition processes and whole word access interact with post-lexical processes.

The production of compounds, as for simple words, does not depend only on the mental representations or on the processes carried out within the mental lexicon, but also on more peripheral events that can modulate central lexical mechanisms and affect the speaker's performance. The role of *buffers* in reading and writing processes, and in particular the existence of a common graphemic buffer for reading and spelling have been the subject of several reports (Caramazza, Miceli, Villa & Romani, 1987; Badecker, Hillis & Caramazza, 1990; Caramazza, Capasso, Miceli, 1996). Within a cognitive architecture, the *buffer* component is conceived as a working memory system that contains the representation that specifies the abstract letter identities and the sequence in which the graphemes in a word appear. This storage system has limited capacity and holds information till further processes come into play.

The buffer component plays a special role in reading and writing and because of its limited capacity the buffer is sensitive to the verbal stimulus length. Thus, a word length effect on reading or on writing is taken as an indication of a disorder in the phonological or the orthographic buffer, respectively (Caramazza et al., 1987). According to Ward's model (Ward, 2003) there could be separate phonological input and output buffers since the input phonemic code (i.e. the word to be spelled) is different from the output phonemic code (i.e. the letter names). The model, however, contains a single orthographic lexicon for both reading and spelling and a common orthographic buffer.

In their landmark study, Caramazza et al. (1987) proposed a full set of criteria to identify the locus of a deficit in the graphemic output buffer: (a) the buffer has a limited space capacity and errors should be quantitatively and qualitatively identical in all types of tasks, irrespective of the input or output modality, as the graphemic buffer is involved in each of these tasks; (b) errors should mostly consist in graphemic deviations from the target (i.e., substitutions, deletions, additions or transposition of letters); (c) errors should appear in both familiar words and in non-words. Thus, the presence of errors should not be affected by the lexical category of stimulus words or by their morphological and semantic features.

However, these criteria are not always met in case reports in the literature, where further contributions were added. Morphological features were in fact found to affect writing in individuals with graphemic buffer impairment. For example, a case (Annoni, Lemmay, de Mattos Pimenta & Lecours, 1998) of a French-speaking woman with acquired dysgraphia, whose deficit could be located at the level of the graphemic buffer, showed spelling errors more frequently with irregular than regular words, although the qualitative type of errors was the same in both categories. The authors discussed these findings in terms of a post-lexical sensitivity to irregular spelling. Representations of irregular words would require a special need for attentional resources by the graphemic buffer level: when focussing attention on the irregularity becomes necessary, this can cause a detriment for the surrounding graphemic constituents. Importantly, and relevant to the present investigation, is the fact that the graphemic buffer plays a role also in the input part of reading (Caramazza, Capasso, Miceli, 1996; Hanley & Key, 1998) in which graphemic representations are input for word recognition that are applied in parallel over the whole graphemic string. Moreover, if damage to the buffer were to interact differently with those processes that involve serial processing (spelling) from those that involve parallel processing (reading) of the graphemes held in the buffer, one would expect damage to this mechanism to have different consequences for reading and for spelling.

The interaction between linguistic performance and short-term memory capacity is very clear when considering patients with general cognitive decline (e.g., patients with senile dementia of Alzheimer type, SDAT). A recent study (Chiarelli, Menichelli & Semenza, 2007) compared the performance of a group of SDAT with a group of aphasic patients and a picture naming task was used to compare the performance of the two groups. Patients with dementia showed omissions and substitutions of the second component of compound words more frequently when they produced compound paraphasias, whereas aphasic patients produced most of their errors on the first constituent. This study, which was the first that examined SDAT in naming compound nouns, highlighted the emergence of processes intervening during compound retrieval. In particular, the detriment of the second component seemed to reflect a clear position effect across types of compounds. According to Chiarelli et al. the second portion may be more sensitive to processing overload and thus pose specific problems to SDAT.

The production of compounds in a deficit of the graphemic buffer was studied by Badecker, Hillis and Caramazza (1990). Their patient (DH) wrote compounds much better than length-matched mono-morphemic words. The authors suggested that in DH compounds had an advantage at the level of the graphemic output buffer since in this temporary memory compounds can be stored in de-composed form, i.e. as smaller units that are not greatly affected by the weakening buffer capacity. Furthermore, whereas in the mono-morphemic words the spelling errors were distributed on the final part of the nouns, the errors on compounds fell in the final positions of both the first and the second component. This finding seems to indicate that compounds pass from the orthographic lexicon to the graphemic output buffer in morpheme-sized units. This would lead to the conclusion that composition happens in the buffer rather than in the lexicon. However, the process must somehow be driven by information about the whole word contained in the lexicon. This said one cannot but observe that a theory about the functioning of the buffer system and its relation with earlier lexical processes is far from being completely outlined. The study of the performance of two Italian-speaking patients, one dyslexic and one dysgraphic, described here, contributes to the understanding of compound processing by highlighting the interaction among compositional processes, whole-word access, and the capacity of the graphemic buffer.

Material

A list of Noun-Noun compound and non-compound words and non-words was prepared and administered in reading aloud, in writing on dictation and in repetition. The list was composed by 24 Noun-Noun compounds (NN: 12 left-headed, NN1, e.g., *pescespada*, 'swordfish'; 12 right-headed, NN2, e.g., *videogioco*, 'videogame') and 24 noncompounds (NC), mono-morphemic nouns that contained an embedded word homograph and homophone to a word, either in the initial position (NC1, e.g., *pellegrino*, 'pilgrim', where *pelle*, 'skin' is a real word while *grino* is a non-word) or in the final position (NC2, e.g., *pavimento,* 'floor', where *pavi* is a non-word while *mento*, 'chin' is a word). The embedded word was not related in meaning to the whole word. Twelve four-syllable NC with a phonological structure resembling that of real compounds (e.g., *damigiana*, 'demijohn') were used as control fillers. Experimental non-words were created by exchanging the position of either the two morphemes of NN (e.g., $pesce_1spada_2$, [lit.] 'fish sword', became the non-word $spada_2pesce_1$) or the two parts of NC (e.g., the $pelle_1grino_2$, 'pilgrim', became the non-word $grino_2pelle_1$). All the words of the experimental material are reported in Appendix 1. These items have been selected from the material used in a recent paper on NN Italian compounds (el Yagoubi et al., 2008) Age of acquisition, Familiarity, Frequency, Imageability and Length (i.e., number of letters) were calculated or collected through questionnaires for the two categories of compounds and the two categories of NC. Age of acquisition (AOA), Familiarity, Imageabilitywere collected via three different groups of 29 Italian speakers for each variable. The subjects judged each item on a 7-point rating scale. Age of Acquisition was different between categories [$F(3,44)=3.42$, $p=0.02$]. Contrasts showed that NN were acquired significantly later than NC [*p*<.001]. Moreover, within compounds left-headed NN resulted to be acquired earlier than right-headed NN (*p*<.001). Familiarity was not different across categories [F(3,44)=1.58, n.s.]Frequency did not show any effect as well [F(3,44)=1.67, n.s.]. No difference was found for Imageability [F(3,42)<1] and for Length, [F(3,44)=2.71]. Table 1 shows the mean values for each variable across categories (see APPENDIX 4).

8.2 Case 1: A patient with phonological dyslexia

RF was a 31 years old, right-handed, Italian-speaking man, with eight years of education. In 2006 (one year before the present study) he underwent surgical treatment for an angioma bleeding into the left fronto-temporo-parietal region. The aphasia examination (Italian version of the AAT, Luzzatti, Wilmes, , Table 2) revealed non-fluent spontaneous speech and spared comprehension. On the ENPA (Esame Neuropsicologico dell'afasia, tr. Neuropsychological examination of aphasia, (Capasso & Miceli, 2001) RF showed spared repetition of both words (100% correct) and non-words (100% correct), whereas in reading performance with words (11/13, 85% correct) was better than with non-words (0/15); he also demonstrated a grammatical class effect whereby nouns were read better than adjectives (15/20, 75% correct and 6/20, 30% correct), verbs (8/20, 40% correct) and function words (7/20, 35% correct). RF read simple words (i.e. in the citation form) better than inflected words (16/20 correct 80% vs. 9/20, 4. Unfortunately, his writing proved hard to investigate. The patient, who used the non-dominant left hand and had additional writing apraxia problems, could not produce but a few all-formed scribbles that were not easily interpretable. For this reason this investigation concerned only repetition and reading.

	Raw score	*PR*	T	Deficit
Token	7	92	64	Slight/minimal
Repetition	145	94	66	Slight/minimal
Written language	86	96	67	Slight/minimal
Naming	115	100	80	Minimal
Comprehension	120	100	80	Minimal

Table 2. Achener Aphasia Test of patient RT.

8.2.1 Procedure

RF was required to repeat and to read aloud the list of experimental stimuli made up of words (i.e., compounds and NC) and non-words described in the material session. RF's performance was tape-recorded and later analyzed in order to assess reading strategies for each type of item.

8.2.2 Results

RF's repetition was flawless for whole experimental material: compounds, NC and all non-words.

RF read aloud correctly 51/60 (85%) words and only 19/60 (46.3%) non-words [$\chi 2(1)= 35.1$; $p<.001$] thus confirming the classical pattern of phonological dyslexia. His errors were mostly substitutions or omissions of letters, especially in the final half of the stimuli.

Within words, RF read all compounds flawlessly (24/24, 100% correct), but he was less effective in reading NC (17/24, 70.8%), [$\chi 2(1)=8.19$; $p<0.01$]. With matched-for-length NC, as well as with non-words, besides errors, RF's production was effortful and slow, sometimes "letter-by-letter". This was clearly not a successful strategy in consideration of his phonological dyslexia, yet he apparently used it as in trying to control the production of a correct long sequence. Table 3 shows the pattern of RT's errors in reading words.

RF Errors							
Type	e.g.	N=	Corr	LbyL	- +	+ -	- -
NN1	*pescespada,* ‘swordfish’	12	12				
NN2	*videogioco,* ‘wordgame’	12	12				
NC1	*pellegrino,* ‘pilgrim’	12	8	3	1		
NC2	*pavimento,* ‘floor’	12	9	2	1		
TOT		*48*	*41*	*5*	*2*		

Table 3. Pattern of RF’errors in reading words (NN1: left-headed compounds; NN2: right-headed compounds; NC1: noncompounds containing an embedded word in the initial position; NC2: noncompounds containing an embedded word in the final position).
Errors: LbyL derive from an overusing of the “letter-by-letter” strategy on the whole word; (- ..+) deletion or substitution of the first constituent/ part; deletion or substitution of the second constituent/part; (- ..-) deletion or substitution of both constituents/parts of the word.

In the case of non-words, RF read correctly 13/24 (54.16%) inverted compounds and only 5/24 (20.8%) non-words derived from NC ($\chi2(1)= 5.69$, $p=0.02$] Furthermore, whereas in reading inverted NC the patient read slowly and often in a “letter-by-letter” fashion (in 13/24 cases), this strategy appeared less frequently with inverted compounds (in 4/24 cases), [($\chi2(1)= 4.75$, $p=0.02$]. In inverted compounds RF made a peculiar kind of error in five cases out of nine errors: the hyper-lexicalization of the written stimulus, i.e. he reversed the order of constituents and uttered the correct compound (e.g., the inverted compound maglia$_2$calza$_1$ was read reversing the constituents and saying the correct compound calza$_1$maglia$_2$ ‘tights’; and banda$_2$capo$_1$ was read correctly as capo$_1$banda$_2$, ‘band leader’). In this way RF thus showed that he segmented the non-word stimulus and re-composed it as a real word.

	Experimental Stimuli	
Type	**e.g.**	**translation**
NN1	*pescespada*	‘swordfish’, lit. ‘fish sword’
NN2	*videogioco*	‘videogame’
NC1	*Pellegrino*	‘pilgrim’, (pelle, ‘skin’)
NC2	*Pavimento*	‘floor’, (mento, ‘chin’)

8.2.3 Discussion

RF read compounds better than length-matched NC; furthermore within non-words, he read inverted compounds better than inverted NC. However, the fact that he could repeat easily and without errors both compounds and NC, makes the possibility of a deficit at the phonological output buffer very unlikely. The locus of impairment could be reasonably placed at the level of the graphemic buffer, in line with the proposal that such processing component plays a role also in reading (Caramazza et al., 1996) . According to this view, when reading compounds, RF could keep the two separate abstract morphemes in the buffer, and then access the orthographic lexicon finding both the lexical units. This interpretation is analogous to that suggested for the graphemic output buffer deficit studied by Badecker et al. (1990). In contrast, matched-for-length NC, which cannot be de-composed into two real words, put an overwhelming load on RF's defective buffer: if halved, the two non-lexical strings cannot find lexical entries. Moreover, the fact that compound constituents have a higher lexical frequency than whole compounds facilitates recognition of singular components.

The hyper-lexicalizations of inverted compounds is an interesting feature of this case. These errors could be explained in two different ways:

1) To reduce the overwhelming load, RF may divide inverted compounds in two parts before entering the buffer. Inverted NC thus access the mental lexicon in a decomposed form. Subsequently, RF may access the phonological form at the word form level where the correct order of constituents is believed to be assigned (Levelt et al., 1999).

According to the dual route model the forms of both the constituents and the whole-word are available in the mental lexicon (Baayen, Dijkstra, Schreuder, 1997) Thus, recognizing the two constituents, RF can occasionally gain access to the only whole-word compound available in the lexicon, which is in the correct order.

8.3 Case 2: A patient with phonological dysgraphia

DA is an 82 years old Italian-speaking woman, right-handed, with a thirteen years of education. She suffered a cerebro-vascular accident resulting in an ischemic lesion in the left corona radiata. The aphasia examination (Italian version of the AAT, (Luzzatti, Wilmes, De Bleser, 1996, Table 3) revealed fluent spontaneous speech with many anomias, i.e. difficulty in retrieving the whole phonological form of items (mainly in the Compound naming section of the battery) and spared comprehension only for short sentences. Reading and repetition abilities were well preserved except for long phrases. Writing was impaired: a clear pattern of phonological dysgraphia, with word over non-word superiority, emerged from the assessment.

	Raw score	*PR*	T	Deficit
Token	23	58	52	Medium/slight
Repetition	146	95	67	Slight/minimal
Written language	82	92	64	Slight/minimal
Naming	96	82	59	Slight/minimal
Comprehension	105	84	60	Slight/minimal

Table 4. Achener Aphasia Test of patient DC

8.3.1 Procedure

DA was administered the whole list of compounds, NC, fillers and non-words in a writing on dictation task. The examiner read each item aloud in a neutral tone, i.e. without underlining the presence of two separate words in the case of compound nouns. Furthermore, DA was administered a list of 48 pairs of words with the conjunction e, 'and', interposed. Twenty-four pairs were made up of the two morphemes of each of the compound stimuli (e.g., for the compound capobanda, 'band leader', the corresponding pair was capo e banda, 'leader and band'), and 24 pairs of two-syllable words matched for length and frequency with the compound constituents (e.g., canto e filtro, 'song and filter').

8.3.2 Results

DA wrote flawlessly 28/58 (48.3%) words and only 11/58 (18.9%) non-words [$\chi 2(1)= 11.16$, $p<0.01$], thus confirming the classical pattern of phonological dysgraphia (two NC items were deleted from the whole list of 60 NC described in the material).

Within words, DA wrote NC (14/22, 64%) significantly better than compounds (7/24, 29.1%; [$\chi 2(1)= 5.5$, $p=0.02$] in which errors (e.g., deletions, substitutions) always fell on the second constituent regardless of headedness. Table 5 shows DA pattern of errors in writing words.

Errors							
Type	**e.g.**	N=	Corr	LbyL	- +	+ -	- -
NN1	*pescespada,* 'swordfish'	12	6	1		5	
NN2	*videogioco,* 'wordgame'	12	1	2		9	
NC1	*pellegrino,* 'pilgrim'	11	6		2	3	
NC2	*pavimento,* 'floor'	11	8	1	1		1
TOT		*46*	*21*	*4*	*3*	*17*	*1*

Table 5. Pattern of DA errors in writing words (NN1: left-headed compounds; NN2: right-headed compounds; NC1: noncompounds containing an embedded word in the initial position; NC2: noncompounds containing an embedded word in the final position).

Errors: LbyL derive from an overusing of the "letter-by-letter" strategy on the whole word; (- ..+) deletion or substitution of the first constituent/ part; deletion or substitution of the second constituent/part; (- ..-) deletion or substitution of both constituents/parts of the word.

In the second condition, when compound constituents were dictated as two separate words mixed with other pairs of words DA made no errors and wrote flawlessly all pairs of words.

8.3.3 Discussion

D.A. showed a phonological dysgraphia and, within this pattern, writing was more correct with NC than with compounds. This clear-cut pattern of compromised written processing of compounds and preserved processing of NC points to a peculiar deficit in accessing compounds in the orthographic lexicon. Because of her phonological dysgraphia D.A. must forcefully use the lexical routine. Her deficit can be explained by positing a difficulty in keeping trace of both constituents of the compound.

The case of an aphasic patient with selective impairment in the retrieval of compound nouns in a picture naming task has already been reported (Delazer & Semenza, 1998). The authors explained her behaviour within the frame of a model by Levelt et al. (1999) according to which the first step is selection of the lexical item and assignment of semantic and syntactic features; subsequently, the phonological feature is retrieved. Thus, compound nouns could be generated by combining two different lexical entries at the first stage. The authors suggested that the deficit in their patient arose when two different constituents were defined by a single lexical entry (one compound noun).

This impairment is very similar to the deficit that DA showed in writing and the same interpretation could also apply. What remains to be explained is where the composition process takes place.

DA could not proceed with composition of compounds and could only occasionally find the whole-word compound form within the lexicon. This impairment in the compositional process is confirmed by the finding that she could write both constituents of a compound when they were dictated as two separate words with an interposed conjunction. In this condition, in fact, DA did not have to do any morphological composition. The separation of the two constituents increased the total length of the stimulus (capo e banda in place of capobanda), but, on the other hand, spread the overall charge over smaller sub-units that DA could more easily hold in her buffer.

8.4 General discussion

The performance of a dyslexic patient and a dysgraphic patient in reading/writing NN compound nouns highlights the interaction between lexical and peripheral processes (buffer). RF, the patient with phonological dyslexia, failed with NC and read compound nouns significantly better. This performance may be surprising for a phonological dyslexic patient, who usually fails with morphologically more complex words (Luzzatti, Mondini, Semenza, 2001). However, R.F.'s particular problem in reading is not with composition as in most of previously described patients. His representation of compounds in the lexicon appears, despite his phonological dyslexia, to be sufficiently strong. RF's performance is, instead, reasonably explained in light of the limited capacity of his deficient graphemic buffer. Another patient with phonological dysgraphia (DH) has been reported in literature (Badecker et al., 1990), who wrote compound nouns significantly better than mono-morphemic nouns. That performance was interpreted as the result of a deficit at the graphemic output buffer: as compound nouns could be divided in morpheme-sized units, it was easier for DH to hold them in his defective buffer. RF seems to match in reading DH's performance in writing. RF's phonological dyslexia caused more difficulties in reading morphologically complex words, but his defective graphemic buffer was advantaged from the possibility of de-composing the long compound nouns into small units (i.e., compound constituents). Thus R.F.'s case provides converging evidence that morphologically complex words access the output buffer in a decomposed form.

In contrast, the defective writing of compounds with respect to noncompounds found in the phonological dysgraphic patient DA points to another locus of functional impairment. The analysis of her performance allows to hypothesize a deficit in compositional processes, in particular in keeping two different representations in one lexical entry. An analogue impairment has already been reported in patient MB, described by Delazer and Semenza (1998), in a picture naming task.

While for the dyslexic patient RF the impairment was located at the graphemic buffer, in DA there was a lexical impairment, her deficit with morphologically complex words seriously affecting compounding. In fact, once compounds were de-composed in two single words the patient wrote these items flawlessly.

The two patients reported here thus crucially differ, albeit in different tasks (RF in reading, DA in writing), in their treatment of compounds, selectively spared in one case (RF) and selectively affected in the other case (DA). The comparison (including reference to previously reported patients DH and MB) between the two cases is useful because it allows distinguishing between aspects of the compounding process that happen and interact at different levels of retrieval. The first process is driven by lexical information about the whole word. The constituents are chosen and put in the right order. However, at this stage, compound words are apparently not assembled yet, and, as RF's and DH's cases show, they are sent to the buffers still as independent units. Thus the next step, the actual assembly, happens in the buffer. If the buffer is defective, as in RF and DH, this favours the production of compounds over monomorphemic words. If, instead, it is the process of choosing the two constituents that fails, then monomorphemic words are produced better than compounds, as it happens in DA and MB.

These cases will hopefully stimulate future research and discussion in the field. The special status of compound nouns (i.e., items that are in between words and sentences) allows to proficiently investigate reading and writing by highlighting the interaction between lexical processes (i.e., operations executed within the mental lexicon) and peripheral mechanisms (e.g., operations of components outside the mental lexicon, like assembling in the buffers). These are cognitive processes external to the lexicon, but they often determine effects in lexical tasks.

9. CONCLUSIONS

In the present dissertation four experiments examining the issue of headness in Noun-Noun compounds have been provided. The possibility of both right headed and left headed compounds represent an ambiguity that the cognitive system must be able to face in everyday language use. This is particular important in Italian, in which syntactic features must be correctly specified. So, for example let's consider the compounds in 24

24.
a) capobanda 'band leader' (lit. 'leader band')
b) astronave 'spaceship' (lit. 'star ship')

Both example in 24 are Italian Noun-Noun compounds. However, the first is left headed while the second is right headed. The same compounds are represented in 25 with gender of each constituent specified

25.
a) [[capo]$_{MASC}$ # [banda]$_{FEM}$]$_{MASC}$
b) [[astro]$_{MASC}$ # [nave]$_{FEM}$]$_{FEM}$

Although the gender of the constituent is the same for both compounds the whole compounds have a different genders: in 25.a) that is left headed, masculine as the head, while in 25.b) that is right headed, feminine as the head (see par 1.3). This characteristic may be important in Italian. For example to associate an adjective (e.g *pericoloso*, 'dangerous') to a compound an appropriate inflectional suffix must be produced. Thus, in Italian it is *un capobanda pericolos**o*** 'a dangerous band leader' (with the thematic vowel *–o* of *pericoloso* indicating singular masculine] and *un'astronave pericolos**a*** 'a dangerous spaceship' (with the thematic vowel *–a* of *pericolosa* indicating singular feminine).
Thus it is fundamental to access both syntactic and semantic information on the whole compounds. In chapter 3 some questions about Italian compound processing have

been raised. Results of the experiments discussed throughout this dissertation, offer a quite clear picture and make possible to answer those questions.

Are left headed and right headed Noun-Noun compounds processed differently?

Since right headed Noun-Noun compounds represent an “irregualarity” compared to the main left-headed, it was possible to hypothesize different mechanisms for compound processing of these two categories. Specifically, as irregular forms, right headed compound could be listed as whole word within the lexicon. Results from Experiment 2 and Experiment 3 tell that this is not the case since evidence of constituency are found also for right headed compounds although with a lesser degree compared to left headed compounds. Early effect found in ERP of Experiment 2 suggest that a strongest semantic relation for constituent in left headed Noun-Noun compounds and hence a higher degree of “constituency”.

Frequency effect from Experiment 1, 2 and 3 suggest that both a whole word representation is availabla and that it influence lexical processing.

When exactly information on head occurs?

Results of Experiment 1 suggest that information on head occurs early and, via a whole word representation, through the activation of morpho-semantic features that relates the whole compounds with the constituents. This activation operate in synergy with morpho-orthographic activation given by the structure of compounds. Right headed compounds are more “frozen” structures, in which morphological structure is less evident and plays a smaller role in comparison to left headed compounds.

“How” the information about head is encoded?

There is no an explicit morphological “marker” of head position. Information on headedness presumably emerges from the activation of the morpho-semantic features of compounds. Stronger semantic connections with the head constituent tell what is the morphological head of compound.

What’s the role of syntax and morphology in compound processing?

In chapter 1 it has been suggested that Italian compounds are morphological structure strongly influenced from syntax. This analogy is confirmed by results of Experiment 1 and 2 that showed similar behaviours for these different compound categories that can be traced in their origin in the Syntactic Plane (Di Sciullo, 2006; Di Sciullo, 2009). Compounds in experiment 2 showed almost no RT costs when they were represented as two separated words. This suggests that they indeed are juxtaposition of words with no hierarchical structure.

How compound processing interacts with more peripheral cognitive process?
Results from Experiment 4 outlined how peripheral structure of the cognitive system interact with lexical processes and how specific impairente in different loci could lead do different performances. An important interaction of lexical characteristics an non-linguistic cognitive processed have been found also in Experiment 2 and 3 in which visuospatial attention was influenced by the morphological structure of words, and specifically, by headedness.

However, there are still issues that require further investigations. All experiments in this thesis used left headed and right headed *lexicalized* compounds. An even biggest challenge for the cognitive system would be dealing with novel compounds. As already pointed in par. 1.7, right headedness shows a certain degree of productivity, although the main compounding production rule for Noun-Noun compounds is left headed. With novel compounds, headedness information cannot be retrieved trough an whole word activation and other mechanisms must necessarily be involved.

APPENDIXES

APPENDIX 1

EXPERIMENT 1 – PSYCHOLINGUISTIC VARIABLES

WHOLE WORDS

Type	n	Length	Familiarity	Frequency
NN1	28	10.32(1.23)	4.87(0.90)	4.05(1.77)
NN2	28	10.86(1.80)	4.55(1.07)	4.68(1.84)
NC1	28	9.32(1.36)	5.29(0.96)	5.61(1.89)
NC2	28	9.31(0.83)	5.44(1.04)	5.60(2.36)

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	28	5.63(1.06)	5.11(1.06)	0.03(0.19)
NN2	28	5.43(1.08)	4.28(1.43)	0.07(0.26)
NC1	28	4.28(1.22)	4.83(1.23)	0.57(0.69)
NC2	28	4.44(1.59)	4.56(1.46)	0.43(0.63)

FIRST CONSTITUENT /EMBEDDED WORD

Type	n	Length	Familiarity	Frequency
NN1	28	4.82(0.55)	5.57(0.87)	8.70(2.23)
NN2	28	4.82(0.77)	5.60(0.99)	8.48(2.15)
NC1	28	4.57(0.74)	5.34(0.96)	7.67(2.14)
NC2	28			

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	28	2.92(0.89)	5.01(1.35)	5.89(3.13)
NN2	28	3.11(0.99)	5.13(1.27)	4.71(3.54)
NC1	28	3.23(1.03)	4.59(1.31)	8.39(8.39)
NC2	28			

SECOND CONSTITUENT /EMBEDDED WORD

Type	n	Length	Familiarity	Frequency
NN1	28	5.50(1.00)	5.39	8.43(2.15)
NN2	28	6.03(1.45)	5.23	8.42(2.44)
NC1	28			
NC2	28	4.64(0.68)	5.13	7.60(1.86)

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	28	3.20(0.96	4.98(1.26)	5.25(4.27)
NN2	28	4.25(1.28)	3.95(1.47	3.00(2.70)
NC1	28			
	28	3.73(1.24)	4.87(1.26)	8.48(3.95)

Mean of the psycholinguistic variables (sd) considered for the experimental items (NN1: left-handed compounds; NN2: right-handed compounds; NC1 noncompounds with a word embedded in the left part of the whole word; NC2: noncompounds with a word embedded in the right part of the whole word.

EXPERIMENT 1 - STIMULI

Experiment 1			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NN1	*Acquavite*	Brandy	*acqua* (water); *vite* (grapes)
NN1	*Arcobaleno*	Rainbow	*arco* (bow); *baleno* (lightning)
NN1	*Bancoposta*	[Lit.] Counter post (the post office counter)	*banco* (counter); *posta* (post)
NN1	*Boccaporto*	Hatchway	*bocca* (mouth); *porto* (harbor)
NN1	*Bordovasca*	[Lit.] The edge of a swimming pool	*bordo* (edge); *vasca* (basin)
NN1	*Burrocacao*	Lipsalve	*burro* (butter); *cacao* (cocoa)
NN1	*Calzamaglia*	Tights	*calza* (sock); *maglia* (knitting)
NN1	*Camposcuola*	School camp	*campo* (camp); *scuola* (school)
NN1	*Capobanda*	Band leader	*capo* (leader); *banda* (band)
NN1	*Ceralacca*	Sealing wax	*cera* (wax); *lacca* (lake)
NN1	*Finecorsa*	Terminal station	*fine* (end); *corsa* (run)
NN1	*Focamonaca*	Monk seal	*foca* (seal); *monaca* (monk)
NN1	*Fondovalle*	Valley bottom	*fondo* (bottom); *valle* (valley)
NN1	*Girocollo*	Round neck	*giro* (round); *collo* (neck)
NN1	*Gommapiuma*	Foam rubber	*gomma* (rubber); *piuma* (feather)
NN1	*Granoturco*	Maize	*grano* (grain); *turco* (Turkish)
NN1	*Grillotalpa*	Mole cricket	*grillo* (cricket); *talpa* (mole)
NN1	*Caciocavallo*	Caciocavallo (a kind of cheese)	*cacio* (cheese) *cavallo* (horse)
NN1	*Metroquadro*	Square metre	*metro* (metre); *quadro* (square)
NN1	*Padrefamiglia [*	Lit.] The head of the household	*padre* (father); *famiglia* (family)
NN1	*Parcomacchine*	[Lit.] The company fleet of cars	*parco* (park); *macchine* (cars)
NN1	*Pastafrolla*	Short pastry	*pasta* (dough); *frolla* (butter dough)
NN1	*Pescespada*	Swordfish	*pesce* (fish); *spada* (sword)
NN1	*Pianoterra*	Ground floor	*piano* (floor); *terra* (ground)
NN1	*Prezzobase*	Starting price	*prezzo* (price); *base* (base)
NN1	*Retrobottega*	Backshop	*retro* (back); *bottega* (shop)
NN1	*Roccaforte*	Fortress	*rocca* (rock); *forte* (fort)
NN1	*Toporagno*	Shrew	*topo* (mouse); *ragno* (spider)
NN2	*Aliscafo*	Hydrofoil	*ali* (wings); *scafo* (hull)
NN2	*Architrave*	Lintel	*archi* (bows); *trave* (beam)
NN2	*Astronave*	Spaceship	*astro* (star); *nave* (ship)
NN2	*Audiofrequenza*	Audio-frequency	*audio* (audio); *frequenza* (frequency)
NN2	*Barbabietola*	Beet	*barba* (beard); *bietola* (chard)
NN2	*Broncospasmo*	Bronchospasm	*bronco* (broncho); *spasmo* (spasm)
NN2	*Calciomercato*	[Lit.] Soccer market	*calcio* (soccer); *mercato* (market)
NN2	*Cartamoneta*	Paper money	*carta* (paper); *moneta* (money)
NN2	*Crocevia*	Crossroads	*croce* (cross); *via* (road)
NN2	*Docciaschiuma*	Shower gel	*doccia* (shower); *schiuma* (foam)
NN2	*Fangoterapia*	Mud therapy	*fango* (mud); *terapia* (therapy)
NN2	*Ferrolega*	Ferroalloy	*ferro* (iron); *lega* (league)
NN2	*Filobus*	Trolley bus	*filo* (yarn); *bus* (bus)
NN2	*Fluidodinamica*	Fluid dynamics	*fluido* (fluid); *dinamica* (dynamics)
NN2	*Fotoromanzo*	Picture story	*foto* (photograph); *romanzo* (romance)
NN2	*Luogotenente*	Lieutenant	*luogo* (place); *tenente* (tenant)
NN2	*Madrepatria*	Motherland	*madre* (mother); *patria* (land)
NN2	*Mondovisione*	World vision	*mondo* (world); *visione* (vision)
NN2	*Montepremio*	Jack-pot	*monte* (mountain); *premio* (prize)

Experiment 1			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NN2	*Motosega*	Chain saw	*moto* (motor); *sega* (saw)
NN2	*Nanosecondo*	Nanosecond	*nano* (nano); *secondo* (second)
NN2	*Pollicoltura*	Poultry farming	*polli* (chickens); *coltura* (farming)
NN2	*Radiocronaca*	Running commentary	*radio* (radio); *cronaca* (commentary)
NN2	*Servosterzo*	Power steering	*servo* (servant); *sterzo* (steering)
NN2	*autostrada*	Highway	*auto* (car) *strada* (road)
NN2	*Vetroresina*	Fibre-glass plastic	*vetro* (glass); *resina* (resin)
NN2	*Videogioco*	Videogame	*video* (video); *gioco* (game)
NN2	*Zootecnica*	Zoo technology	*zoo* (zoo); *tecnica* (technology)
NC1	*Barracuda*	Barracuda	*barra* (bar)
NC1	*Cavaliere*	Horse-rider	*cava* (mine)
NC1	*Clorofilla*	Chlorophyll	*cloro* (chloro)
NC1	*Coccodrillo*	Crocodile	*cocco* (coconut)
NC1	*Collaudo*	Test/inspection	*colla* (glue)
NC1	*Filastrocca*	Rigmarole	*fila* (row)
NC1	*Formalina*	Formalin	*forma* (shape)
NC1	*Funerale*	Funeral	*fune* (cable)
NC1	*Gelosia*	Jealousy	*gelo* (chill)
NC1	*Maleficio*	Spell	*male* (ill)
NC1	*Maresciallo*	Marshal	*mare* (sea)
NC1	*Melanoma*	Melanoma	*mela* (apple)
NC1	*Melodia*	Melody	*melo* (apple-tree)
NC1	*Mercenario*	Mercenary	*merce* (goods)
NC1	*Meteorite*	Meteorite	*meteo* (weather-report)
NC1	*Oratore*	Orator	*ora* (hour)
NC1	*ortogonale*	Orthogonal	*orto* (vegetable plot)
NC1	*Paladino*	Paladin	*pala* (shovel)
NC1	*Pappagorgia*	Double chin	*pappa* (baby-food)
NC1	*Pastorizia*	Sheep farming	*pasto* (meal)
NC1	*Pellegrino*	Pilgrim	*pelle (skin)*
NC1	*Peperone*	Pepper	*pepe* (pepper)
NC1	*Polpastrello*	Pulp	*polpa* (pulp)
NC1	*Pontefice*	Pontiff	*ponte* (bridge)
NC1	*Salamandra*	Salamander	*sala* (hall)
NC1	*Serratura*	Lock	*serra* (greenhouse)
NC1	*Tassonomia*	Tassonomy	*tasso* (badger)
NC1	*Temperatura*	Temperature	*tempera* (distemper)
NC2	*Accidente*	Accident	*dente* (tooth)
NC2	*Accredito*	Crediting/credit	*dito* (finger)
NC2	*Catastrofe*	Catastrophe	*strofe* (strophes)
NC2	*conguaglio*	Balance	*aglio* (garlic)
NC2	*Dirigente*	Manager/director	*gente* (people)
NC2	*Discepolo*	Disciple	*polo* (pole)
NC2	*Fazzoletto*	Handkerchief	*letto* (bed)
NC2	*Imbarazzo*	Embarrassment	*razzo* (rocket)
NC2	*Logaritmo*	Logarithm	*ritmo* (rythm)
NC2	*Mandragola*	Mandrake	*gola* (throat)
NC2	*Marzapane*	Marzipan	*pane* (bread)
NC2	*Megalite*	Megalith	*lite* (quarrel)
NC2	*Patriarca*	Patriarch	*arca* (arch)
NC2	*Pavimento*	Floor	*mento* (chin)

Experiment 1			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NC2	*Pentecoste*	Pentecost	*coste* (coasts)
NC2	*Prezzemolo*	Parsley	*molo* (pier)
NC2	*Pugilato*	Boxing	*lato* (side)
NC2	*Recidiva*	Relapse	*diva* (goddes)
NC2	*Requisito*	Requirement	*sito* (site)
NC2	*Rotocalco*	Illustrated magazine	*calco* (impression)
NC2	*Scarafaggio*	Cockroach	*faggio* (beech tree)
NC2	*Schiamazzo*	Din	*mazzo* (bunch)
NC2	*Semaforo*	Raffic	*light* foro (hole)
NC2	*Tartaruga*	Tortoise	*ruga* (wrinkle)
NC2	*Varicella*	Chickenpox	*cella* (cell)
NC2	*Vegetale*	Vegetable	*tale* (someone)
NC2	*Vettovaglia*	Provisions	*vaglia* (money order)
NC2	*Virulenza*	Virulence	*lenza* (fishing-line)

List of the experimental stimuli: 28 NN1, left headed compounds; 28 NN2, right headed compounds; 28 NC1, noncompounds with a real word embedded on the left side of the whole word; 28 NC2, noncompounds with a real word embedded on the right side of the whole word (letter strings corresponding to real Italian words are underlined and translated).

APPENDIX 2

EXPERIMENT 2 - PSYCHOLINGUISTIC VARIABLES

WHOLE WORDS

Type	n	Length	Neighborhood size	Frequency
NN1	24	10.33(1.31)	1.12(0.99)	4.62(2.03)
NN2	24	10.04(1.54)	1.50(1.25)	5.27(1.29)
VN	24	11.12(1.87)	0.71(0.81)	4.61(1.40)
NC1	24	9.62(9.62)	2.92(1.95)	5.44(2.09)
NC2	24	9.21(9.21)	2.92(2.26)	6.01(2.19)

FIRST CONSTITUENT

Type	n	Length	Neighborhood size	Frequency
NN1	24	4.96(0.62)	29.00(11.81)	9.47(1.93)
NN2	24	4.79(0.78)	30.12(16.53)	9.46(1.48)
VN	24	5.62(1.21)	23.37(17.25)	6.48(2.11)
NC1	24	4.54(0.51)	37.04(14.35)	7.50(1.69)

SECOND CONSTITUENT

Type	n	Length	Neighborhood size	Frequency
NN1	24	5.37(1.05)	25.04(15.45)	8.69(1.72)
NN2	24	5.25(1.22)	26.04(17.97)	9.08(2.27)
VN	24	5.50(1.38)	24.04(17.21)	8.44(2.01)
NC2	24	4.71(0.62)	33.54(14.07)	7.21(2.00)

Mean of the psycholinguistic variables considered for the experimental items (NN1: left-headed compounds; NN2: right-headed compounds; NC1 noncompounds with a word embedded in the left part of the whole word; NC2: noncompounds with a word embedded in the right part of the whole word. VN: Verb Noun compounds.

EXPERIMENT 2 - STIMULI

Experiment 2			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NN1	*Acquavite*	Brandy	*acqua* (water); *vite* (grapes)
NN1	*Arcobaleno*	Rainbow	*arco* (bow); *baleno* (lightning)
NN1	*Boccaporto*	Hatchway	*bocca* (mouth); *porto* (harbor)
NN1*	*bancarotta*	Bankruptcy	*banca* (banca); *rotta* (broken)
NN1	*Bordovasca*	[Lit.] The edge of a swimming pool	*bordo* (edge); *vasca* (basin)
NN1	*Burrocacao*	Lipsalve	*burro* (butter); *cacao* (cocoa)
NN1	*Caciocavallo*	(A kind of cheese)	*cacio* (cheese); *cavallo* (horse)
NN1*	*Camposanto*	Cemetery	*campo* (filed); *santo* (holy)
NN1	*Capobanda*	Band leader	*capo* (leader); *banda* (band)
NN1	*Capoclasse*	Class leader	*capo* (leader); *classe* (class)
NN1	*Carroattrezzi*	Breakdown vehicle	*carro* (cart); *attrezzi* (tools)
NN1	*Cartacarbone*	Carbon paper	*carta* (paper); *carbone* (coal)
NN1	*centrotavola*	Centre-piece	*centro* (centre); *tavola* (table)
NN1	*Fondovalle*	Valley bottom	*fondo* (bottom); *valle* (valley)
NN1	*Girovita*	Waistline	*giro* (round); *vita* (waistline)
NN1	*Gommapiuma*	Foam rubber	*gomma* (rubber); *piuma* (feather)
NN1	*guardiacaccia*	Gamekeeper	*guardia* (guard); *caccia* (hunt)
NN1	*Montepremio*	Jack-pot	*monte* (mountain); *premi* (prizes)
NN1	*Pallavolo*	Volleyball	*palla* (ball); *volo* (flight)
NN1	*Pescecane*	Shark	*pesce* (fish); *cane* (dog)
NN1	*Pescespada*	Swordfish	*pesce* (fish); *spada* (sword)
NN1	*Pianoterra*	Ground floor	*piano* (floor); *terra* (ground)
NN1	*Roccaforte*	Fortress	*rocca* (rock); *forte* (fort)
NN1*	*Terraferma*	Dryland	*terra* (land); *ferma* (still)
NN2	*Aliscafo*	Hydrofoil	*ali* (wings); *scafo* (hull)
NN2	*Astronave*	Spaceship	*astro* (star); *nave* (ship)
NN2	*Autocarro*	Lorry	*auto* (car); *carro* (cart)
NN2	*Autotreno*	Roadtrain	*auto* (car); *treno* (train)
NN2	*Banconota*	Banknote	*banco* (bank); *nota* (note)
NN2	*Barbabietola*	Beet	*barba* (beard); *bietola* (chard)
NN2	*Bassorilievo*	Bas-relief	*basso* (low); *rilievo* (relief)
NN2	*Bagnoschiuma*	Foam bath	*bagno* (bath); *schiuma* (foam)
NN2	*Calciomercato*	[Lit.] Soccer market	*calcio* (soccer); *mercato* (market)
NN2	*Ciclomotore*	Motor	*ciclo* (cycle); *motore* (motor)
NN2	*Filovia*	Trolley line	*filo* (string); *via* (way)
NN2	*Fotocopia*	Photocopy	*foto* (photo); *copia* (copy)
NN2	*Francobollo*	Stamp	*franco* (frank); *bollo* (stamp)
NN2	*Madrepatria*	Motherland	*madre* (mother); *patria* (land)
NN2	*Madrepatria*	Manuscript	*mano* (hand); *scritto* (writing)
NN2	*Maremoto*	Seaquake	*mare* (sea); *moto* (movement)
NN2	*Motosega*	Chainsaw	*moto* (motor); *sega* (saw)
NN2*	*Primadonna*	Queen bee	*prima* (first); *donna* (woman)
NN2	*Radiocronaca*	Running commentary	*radio* (radio); *cronaca* (commentary)
NN2	*Retromarcia*	Reverse gear	*retro* (behind); *marcia* (gear)
NN2	*Seggiovia*	Chairlift	*seggio* (chair); *via* (way)
NN2	*Stratosfera*	Stratosphere	*strato* (layer); *sfera* (sphere)
NN2	*Vetroresina*	Fibre-glass plastic	*vetro* (glass); *resina* (resin)

Experiment 2			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NN2	*Videogioco*	Videogame	*video* (video); *gioco* (game)
VN	*Apriscatole*	Can opener	*apri* (open); *scatole* (boxes)
VN	*Asciugamani*	Towel	*asciuga* (dry) *mani* (hands)
VN	*Aspirapolvere*	Vacuum cleaner	*aspira* (suck) *polvere* (dust)
VN	*Attaccabrighe*	Troublemaker	*attacca* (attack) *brighe* (troubles)
VN	*Battipanni*	Carpetbeater	*batti* (beat) *panni* (clothes)
VN	*Cascamorto*	Lounge lizard	*casca* (fall) *morto* (dead)
VN	*Contagiri*	Counter	*conta* (count) *giri* (cycle)
VN	*Grattacielo*	Skyscraper	*gratta* (scratch) *cielo* (sky)
VN	*Guardaroba*	Wardrobe	*guarda* (watch) *roba* (things)
VN	*Lustrascarpe*	Shoe-shine boy	*lustra* (polish) *scarpe* (shoes)
VN	*Mangianastri*	Cassette player	*mangia* (eat) *nastri* (ribbons)
VN	*Parapetto*	Parapet	*para* (shield) *petto* (chest)
VN	*Passamontagna*	Balaclava	*passa* (pass) *montagna* (mountain)
VN	*Pelapatate*	Potato peeler	*pela* (peel) *patate* (potatos)
VN	*Reggiseno*	Bra	*reggi* (hold) *seno* (breast)
VN	*Salvagente*	Life belt	*salva* (save) *gente* (people)
VN	*Scacciapensieri*	Jew's arp	*scaccia* (fight-off) *pensieri* (thoughts)
VN	*Scansafatiche*	Lazybones	*scansa* (avoid) *fatiche* (efforts)
VN	*Schiaccianoci*	Nutcracker	*schiaccia* (crush) *noci* (nuts)
VN	*Spazzaneve*	Snowplough	*spazza* (sweep) *neve* (snow)
VN	*Strizzacervelli*	Shrink (psychiatrist)	*strizza* (shrink) *cervelli* (brains)
VN	*Tirapiedi*	Hanger-on	*tira* (pull) *piedi* (feet)
VN	*Tostapane*	Toaster	*tosta* (toast) *pane* (bread)
VN	*Tritacarne*	Meat grinder	*trita* (grind) *carne* (meat)
NC1	*Barracuda*	Barracuda	*barra* (bar)
NC1	*Cavaliere*	Horse-rider	*cava* (mine)
NC1	*Circospetto*	Wary	*circo* (circus)
NC1	*Clorofilla*	Chlorophyll	*cloro* (chloro)
NC1	*Coccodrillo*	Crocodile	*cocco* (coconut)
NC1	*Contenuto*	Contents	*conte* (count)
NC1	*Filastrocca*	Rigmarole	*fila* (row)
NC1	*Formalina*	Formalin	*forma* (shape)
NC1	*Funerale*	Funeral	*fune* (cable)
NC1	*Melanoma*	Melanoma	*mela* (apple)
NC1	*Mercenario*	Mercenary	*merce* (goods)
NC1	*Moratoria*	Moratorium	*mora* (mulberry)
NC1	*Mulinello*	Whirlpool	*muli* (donkeys)
NC21	*Ortogonale*	Orthogonal	*orto* (vegetable plot)
NC1	*Paladino*	Paladin	*pala* (shovel)
NC1	*Pappagorgia*	Double chin	*pappa* (baby-food)
NC1	*Pastorizia*	Sheep farming	*pasto* (meal)
NC1	*Pellegrino*	Pilgrim	*pelle (skin)*
NC1	*Peperone*	Pepper	*pepe* (pepper)
NC1	*Poliziotto*	Policeman	*poli (poles)*
NC1	*Polpastrello*	Pulp	*polpa* (pulp)
NC1	*Pontefice*	Pontiff	*ponte* (bridge)
NC1	*Salamandra*	Salamander	*sala* (hall)
NC1	*Tassonomia*	Tassonomy	*tasso* (badger)
NC2	*Accredito*	Crediting/credit	*dito* (finger)
NC2	*Amarena*	Black cherry	*rena* (sand)

Experiment 2			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NC2	*Camerata*	Dormitory	*rata* (instalment)
NC2	*Catastrofe*	Catastrophe	*strofe* (strophes)
NC2	*Centurione*	Centurion	*rione* (district)
NC2	*conguaglio*	Balance	*aglio* (garlic)
NC2	*Dirigente*	Manager/director	*gente* (people)
NC2	*Discepolo*	Disciple	*polo* (pole)
NC2	*Elefante*	Elephant	*fante* (infantryman)
NC2	*Fazzoletto*	Handkerchief	*letto* (bed)
NC2	*Imbarazzo*	Embarrassment	*razzo* (rocket)
NC2	*Logaritmo*	Logarithm	*ritmo* (rythm)
NC2	*Patriarca*	Patriarch	*arca* (arch)
NC2	*Pavimento*	Floor	*mento* (chin)
NC2	*Pentecoste*	Pentecost	*coste* (coasts)
NC2	*Prezzemolo*	Parsley	*molo* (pier)
NC2	*Protocollo*	Protocol	collo (neck)
NC2	*Requisito*	Requirement	*sito* (site)
NC2	*Scarafaggio*	Cockroach	*faggio* (beech tree)
NC2	*Schiamazzo*	Din	*mazzo* (bunch)
NC2	*Semaforo*	Raffic	*light* foro (hole)
NC2	*Tartaruga*	Tortoise	*ruga* (wrinkle)
NC2	*Varicella*	Chickenpox	*cella* (cell)
NC2	*Virulenza*	Virulence	*lenza* (fishing-line)

List of the experimental stimuli: 24 NN1, left headed Noun-Nouncompounds; 24 NN2, right headed Noun-Noun compounds; 24 NC1, noncompounds with a real word embedded on the left side of the whole word; 24 NC2, noncompounds with a real word embedded on the right side of the whole word (letter strings corresponding to real Italian words are underlined and translated). 24 VN: verb Noun compounte
* Noun-Adjective or Adjective-Noun included in the experiemental list.

APPENDIX 3

EXPERIMENT 3 - PSYCHOLINGUISTIC VARIABLES

WHOLE WORDS

Type	n	Length	Familiarity	Frequency
NN1	28	10.21(1.10)	4.89(0.89)	3.22(2.30)
NN2	28	10.82(1.82)	4.86(1.09)	4.18(2.28)
VN	32	10.53(1.24)	5.25(1.10)	4.71(2.72)

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	28	4.60(1.05)	5.16(1.08)	0.89(1.11)
NN2	28	5.34(1.65)	4.21(1.42)	1.14(1.30)
VN	32	4.29(1.25)	5.29(1.05)	1.12(1.22)

FIRST CONSTITUENT

Type	n	Length	Familiarity	Frequency
NN1	28	4.78(0.57)	5.53(0.88)	8.71(2.21)
NN2	28	4.86(0.85)	5.59(0.97)	8.39(2.09)
VN	32	5.28(0.96)		7.48(2.40)

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	28	2.92(0.87)	5.02(1.32)	28.62(11.55)
NN2	28	3.11(0.97)	5.09(1.25)	27.18(14.94)
VN	32			34.60(19.09)

SECOND CONSTITUENT

Type	n	Length	Familiarity	Frequency
NN1	28	5.42(0.96)	5.38(0.98)	8.38(2.14)
NN2	28	5.96(1.50)	5.28(1.01)	8.36(2.38)
VN	32	5.25(1.13)		8.91(1.38)

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	28	3.21(0.95)	5.00(1.24)	22.21(12.64)
NN2	28	4.20(1.29)	4.02(1.50)	18.5(16.40)
VN	32			26.28(20.21)

Mean of the psycholinguistic variables considered for the experimental items (NN1: left-handed compounds; NN2: right-handed compounds; NC1 noncompounds with a word embedded in the left part of the whole word; NC2: noncompounds with a word embedded in the right part of the whole word.

EXPERIMENT 3 - STIMULI

Experiment 3			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NN1	*Acquavite*	Brandy	*acqua* (water); *vite* (grapes)
NN1	*Arcobaleno*	Rainbow	*arco* (bow); *baleno* (lightning)
NN1	*Bancoposta*	[Lit.] Counter post (the post office counter)	*banco* (counter); *posta* (post)
NN1	*Boccaporto*	Hatchway	*bocca* (mouth); *porto* (harbor)
NN1	*Bordovasca*	[Lit.] The edge of a swimming pool	*bordo* (edge); *vasca* (basin)
NN1	*Burrocacao*	Lipsalve	*burro* (butter); *cacao* (cocoa)
NN1	*Calzamaglia*	Tights	*calza* (sock); *maglia* (knitting)
NN1	*Camposcuola*	School camp	*campo* (camp); *scuola* (school)
NN1	*Capobanda*	Band leader	*capo* (leader); *banda* (band)
NN1	*Ceralacca*	Sealing wax	*cera* (wax); *lacca* (lake)
NN1	*Finecorsa*	Terminal station	*fine* (end); *corsa* (run)
NN1	*Focamonaca*	Monk seal	*foca* (seal); *monaca* (monk)
NN1	*Fondovalle*	Valley bottom	*fondo* (bottom); *valle* (valley)
NN1	*Girocollo*	Round neck	*giro* (round); *collo* (neck)
NN1	*Gommapiuma*	Foam rubber	*gomma* (rubber); *piuma* (feather)
NN1	*Granoturco*	Maize	*grano* (grain); *turco* (Turkish)
NN1	*Grillotalpa*	Mole cricket	*grillo* (cricket); *talpa* (mole)
NN1	*Melograno*	Pomegranate	*melo* (apple tree); *grano* (grain)
NN1	*Metroquadro*	Square metre	*metro* (metre); *quadro* (square)
NN1	*Padrefamiglia [*	Lit.] The head of the household	*padre* (father); *famiglia* (family)
NN1	*Parcomacchine*	[Lit.] The company fleet of cars	*parco* (park); *macchine* (cars)
NN1	*Pastafrolla*	Short pastry	*pasta* (dough); *frolla* (butter dough)
NN1	*Pescespada*	Swordfish	*pesce* (fish); *spada* (sword)
NN1	*Pianoterra*	Ground floor	*piano* (floor); *terra* (ground)
NN1	*Prezzobase*	Starting price	*prezzo* (price); *base* (base)
NN1	*Retrobottega*	Backshop	*retro* (back); *bottega* (shop)
NN1	*Roccaforte*	Fortress	*rocca* (rock); *forte* (fort)
NN1	*Toporagno*	Shrew	*topo* (mouse); *ragno* (spider)
NN2	*Aliscafo*	Hydrofoil	*ali* (wings); *scafo* (hull)
NN2	*Architrave*	Lintel	*archi* (bows); *trave* (beam)
NN2	*Astronave*	Spaceship	*astro* (star); *nave* (ship)
NN2	*Audiofrequenza*	Audio-frequency	*audio* (audio); *frequenza* (frequency)
NN2	*Barbabietola*	Beet	*barba* (beard); *bietola* (chard)
NN2	*Broncospasmo*	Bronchospasm	*bronco* (broncho); *spasmo* (spasm)
NN2	*Calciomercato*	[Lit.] Soccer market	*calcio* (soccer); *mercato* (market)
NN2	*Cartamoneta*	Paper money	*carta* (paper); *moneta* (money)
NN2	*Crocevia*	Crossroads	*croce* (cross); *via* (road)
NN2	*Docciaschiuma*	Shower gel	*doccia* (shower); *schiuma* (foam)
NN2	*Fangoterapia*	Mud therapy	*fango* (mud); *terapia* (therapy)
NN2	*Ferrolega*	Ferroalloy	*ferro* (iron); *lega* (league)
NN2	*Filobus*	Trolley bus	*filo* (yarn); *bus* (bus)
NN2	*Fluidodinamica*	Fluid dynamics	*fluido* (fluid); *dinamica* (dynamics)
NN2	*Fotoromanzo*	Picture story	*foto* (photograph); *romanzo* (romance)
NN2	*Luogotenente*	Lieutenant	*luogo* (place); *tenente* (tenant)
NN2	*Madrepatria*	Motherland	*madre* (mother); *patria* (land)

Experiment 3			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NN2	*Mondovisione*	World vision	*mondo* (world); *visione* (vision)
NN2	*Montepremio*	Jack-pot	*monte* (mountain); *premio* (prize)
NN2	*Motosega*	Chain saw	*moto* (motor); *sega* (saw)
NN2	*Nanosecondo*	Nanosecond	*nano* (nano); *secondo* (second)
NN2	*Pollicoltura*	Poultry farming	*polli* (chickens); *coltura* (farming)
NN2	*Radiocronaca*	Running commentary	*radio* (radio); *cronaca* (commentary)
NN2	*Servosterzo*	Power steering	*servo* (servant); *sterzo* (steering)
NN2	*Terremoto*	Earthquake	*terre* (lands); *moto* (motion)
NN2	*Vetroresina*	Fibre-glass plastic	*vetro* (glass); *resina* (resin)
NN2	*Videogioco*	Videogame	*video* (video); *gioco* (game)
NN2	*Zootecnica*	Zoo technology	*zoo* (zoo); *tecnica* (technology)
VN	*Accendigas*	Gas lighter	*accendi* (light) *gas* (gas)
VN	*Affittacamere*	Landlord	*affitta* (rent) *camere* (lord)
VN	*Alzabandiera*	Flag-raising	*alza* (riase) *bandiera* (flag)
VN	*Appendiabiti*	Hallstand	*appendi* (hang) *abiti* (clothes)
VN	*Asciugamani*	Towel	*asciuga* (dry) *mani* (hands)
VN	*Buttafuori*	Bouncer	*butta* (trow) *fuori* (outside)
VN	*Calzascarpe*	Shoehorn	*calza* (fit) *scarpe* (shoes)
VN	*Cantastorie*	Story teller	*canta* (sing) *storie* (stories)
VN	*Cavatappi*	Corkscrew	*cava* (take out) *tappi* (corks)
VN	*Contagiri*	Counter	*conta* (count) *giri* (cycle)
VN	*Contagocce*	Dropper	*conta* (count) *gocce* (drops)
VN	*Coprifuoco*	Curfew	*copri* (cover) *fuoco* (fire)
VN	*Fendinebbia*	Fog lamp	*fendi* (split) *nebbia* (fog)
VN	*Fermacapelli*	Hair slide	*ferma* (stop) *capelli* (hair)
VN	*Giradischi*	Record-player	*gira* (spin) *dischi* (records)
VN	*Girasole*	Sunflower	*gira* (spin) *sole* (sun)
VN	*Grattacielo*	Skyscraper	*gratta* (scratch) *cielo* (sky)
VN	*Guardaroba*	Wardrobe	*guarda* (watch) *roba* (things)
VN	*Guastafeste*	Spoilsport	*guasta* (waste) *feste* (parties)
VN	*Lanciafiamme*	Flamethrower	*lancia* (throw) *fiamme* (flames)
VN	*Macinapepe*	Pepper mill	*macina* (mill) *pepe* (pepper)
VN	*Marciapiede*	Sidewalk	*marcia* (march) *piede* (foot)
VN	*Montacarichi*	Hoist	*monta* (mount) *carichi* (loads)
VN	*Paraocchi*	Blinkers	*para* (shield) *occhi* (eyes)
VN	*Parasole*	Parasol	*para* (shield) *sole* (sun)
VN	*Portabagagli*	Car trunk	*porta* (carry) *bagagli* (luggage)
VN	*Posacenere*	Ash trasher	*posa* (put) *cenere* (ash)
VN	*Reggiseno*	Bra	*reggi* (hold) *seno* (breast)
VN	*Spazzaneve*	Snowplough	*spazza* (sweep) *neve* (snow)
VN	*Tagliaunghie*	Nail clippers	*taglia* (cut) *unghie* (nails)
VN	*Tappabuchi*	Stopgap	*tappa* (plut) *buchi* (holes)
VN	*Voltafaccia*	Volte-face	*volta* (turn) *faccia* (face)

List of the experimental stimuli: 28 NN1, left headed compounds; 28 NN2, right headed compounds; 32 VN Verb Noun compounds

APPENDIX 4

EXPERIMENT 4 - PSYCHOLINGUISTIC VARIABLES

WHOLE WORDS

Type	n	Length	Familiarity	Frequency
NN1	12	9.92(0.79)	5.11(0.77)	3.92(1.58)
NN2	12	10.75 (1.71)	5.01(0.90)	4.98(1.32)
NC1	12	9.75(1.66)	5.67(1.22)	5.82(1.85)
NC2	12	9.25(0.75)	5.84(0.99)	5.64(2.35)

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	12	4.63(0.97)	5.15(1.00)	1.00(0.70)
NN2	12	5.09(0.98)	5.06(1.00)	0.85(0.69)
NC1	12	3.74(0.69)	5.20(0.88)	2.62(1.99)
NC2	12	3.91(0.66)	5.09(0.94)	2.62(2.44)

FIRST CONSTITUENT /EMBEDDED WORD

Type	n	Length	Familiarity	Frequency
NN1	12	4.67(0.49)	5.64(0.53)	9.52(1.96)
NN2	12	4.91(0.54)	5.96(0.92)	9.31(2.23)
NC1	12	4.17(0.83)	5.34(1.00)	7.54(2.17)

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	12	3.05(0.64)	4.55(1.47)	6.5(3.00)
NN2	12	2.86(0.75)	5.55(0.76)	4.54(3.61)
NC1	12	3.32(1.00)	4.70(1.32)	8.73(5.57)

SECOND CONSTITUENT /EMBEDDED WORD

Type	n	Length	Familiarity	Frequency
NN1	12	5.25(0.45)	5.72(0.80)	8.62 (1.77)
NN2	12	5.64(1.43)	5.68(0.87)	9.35(2.26)
NC1	12	4.68(1.00)	5.6(0.94)	9.20(2.24)

Type	n	Age of Acquisition	Imageability	Neighborhood size
NN1	12	2.87(0.64)	5.34(0.78)	6.33(4.10)
NN2	12	3.61(1.30)	4.47(1.39)	3.45(3.59)
NC1	12	3.29(1.28)	4.88(1.51)	8.91(4.08)

Mean of the psycholinguistic variables considered for the experimental items (NN1: left-handed compounds; NN2: right-handed compounds; NC1 noncompounds with a word embedded in the left part of the whole word; NC2: noncompounds with a word embedded in the right part of the whole word.

EXPERIMENT 4 - STIMULI

EXPERIMENT 4			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NN1	*Arcobaleno*	Rain-bow	*arco* (bow); *baleno* (lightning)
NN1	*Bancoposta*	[lit.] Counter post (the post office counter)	*banco* (counter); *posta* (post)
NN1	*Calzamaglia*	Tights	*calza* (sock); *maglia* (knitting)
NN1	*Camposcuola*	School camp	*campo* (camp); *scuola* (school)
NN1	*Capobanda*	Band leader	*capo* (leader); *banda* (band)
NN1	*Ceralacca*	Sealing wax	*cera* (wax); *lacca* (lake)
NN1	*Finecorsa*	Terminal station	*fine* (end); *corsa* (run)
NN1	*Fondovalle*	Valley bottom	*fond* (bottom); *valle* (valley)
NN1	*Girocollo*	Round neck	*giro* (round); *collo* (neck)
NN1	*Gommapiuma*	Foam rubber	*gomma* (rubber); *piuma* (feather)
NN1	*Metroquadro*	Square metre	*metro* (metre); *quadro* (square)
NN1	*Pescespada*	Swordfish	*pesce* (fish); *spada* (sword)
NN2	*Astronave*	Spaceship	*astro* (star); *nave* (ship)
NN2	*Calciomercato*	[lit.] Soccer market	*calcio* (soccer); *mercato* (market)
NN2	*Cartamoneta*	Paper money	*carta* (paper); *moneta* (money)
NN2	*Crocevia*	Crossroads	*croce* (cross); *via* (road)
NN2	*Docciaschiuma*	Shower gel	*doccia* (shower); *schiuma* (foam)
NN2	*Fangoterapia*	Mud therapy	*fango* (mud); *terapia* (therapy)
NN2	*Fotoromanzo*	Picture story	*foto* (photograph); *romanzo* (romance)
NN2	*Madrepatria*	Motherland	*madre* (mother); *patria* (land)
NN2	*Mondovisione*	World vision	*mondo* (world); *visione* (vision)
NN2	*Montepremio*	Jack-pot	*monte* (mountain); *premio* (prize)
NN2	*Motosega*	Chain saw	*moto* (motor); *sega* (saw)
NN2	*Radiocronaca*	Running commentary	*radio* (radio); *cronaca* (commentary)
NN2	*Vetroresina*	Fibre-glass plastic	*vetro* (glass); *resina* (resin)
NN2	*Videogioco*	Videogame	*video* (video); *gioco* (game)
NC1	*Cavaliere*	Horse-rider	*cava* (mine)
NC1	*Clorofilla*	Chlorophyll	*cloro* (chloro)
NC1	*Coccodrillo*	Crocodile	*cocco* (coconut)
NC1*	*Cremagliera*	Rack	*crema* (cream)
NC1	*Filastrocca*	Rigmarole	*fila* (row)
NC1	*Gelosia*	Jealousy	*gelo* (chill)
NC1	*Maresciallo*	Marshal	*mare* (sea)
NC1	*Melodia*	Melody	*melo* (apple-tree)
NC1	*Pellegrino*	Pilgrim	*pelle* (skin)
NC1	*Peperone*	Pepper	*pepe* (pepper)
NC1	*Polpastrello*	Pulp	*polpa* (pulp)
NC1	*Salamandra*	Salamander	*sala* (hall)
NC1	*Serratura*	Lock	*serra* (greenhouse)
NC1	*Temperatura*	Temperature	*tempera* (distemper)
NC2	*Accidente*	Accident	*dente* (tooth)
NC2*	*Catafalco*	Catafalque	*falco* (hawk)
NC2	*Dirigente*	Manager / director	*gente* (people)
NC2	*Fazzoletto*	Handkerchief	*letto* (bed)
NC2	*Logaritmo*	Logarithm	*ritmo* (rythm)
NC2	*Marzapane*	Marzipan	*pane* (bread)
NC2	*Pavimento*	Floor	*mento* (chin)

EXPERIMENT 4			
Type	**Stimulus**	**Translation**	**Translation of stimulus constituents**
NC2	*Scarafaggio*	Cock-roach	*faggio* (beech tree)
NC2	*Schiamazzo*	Din	*mazzo* (bunch)
NC2	*Tartaruga*	Tortoise	*ruga* (wrinkle)
NC2	*Varicella*	Chicken pox	*cella* (cell)
NC2	*Vegetale*	Vegetable	*tale* (someone)

*Items that did not used in the writing on dictation task administered to DA

List of the experimental stimuli: 12 NN1, left headed compounds; 12 NN2, right headed compounds; 12 NC1, noncompounds with a real word embedded on the left side of the whole word; 12 NC2, noncompounds with a real word embedded on the right side of the whole word (letter strings corresponding to real Italian words are underlined and translated).

References

- Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compounds: There is a mouse in mousetrap. *European Journal of Cognitive Psychology,* 16, 285 -311.
- Annoni, J. M., Lemmay, M. A., de Mattos Pimenta, M. A. & Lecours A. R. (1998) The contribution of attentiona mechanisms to an irregularity effect at the graphemic buffer level, *Brain and Language,* 63, 64-78.
- Arduino, L. S., Burani, C., & Vallar, G. (2002). Lexical effects in left neglect dyslexia: A study in Italian patients. *Cognitive Neuropsychology. 19*, 421–444.
- Baayen, R. H., (2004) Statistics in Psycholinguistics: A critique of some current gold standards. *Mental Lexicon Working Papers I* , 1-45
- 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., Dijkstra, T., and Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, 36:94, 117.
- Baayen, R. H., Schreuder, R., De Jong, N. H., & Krott, A. (2002). Dutch inflection: the rules that prove the exception. In S. Nooteboom, F. Weerman, & F. Wijnen (Eds.), *Storage and Computation in the Language Faculty* (pp. 61–92). Kluwer Academic Publishers, Dordrecht.
- Baayen, R. H., Tweedie, F. J., Schreuder, R., (2002) The Subjects as a Simple Random Effect Fallacy: Subject Variability and Morphological Family Effects in the Mental Lexicon. *Brain and Language*, 81, 55-65.
- Baayen, R.H. (2007). Storage and computation in the mental lexicon. In G. Jarema and G. Libben (Eds.) *The Mental Lexicon: Core Perspectives*, Elsevier, 81-104. .
- Baayen, R.H., (2007) *Analyzing Linguistic Data*. A practical introduction to statistics. Cambridge: Cambridge University Press.
- Badecker, W. (2001): "Lexical composition and the production of compounds: Evidence from errors in naming". *Language and Cognitive Processes* 16, 337–366.
- Badecker, W., Hillis, W., Caramazza W., (1990) Lexical morphology and its role

in the writing process: Evidence from a case of acquired dysgraphia, *Cognition*, 35, 205-243.

- Behrmann, M., Moscovitch, M., Black, S. E., & Mozer, M. (1990). Perceptual and conceptual factors in neglect dyslexia: Two contrasting case studies. *Brain, 113*, 1163–1183.
- 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.
- Bien, H., Levelt, W., and Baayen, R. H. (2005). Frequency effects in compound production. *PNAS* 102, 17876-17881.
- Bisetto, A., Scalise, S. (2005) Classification of Compounds. *Lingue e Linguaggio* 4(2), 319-332
- Bisetto, A., Scalise, S., (1999) 'Compounding: Morphology and/or syntax?' in Mereu, L., (Ed.) *Boundaries of Morphology and Syntax.* Amsterdams: John Benjamins, 31-48.
- Bisiach, E., Meregalli, S., & Berti, A. (1990). Mechanisms of production control and belief fixation in human visuospatial processing: Clinical evidence from unilateral neglect and misrepresentation. In M.L. Commons, R.J. Herrnstein, S.M. Kosslyn, & D.B. Mumford (Eds.), *Models of behaviour: Computational and clinical approaches to pattern recognition and concept formation.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Bisiach, E., Valla, G., Perani, D., Papagno, C., Berti, A. (1986) Unawareness of disease following lesions of the right hemisphere: anosognosia for hemiplegia and anosognosia for hemianopia. *Neuropsychologia, 24*, 471-82.
- Blanken, G. (2000): "The production of nominal compounds in aphasia". *Brain and Language*. 74, 84–102.
- Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production*: Vol.2 London: Academic Press.
- Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425-455.
- Caramazza, A., Miceli, G., Villa, G., Romani, C., (1987) The role of the graphemic buffer in spelling: evidence from a case of acquired dysgraphia, *Cognition*, 26, 59-85.
- Caramazza, A., Capasso, R., Miceli, G., (1996) The role of the graphemic buffer

in reading *Cognitive Neuropsychology,* 13 , 673-698.

- Caramazza, A., Hillis, A.E., (1990) Levels of representation, co-ordinate frames, and unilateral neglect. *Cognitive neuropsychology,* 7 (1990), 391-445.
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28, 297–332.
- Chiarelli, V., Menichelli, A., Semenza, C., Naming compounds in Alzheimer disease, *The Mental Lexicon,* (2007), 261-172.
- Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. Journal of Verbal Learning and Verbal Behavior, 12, 335–359.
- Cohen, J., (1983) The Cost of Dichotomization. *Applied Psychological Measurement,* 7, 249-254
- Cubelli, R., Beschin, N. (2005). The processing of the right-sided accent mark in left neglect dyslexia, *Brain and Language, 95,* 319–326.
- Davidson, D. J., (2009) Functional Mixed-effect Models for electrophysiological responses. *Neurophysiology,* 41(1), 71-79.
- De Jong, N. H., Feldman, L. B., Schreuder, R., Pastizzo, M., & Baayen, R. H. (2002). The processing and representation of Dutch and English compounds: Peripheral morphological, and central orthographic effects. *Brain and Language*, 81, 555–567.
- Delazer, M., & Semenza, C., (1998) The processing of compound words. A study in Aphasia. *Brain and Language*, 61, 54-62.
- Dell'Acqua, R., Pesciarelli, F., Jolicœur, P., Eimer, M., Peressotti, F., (2007) The interdependence of spatial attention and lexical access as revealed by early asymmetries. *Psychophysiology* (44), 1-8
- in occipito-parietal ERP activity
- Deutsch, J.A., Deutsch, D. (1963). Some theoretical considerations. *Psychol Review, 70*, 80-90.
- Deutsch, J.A., Deutsch, D. (1963). Some theoretical considerations. *Psychological Review, 70*, 80-90.
- Di Sciullo (2005) *Asymmetry in Morphology*. Cambridge, Mass.: The MIT Press
- Di Sciullo (2006) *Decomposing compounds*. Unpublished Manuscript

- Diependaele, K., Sandra, D., Grainger, J., (2009) Semantic priming and masked morphological priming: The case of prefixed words. *Memory & Cognition* 37(6), 895-908.
- Dipendaele, K., Sandra, D., Grainger, J., Lexical and sublexical morphological influences in visual word recognition. *Work presented at the Sixth International conference on the Mental Lexicon.* Banff, October 7-10, Alberta, Canada.
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*, 357–374.
- Dressler, W. U., (2006). Compound Types. In Libben, G., Jarema, G., (Eds.) (2006) *The Representation and Processing of Compound Words*. Oxford, UK: Oxford University Press.
- Duñabeita, J. A., Laka, I., Perea, M., Carreiras, M., (2009) Is *Milkman* a superhero like *Batman*? Constituent morphological priming in compound words. *European Jouranl of Cognitive Psychology*. 21(4) 615-640
- Duñabeitia, J. A., Perea, M., Carreiras, M., (2007) The role of the frequency of constituents in compound words: Evidence from Basque and Spanish. *Psychonomic Bulletin & Review*. 14 (6), 1171-1176.
- Fiorentino, R., Fund-Reznicek, E. (2008) Dissociating morphology, form and meaning in compound processing. *Work presented at the Sixth International conference on the Mental Lexicon.* Banff, October 7-10, Alberta, Canada.
- Forster, K. I. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research*, 27, 203-233.
- Friederici, A., (2004) Event-related Brain Potential Studies in Language. *Current Neurology and Neuroscience Reports*, 4, 466–470
- G. Vallar, C. Guariglia, D. Nico and P. Tabossi, (1996) Left neglect dyslexia and the processing of neglected information. *Journal of Clinical and Experimental Neuropsychology,* 18 1–14.
- Gagné, C. L., & Spalding, T. L., (2006) Conceptual combination: implication for the mental lexicon. In In Libben, G., Jarema, G., (Eds.) *The Representation and Processing of Compound Words*. Oxford, UK: Oxford University Press.
- Geisser, S., & Greenhouse, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–112.
- Giraudo, H., & Grainger, J. (2000) Priming complex words: Evidence for

supralexical representation of morphology. *Psychonomic Bullettin & Review* 8(1), 127-131

- Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, 15, 421-444.
- Gold, B. T., Rastle, K., (2007) Neural Correlates of Morphological Decomposition during Visual Word Recognition.
- Grossmann, M., Rainer, F., (Eds) (2004) *La Formazione delle Parole in Italiano*. Tübingen: Niemeyer.
- Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 194–205.
- Hamilton, A. C., Coslett, H. B., (2007) Impairment in writing, but not reading, morphologically complex words. 45(7), 1586-1590
- Hanley, J. R., Key, J. Note: Does the graphemic buffer play a role in reading? *Cognitive neuropsychology,* 15, 313-318.
- Harley (2008) *Psychology of Language: from Data to Theory, Third Edition.* Psychology Press: London.
- Hittmair-Delazer, M., Andree, B., Semenza, C., De Bleser, R., & Benke, T. (1994): "Naming by German compounds". *Journal of Neurolinguistics* 8, 27–41.
- Holcomb, P. J., (1993) Semantic priming and stimulus degradation: Implication for the role of N400 in language Processing. *Psychophysiology* 30, 47-61
- Holcomb, P. J., Reder, L., Misra, M., Grainger, J., (2005) The effects of prime visibility on ERP measures of masked priming. *Cognitive Brain Research* 24, 155–172
- Holcomb, P.J.,Grainger, J.,O'Rourke,T., 2002.Anelectrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*. 14, 938–950.
- Inhoff, A. W., Briihl, D., & Schwartz, J. (1996). Compound word effects differ in reading, on-line naming, and delayed naming tasks. *Memory and Cognition*, 24, 466-476.
- 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. (2007): "The processing of compounds in bilingual aphasia". *Brain and Language* 103, 22–23.
- 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
- Kinoshita, S., Lupker, J. S., (2003) *Masked Priming. The State of Art.* New York: Pyshology Press.
- Koester, D., Gunter, T. C., & Wagner, S. (2007). The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPs. *Brain and Language*, 102, 64–79.
- Koester, D., Gunter, T. C., Wagner, S., & Friederici,A. D. (2004). Morphosyntax, prosody, and linking elements: The auditory processing of Germannominal compounds. *Journal of CognitiveNeuroscience*, 16, 1647–1668.
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38,* 557–577.
- Kutas, M., Federmeir, K. D., (2005) Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Sciences, 4(12), 463-470.
- Làdavas, E. (1998). Interaction of spatial attention and reading processing in neglect dyslexia. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 723–743). Cambridge, MA: MIT Press.
- Libben, G., (1998) Semantic Transparency and the Processing of Compounds: Consequences for Representation, Processing and Impairment. *Brain and Language* 61, 30-44.
- 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*. Oxford, UK: Oxford University Press.
- Libben, G., Jarema, G., (Eds.) (2006). *The Representation and Processing of Compound Words*. Oxford, UK: Oxford University Press.
- Longtin, C. M., Segui, J., Hallé, P. A. (2003) Morphological priming without morphological relationship, *Language and Cognitive Processes*. 18, 313-334

- Longtin, C., Meunier, F. (2005). Morphological decomposition in early visual word processing. *Journal of Memory and Language, 53*, 26–41.
- Luck, S., J., (2005) *An Introduction to the Event-Related Potential Technique*. Cambridge, Mass.: The MIT Press
- Luzzatti, C., Mondini, S., & Semenza, C., (2001) Lexical representation of morphologically complex words: Evidence from an Italian agrammatic patient, *Brain and Language*, 79, 345-359.
- Luzzatti, C., Wilmes, W., De Bleser, R., (1996) Aachener Aphasia Test (AAT): Versione Italiana 2nd ed. Firenze: Organizzazioni speciali
- M. Kiefer, The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400. priming effects, *Cognitive Brain Research* 13 (1) (2002), 27– 39.
- M.L. Commons, R.J. Herrnstein, S.M. Kosslyn, & D.B. Mumford (Eds.) (1990), *Models of behaviour: Computational and clinical approaches to pattern recognition and concept formation.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc
- Marelli, M., Aggujaro, S., Molteni, F., Luzzatti, C. (2009b) Morphological structure and neglect dyslexia: A study on compound-word reading. *MOPROC .6th Morphological Processing Conferente.* June 14-17, Turku, Finland.
- 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(3), 430-454.
- McKinnon, R., Allen, M., & Osterhout, L. (2003). Morphological decomposition involving non productive morphemes: ERP evidence. *NeuroReport*, 14, 883–886.
- Mereu, L., (Ed.) *Boundaries of Morphology and Syntax.* Amsterdams: John Benjamins, 31-48.
- Mondini, S., Jarema, G., Luzzatti, C., Burani, C., & Semenza, C. (2002): "Why is 'Red Cross' different from 'Yellow Cross'? A neuropsychological study on noun-adjective agreement within Italian compounds". *Brain and language* 81, 621–634.

- Mondini, S., Kehayia, E., Gillon, B., Arcara, G., Jarema, G., (2009) Lexical access of mass and count nounds. How word recognition times corretale with lexical and morpho-syntactic processing. *The mental Lexicon*. 4(3), 354-379
- Mondini, S., Luzzatti, C., Saletta, C., Allamano, N., Semenza, C., (2005) Mental representations of prepositional compounds: Evidence of Italian agrammatic patients. *Brain and Language.* 94, 178-187.
- Mondini, S., Luzzatti, C., Zonca, G., Pistarini, C. & Semenza, C. (2004): "The mental representation of Verb-Noun Compounds in Italian: Evidence from a multiple single-case study in aphasia". *Brain and Language* 90, 470–477.
- Nittono, H., Suehiro, M., Hori T. (2002) Word imageability and N400 in an incidental memory paradigm. *International Journal of Psychophisiology* 44(3) 219-229.
- Nobre, A. C., McCharty, G. (1994) Language-Related ERPs: Scalp Distributions and Modulation by Word Type and Semantic Priming. *Journal of Cognitive Neuroscience*, 63, 233-255
- Pinker, S., (1999) Word and Rules. *Lingua* 106, 219-242
- Pollatsek, A, & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*, 20, 261 -290
- Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences*, 7, 187 -189.
- R Development Core Team (2007). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.
- Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. Journal of Memory and Language, 41, 416–426.
- Ralli, A., (1992) Greek Compounding, in Scalise, S., (Ed.) *Morphology of compounding*. Monography on *Rivista di Linguistica*
- Rastle, K., Davis, M. H., New, B. (2004) The broth in my brother's brothel: morpho-orthographic segmentation in visual word recognition, *Psychonomic Bulletin & Review* 11, 1090-1098

- Riddoch, J., Humphreys, G., Cleton, P., Fery, P. (1990). Interaction of attentional and lexical processes in neglect dyslexia. *Cognitive Neuropsychology, 7*, 479 – 517.
- Scalise, S., (1992a) Compounding in Italian. *Rivista di Linguistica*. 4, 175-199
- Scalise, S., (1992b) *Morphology of compounding*. Monography on *Rivista di Linguistica*
- Scalise, S., (1994) *Morfologia*. Bologna: Il Mulino.
- Scalise, S., Guevara, G., (2005) The Lexicalist Approach of Word Formation and the Notion of the Lexicon *in* Štekauer, P., Lieber, R. (Eds.) *Handbook of Word-Formation.* Dordrecht: The Netherlands. Springer
- Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Semenza, C. & Mondini, S., (in press). Compound Words in Neuropsychology.
- Semenza, C., & Mondini, S. (2006). Neuropsychology of compound words. In G. Libben & G. Jarema (Eds.), *The Representation and Processing of Compound Words*. Oxford, UK: Oxford University Press.
- Semenza, C., Butterworth, B. , Panzeri, M. & Hittmair-Delazer, M. (1992): Derivational rules in aphasia. *Berkeley Linguistic Society* 18 , Berkeley, CA.
- Semenza, C., Luzzatti, C., & Carabelli, S. (1997): “Morphological representation of a nouns: A study on Italian aphasic patients”. *Journal of neurolinguistics* 10, 33–43.
- Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts applied to the P300 ERP. *Psychophysiology*, 23, 695–703.
- Shoolman, N., Andrews, S., (2003) Racehorses Reindeer and Sparrows. Using Masked Priming to Investigate Morphological Influences on Compound words Identification *in* Kinoshita, S., Lupker, J. S., *Masked Priming. The State of Art.* New York: Pyshology Press.
- Siéreff, E., Pollatsek, A., Posner, M. (1988). Recognition of visual letter strings following injury to the posterior visual spatial attention system, *Cognitive Neuropsychology, 5,* 427-449.
- Spencer, A., (2005) Word formation and syntax, in Štekauer, P., Lieber, R. (Eds.) *Handbook of Word-Formation.* Dordrecht: The Netherlands. Springer.

- Steinhauer, K., Connolly, J. F., (2008) Event-related potential in the study of language in Stemmer, B., Whitaker, H. A. (Eds.) *The Handbook of Neuroscience of Language.* Elsevier: Amsterdam.
- Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 57, 745–765.
- 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.
- Umiltà, C. (2001). Mechanisms of attention. In B.Rapp (Ed.), *The Handbook of Cognitive Neuropsychology* (pp. 135-158). New York: Psychology Press.
- Ward, J., (2003) Understainding Oral spelling: a review and synthesis, *Neurocase* 9, 1.14
- Williams, E., (1981) Language, Structure and Morphology. *Linguistic Review*, 1, 81-114.