Published in final edited form as:

J Commun Disord. 2019 ; 79: 58–75. doi:10.1016/j.jcomdis.2019.03.001.

Verb and sentence processing patterns in healthy Italian participants: Insight from the Northwestern Assessment of Verbs and Sentences (NAVS)

Elena Barbieri^{a,*}, Irene Brambilla^b, Cynthia K. Thompson^{a,c,d}, Claudio Luzzatti^{b,e}

^aCenter for the Neurobiology of Language Recovery, Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL, 60208, United States

^bDepartment of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo, 1, 20126, Milano, Italy

^cMesulam Cognitive neurology and Alzheimer's Disease Center, Northwestern University, Chicago, IL, United States

^dDepartment of Neurology, Northwestern University, Chicago, IL, United States

^eMilan Center for Neuroscience, Milano, Italy

Abstract

We developed an Italian version of the Northwestern Assessment of Verbs and Sentences (NAVS, Thompson, 2011), a test assessing verb and sentence deficits typically found in aphasia, by focusing on verb-argument structure and syntactic complexity effects, rarely captured by standard language tests.

Twenty-one young healthy individuals underwent a computerized experimental version of the NAVS, including three subtests assessing production/comprehension of verbs with different number (one, two, three) and type (obligatory or optional) of arguments, and two investigating production/comprehension of sentences with canonical/non-canonical word order.

The number of verb arguments affected participants' reaction times (RTs) in verb naming and comprehension. Furthermore, verbs with optional arguments were processed faster than verbs with only obligatory arguments. Comprehension accuracy was lower for object-cleft vs. subject-cleft sentences. Object clefts and object relatives also elicited longer RTs than subject clefts and subject relatives, respectively.

*Corresponding author.: elena.barbieri@northwestern.edu (E. Barbieri).

Authorship roles

Elena Barbieri: Conceptualization and design of the study; analysis, interpretation; drafting of the article for intellectual content; agreement to be accountable for all aspects of the work; approval of the final version.

Irene Brambilla: Data acquisition, analysis and interpretation; revising the article for intellectual content; agreement to be accountable for all aspects of the work; approval of the final version.

Cynthia K. Thompson: Conceptualization and design of the study; interpretation; revising the article for intellectual content; agreement to be accountable for all aspects of the work; approval of the final version.

Claudio Luzzatti: Conceptualization and design of the study; interpretation; revising the article for intellectual content; agreement to be accountable for all aspects of the work; approval of the final version.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.jcomdis.2019.03.001>.

The study shows that the NAVS is sensitive to linguistic aspects of verb/sentence processing in Italian as in the English language. The study also highlights some differences between languages in the verb/sentence processing patterns of healthy individuals. Finally, the study contributes to the understanding of how information about verb-argument structure is represented and processed in healthy individuals, with reference to current models of verb processing.

Keywords

Verb; Sentence; Production; Comprehension; Aphasia

1. Introduction

Despite the frequent occurrence of deficits in verb and sentence processing in aphasia, standard aphasia assessment batteries both in English (e.g. Western Aphasia Battery, WAB, Kertesz, 1982; Boston Diagnostic Aphasia Examination, BDAE, Goodglass & Kaplan, 1983; Comprehensive Aphasia Test, CAT, Porter & Howard, 2004) and in Italian (Aachen Aphasia Test – Italian Version, AAT, Luzzatti, Willmes, & De Bleser, 1996; Esame Neuropsicologico per l’Afasia, ENPA, Capasso & Miceli, 2001; Esame del Linguaggio II, Ciurli, Marangolo, & Basso, 1996) lack an extensive assessment of verb and sentence production/comprehension.

A few available tests in English address either verb (Object and Action Naming Battery, Druks & Masterson, 2000) or sentence processing (Philadelphia Comprehension Battery for Aphasia, Saffran, Schwartz, Linebarger, Martin, & Bochetto, 1988; Subject-relative, Object-relative, Active, Passive battery, Love & Oster, 2002), and only two (i.e., Verb and Sentence Test, VAST, Bastiaanse, Edwards, Maas, & Rispens, 2003; Northwestern Assessment of Verbs and Sentences, NAVS, Thompson, 2011) assess both, although different sentence structures are examined for comprehension and production in the VAST. Turning to Italian, Luzzatti and coworkers (Crepaldi et al., 2006; Luzzatti et al., 2002) developed two tests assessing picture naming of nouns and verbs, where the latter were classified based on their argument properties (although no differentiation was made between optional and obligatory arguments, or between two-argument and three-argument transitive verbs). A more recent battery (Cecchetto, Di Domenico, Garraffa, & Papagno, 2012) focuses on the diagnosis of sentence comprehension deficits by investigating the influence of syntactic complexity and working memory on sentence processing.

The NAVS (Thompson, 2011) was developed to investigate difficulties in comprehension and production of verbs with different argument structures and of sentences with different levels of syntactic complexity. The test includes five subtests: Verb Naming Test (VNT), Verb Comprehension Test (VCT), Argument Structure Production Test (ASPT), Sentence Production Priming Test (SPPT) and Sentence Comprehension Test (SCT). Three subtests (VNT, VCT, ASPT) assess the ability to produce and comprehend verbs in isolation and within a sentence context, with respect to the number of verb arguments and argument optionality. Two subtests (SPPT, SCT) investigate production and comprehension of sentences with different syntactic complexity, respectively, based on the order of constituents

(canonical vs. non-canonical) and on the type of syntactic movement (NP-movement vs. Wh-movement).

Cho-Reyes and Thompson (2012) found that individuals with agrammatic or anomic aphasia encountered greater difficulty in producing three-argument than two- and one-argument verbs on both the VNT and the ASPT. In addition, agrammatic (but not anomic) individuals demonstrated difficulty in producing verbs with optional arguments on the ASPT. These findings are in line with the cross-linguistic evidence showing that verb production in agrammatic aphasia is affected by verb-argument structure complexity (see below). With respect to sentence processing, production and comprehension of non-canonical sentences were largely impaired in the agrammatic group (as previously reported in the literature, see below), whereas anomic individuals showed lower accuracy only for the most complex structures, i.e. object-relative sentences, and only in production. Results obtained by Cho-Reyes and Thompson (2012) point to the NAVS as a useful measure for capturing verb and sentence processing difficulties that are typical of agrammatism. However, a comparable measure of verb and syntactic processing has not yet been developed for use with native speakers of Italian.

1.1. Argument structure processing in healthy and aphasic individuals

Verb-argument structure specifies the number of participants in the action described by a verb, as well as their thematic role (Agent, Theme, etc.) and the type of syntactic constituents that need to be assigned to the verb arguments. Based on the number of subcategorized arguments, transitive verbs, like carry (e.g. *Mary was carrying many bags*), require more arguments than intransitive verbs, like sleep (e.g. *John was sleeping*). For some verbs (e.g. *eat*), arguments can be optional, as they can be omitted without generating a violation (e.g. *Sam was eating* is equally acceptable as *Sam was eating a sandwich*). Among intransitive verbs, unergative verbs (e.g. *sleep*) differ from unaccusative verbs (e.g. *disappear*) in the thematic roles of their grammatical subjects (Agent for unergatives, Theme for unaccusatives).

Theoretical approaches to argument structure are generally classifiable into ‘lexicalist’ and ‘constructivist’. The lexicalist approach assumes that verb-argument structure is part of the verb lexical representation, and is stored within the lemma (Levelt, Roelofs, & Meyer, 1999; Rappaport Hovav & Levin, 1998). According to this view, verb-argument structure information becomes available as soon as the verb is accessed, and is then mapped onto the sentence structure. Based on the evidence that healthy individuals tend to use a verb (e.g. *give*) with a double-object construction (e.g., *John gave Mary the book*) following a prime sentence with double-object construction but a different verb, some authors (see Pickering & Branigan, 1998) have proposed models where verb-argument structure information is encoded in “combinatorial nodes” that are activated only when a verb is used with a particular construction. According to the constructivist approach, verb lexical entries do not entail a separate storage for argument structure properties, rather verb-argument structure is determined by specific meanings and/or discourse functions, and stored as a construction, i.e. as “pair of form and meaning” (Goldberg, 2003; Hale & Keyser, 2002).

Experimental evidence for an effect of argument structure complexity on verb and sentence processing in healthy individuals is limited to a few psycholinguistic studies (Shapiro, Zurif, & Grimshaw, 1989; Ahrens & Swinney, 1995; Ahrens, 2003; Friedmann, Taranto, Shapiro, & Swinney, 2008; McAllister, Bachrach, Waters, Michaud, & Caplan, 2009; Shapiro & Levine, 1990; Shapiro, Zurif, & Grimshaw, 1987), indicating longer reaction times (RTs) to two- or three-argument verbs than to one-argument verbs, to unaccusative (vs. unergative) verbs, and to optionally transitive (e.g. *move*) than to obligatory transitive (e.g. *exhibit*) verbs. Results suggesting that two- and three-argument verbs are associated with a greater processing cost than one-argument verbs, and that unaccusative verbs require greater processing resources than unergative verbs, have also been provided by neuroimaging studies, while neuroimaging findings regarding the greater processing cost associated with optional than obligatory transitive verbs are largely inconsistent (Den Ouden, Fix, Parrish, & Thompson, 2009; Meltzer-Asscher, Mack, Barbieri, & Thompson, 2015; Shetreet, Friedmann, & Hadar, 2010)

Clear evidence that argument structure complexity affects verb/sentence processing comes from the literature on agrammatic aphasia, an acquired language disorder often caused by stroke and characterized by production of short simple sentences (2–3 words) in which function words (e.g. prepositions) and bound morphemes (e.g. verb suffixes) are often omitted, and by a general difficulty in retrieving verbs (vs. nouns). Processing of verbs and sentences in agrammatic aphasia (across tasks and across languages) is increasingly difficult as the number of arguments increases (Barbieri, Basso, Frustaci, & Luzzatti, 2010; Barbieri, Aggujaro, Molteni, & Luzzatti, 2015; De Bleser & Kauschke, 2003; Dragoy & Bastiaanse, 2010; Kiss, 2000; Luzzatti et al., 2002; Sanchez-Alonso, Martinez-Ferreiro, & Bastiaanse, 2011; Thompson, Lange, Schneider, & Shapiro, 1997; Thompson, 2003). Furthermore, individuals with aphasia evince more difficulty when processing unaccusative than unergative verbs (Bastiaanse & van Zonneveld, 2005; Kegl, 1995; Lee & Thompson, 2011; Luzzatti et al., 2002; McAllister et al., 2009; Sanchez-Alonso et al., 2011; Thompson, 2003), and obligatory transitive verbs are more difficult than optional transitive verbs (Shapiro & Levine, 1990; Shapiro et al., 1987, 1989). These findings led to the formulation of the Argument Structure Complexity Hypothesis (Thompson, 2003), according to which the pattern of verb production deficits associated with agrammatism is partly accounted for by argument structure complexity, i.e. verbs with more complex argument structure are more difficult to produce as they entail a more complex lexical representation.

1.2. Syntactic complexity in healthy and aphasic individuals

Following the Government and Binding Theory (Chomsky, 1981), syntactically complex sentences are derived from simpler structures through syntactic operations that result in changes in the constituent order within the sentence. So for example, passive sentences (see (1)) are derived via movement of the post-verbal noun phrase (NP, i.e. *the cat*) from the position occupied by the direct object to the position occupied by the grammatical subject. As a result of this process (NP-movement), a trace, i.e. a phonologically silent element co-indexed with the moved NP (t_i), is left behind, thereby ensuring correct assignment of thematic roles, i.e. Theme to the moved NP (i.e. *the cat*) and Agent to the post-verbal prepositional phrase (PP, *by the dog*). Similarly, other non-canonical sentences such as

object-relative sentences (2) are derived through movement (Wh-movement) of the relative pronoun *who* from the direct object position to a non-argument position.

(1). *The cat*_i is followed *t*_i by the dog

(2). The cat *who*_i the dog is following *t*_i is black

Contrary to the Government and Binding Theory, some theoretical approaches (see Bresnan, 2000; Pollard & Sag, 1994) attribute the complexity of sentences in (1) and (2) to lexical-semantic processes, without postulating derivation from simpler structures through syntactic movement. Both approaches define, in Subject-Verb-Object (SVO) languages like English and Italian, sentences in (1) and (2) as *non-canonical*, because the Agent (*the dog*) follows the Theme (*the cat*), as opposed to *canonical* sentences where the Agent precedes the Theme. Studies on healthy English speakers have shown that non-canonical sentences require greater processing demands than canonical sentences (Choy & Thompson, 2010; Ferreira, 2003; Ford, 1983; Gordon, Hendrick, Johnson, & Lee, 2006; Osterhout & Swinney, 1993; Stromswold, Caplan, Alpert, & Rauch, 1996). Non-canonical sentences are also more often impaired (both in production and comprehension) than canonical sentences in agrammatic aphasia (Bastiaanse & Edwards, 2004; Caplan & Futter, 1986; Faroqi-Shah & Thompson, 2003; Grodzinsky, Piñango, Zurif, & Drai, 1999; Linebarger, Schwartz, & Saffran, 1983; Thompson, Tait, Ballard, & Fix, 1999; Yarbey Duman, Altinok, Özgür, & Bastiaanse, 2011). In addition, treatment studies have provided evidence in favor of separate neurocognitive mechanisms underlying NP- and Wh-movement by showing lack of generalization of treatment gains from Wh-movement structures to NP-structures (Thompson & Shapiro, 2005; Thompson, Den Ouden, Bonakdarpour, Garibaldi, & Parrish, 2010; but see Stadie et al., 2008).

1.3. Syntactic complexity and structural ambiguity in Italian

Although both Italian and English are SVO languages, there are multiple differences between the two languages. Italian allows for SOV and OVS constructions more often than English, although mostly when a marked context is provided. In addition, Italian is a pro-drop (or null subject) language, where the grammatical subject can be omitted because the referent of the null subject can be inferred from verb morphology, and the thematic role is assigned to a phonologically silent element. This phenomenon may generate structural ambiguity as in (3a), where both an Agent-first and a Theme-first interpretation are plausible (although, in the absence of a preceding context, an Agent-first interpretation is preferred). Thematic role disambiguation can be achieved through morphological marking, as shown in (3b).

(3a). *Ha*_{Sing} *chiamato* *Maria*_{Sing}

(He/she) has called Maria / Maria has called (lit. Has called Maria)

(b). *Hanno*_{Plur} *chiamato* *Maria*_{Sing}

(They) have called Maria

The sentence in (3a) is an example of free inversion, a phenomenon typical of null subject languages (see Belletti, 2008), which is the overwhelming preference that Italian speakers use when answering questions like *Chi ha chiamato?* (Who called?) with sentences where the grammatical subject (and Agent) moves after the verb.

Structural ambiguity can be observed in Italian even when the grammatical subject is phonologically realized. Interrogative sentences introduced by the pronoun *chi* (see (4)) are particularly ambiguous in Italian (unless disambiguation through morphological marking is provided), as the only acceptable form for both subject Wh- and object Wh-questions is the “Chi-V-NP” structure (see Rizzi, 1982 for a discussion).

(4). *Chi*_{Subj/Obj} *sta baciando la donna?*

Who is kissing the woman? / Who is the woman kissing?

(5). *Mi mostri l'orso*_{Sing} *che morde*_{Sing} *il cane*_{Sing}

Show me the bear who is biting the dog / Show me the bear who the dog is biting

Relative clauses may also be structurally ambiguous in absence of morphological/pragmatic information (5), although studies conducted on adult healthy individuals (Carminati, Guasti, Schadee, & Luzzatti, 2006) found a preference for an Agent-first interpretation of (5), i.e. an interpretation where the NP *l'orso* (the bear) is the Agent. When disambiguating information is provided, studies on Italian children (see Guasti, Stavrakaki, & Arosio, 2007) have reported greater difficulty in comprehension of unambiguous object (vs. subject) relatives.

While processing of non-canonical sentences has been widely analyzed in English, only a few studies have investigated processing of NP- and Wh-movement in Italian agrammatic individuals, with results indicating impaired comprehension of passive sentences (Luzzatti et al., 2001) and sentences entailing Wh-movement (Garraffa & Grillo, 2008) in the face of a preserved comprehension of canonical sentences such as active and subject-relative structures.

1.4. Aim of the study

The present study examined young, healthy native Italian speakers' performance on an adaptation of the NAVS into Italian, to ascertain normal verb and sentence processing patterns with respect to verb argument structure and syntactic complexity, prior to development and standardization of the test with Italian aphasic speakers. The purpose of the present research is to provide preliminary data that are necessary to build a reliable Italian version of the NAVS. Because the demographics of the participant group recruited for this study are not representative of the range of age and education typically found in aphasic individuals, the findings of the present study are not immediately applicable to the clinical population. Nevertheless, the present research provides valuable information on the cross-linguistic reliability of the materials included in the NAVS.

Verb-argument structure complexity was investigated in terms of the number of arguments selected by the verb (i.e., one-, two-, or three-argument verbs) as well as argument

optionality, by contrasting transitive verbs that optionally take a second or third argument with transitive verbs that obligatorily require two or three arguments. Based on the available evidence from the psycholinguistic and neuroimaging literature on normal language processing in both English and Italian (Agnew, van de Koot, McGettigan, & Scott, 2014; Barbieri et al., 2011; Den Ouden et al., 2009; Meltzer-Asscher et al., 2015; Thompson et al., 2007), and following the Argument Structure Complexity Hypothesis (Thompson, 2003), longer reaction times for three- and two-argument verbs (vs. one-argument verbs) were expected. We also expected greater difficulty for optional compared to obligatory argument verbs based on the results of a few studies with healthy English speakers (Shapiro & Levine, 1990; Shapiro et al., 1987, 1989). As verb-argument structure properties largely overlap between Italian and English, we expected that verb-argument structure complexity would affect processing times of healthy Italian individuals in the same way as in English (see Table 1). However, no differences in accuracy were expected, given the simplicity of the tasks.

The present study also investigated the effects of syntactic complexity on comprehension and production of relative clauses, cleft, and simple declarative structures. Syntactic complexity was examined with respect to the order of constituents in the sentence, by comparing canonical (i.e. Agent-Theme) to non-canonical (Theme-Agent) order. In line with the original English version of the NAVS, sentences included in the SCT and SPPT subtests did not allow for disambiguation based on morphological or pragmatic information. Thus, in the Italian version, Wh-questions (included in the English version) were eliminated and replaced with cleft sentences because of language-specific constraints that render the word order of object Wh-questions identical to that of subject Wh-questions (see Rizzi, 1982). Based on evidence that healthy Italian individuals show a preference for an Agent-first interpretation of relative sentences in absence of disambiguating morphological information (Carminati et al., 2006), we did not expect any difference in accuracy between subject and object relatives, as well as between subject and object clefts. In addition, based on the literature on healthy speakers of English and Italian (see Arosio, Adani, & Guasti, 2009; Guasti et al., 2007, for studies with children), we expected longer reaction times for non-canonical (vs. canonical) sentences (see Table 1), and specifically for passive vs. active, object-cleft vs. subject-cleft and object-relative vs. subject-relative sentences.

2. Materials and methods

2.1. Participants

Twenty-one neurologically unimpaired participants (mean age = 22.4, mean education = 14.8), all students at the Department of Psychology at University of Milano-Bicocca, participated in the study. Participants were recruited through an automated system (Sona System, <https://psicologia.unimib.it/en/research/research-services>) and selected according to the following criteria: monolingual native speakers of Italian, no history of learning disabilities and/or neurological/psychiatric disorders. All participants were right-handed (with the exception of one left-handed participant) and had normal or corrected-to-normal vision. All participants provided informed consent prior to enrollment in the study, which

was approved by the Institutional Review Board of University of Milano-Bicocca. Participation in the study was compensated with student credits.

2.2. Materials

Tasks and stimuli were translated (or adapted) from the original version of the NAVS by the first author and the senior author, who are both native Italian speakers.

2.2.1. Subtests assessing verb and verb argument structure—Verbs from the original (English) NAVS VNT as well as distractors included in the original VCT, which did not differ in argument structure when translated into Italian, were pre-tested for naming agreement in a group of 20 healthy Italian participants. Based on this, for the Italian VNT, five verbs were replaced (crawl→walk, *camminare*; shave→drink, *bere*; tickle→water, *innaffiare*; send→plant, *piantare*; put→insert, *inserire*) and the final 22 verbs were used in all verb and verb argument subtests (see Appendix I, II and III), i.e. the Verb Naming Test (VNT), the Verb Comprehension Test (VCT) and the Argument Structure Production Test (ASPT). These were classified as one-argument ($n = 5$), two-argument ($n = 10$) or three-argument verbs ($n = 7$). Two- and three-argument verbs were further classified as taking optional ($n = 5$ two-argument verbs and $n = 5$ three-argument verbs) or obligatory arguments ($n = 5$ two-argument verbs and $n = 2$ three-argument verbs). Argument optionality (for two- and three-argument verbs) was determined by a corpus analysis performed using “NoSketch Engine” (Kilgariff et al., 2014) and containing ~1.9 million tokens derived from Italian websites. For each verb lemma, 100 occurrences were extracted and each occurrence was coded to determine if the optional argument (i.e. the second argument for two-argument verbs and the third argument for three-argument verbs) was explicitly realized in a sentence. The proportion of occurrences with a second or a third argument was then computed for each verb and averaged across verbs of the same class (obligatory: $M = 82$, $SD = 10.9$; optional: $M = 52$, $SD = 25.3$). A non-parametric two-tailed Kolmogorov-Smirnov test demonstrated that the distribution of argument optionality was significantly different between the two verb classes ($D = 0.7$, $p = .018$), and showed that instances where the optional argument was realized were significantly greater in the ‘obligatory’ than in the ‘optional’ category, as expected. Imageability norms were obtained from a group of Italian students in the Department of Psychology ($N = 20$; age: 22–31, education: 16–18 years) at University of Milano-Bicocca, who were asked to judge the ability of verbs to elicit mental images on a scale ranging from 1 (not imageable at all) to 5 (highly imageable). Frequency values were derived from a large corpus (350 million tokens) of Italian newspaper text (“La Repubblica”; Baroni et al., 2004). Non-parametric Wilcoxon test statistics indicated that verbs were matched across categories for word frequency (optional vs. obligatory: $p = .307$; one- vs. two-arguments: $p = .624$; two- vs. three-arguments: $p = .161$; one- vs. three-arguments: $p = .106$). Length in syllables was matched across argument number categories (one vs. two-arguments: $p = .269$; two- vs. three-arguments: $p = .912$; one- vs. three-arguments: $p = .278$), although verbs taking optional arguments were slightly longer than verbs taking only obligatory arguments ($p = .049$). No differences in imageability were found between verbs taking obligatory vs. optional arguments ($p = .467$) or between one- and two-argument verbs ($p = .123$), whereas three-argument verbs were less imageable than two-argument ($p = .063$) and one-argument verbs ($p = .042$). Given that different phones

may be more or less effective in triggering the voice key (Kessler, Treiman, & Mullennix, 2002), a measure of how fast the first phone triggered the voice key (compared to other phones) was computed based on the data provided by Kessler et al. (2002), reflecting the number of phones (in the English phonetic alphabet) that elicited faster RTs than the given phone (i.e., the higher the number, the slower the phone is in triggering the voice key). Non-parametric Wilcoxon test statistics indicated that verb types were matched for voice-onset time (optional vs. obligatory: $p = .499$; one vs. two arguments: $p = .722$; two vs. three arguments: $p = .522$; one vs. three arguments: $p = 1$).

For each verb, pictures of actions (640×480 pixels for VNT and ASPT, 188×193 to 251×188 pixels for VCT) were either taken from the English NAVS (for verbs that were shared between the two versions) or were drawn matching the style and size of the original drawings (for the newly introduced verbs). The same pictures were used for target verbs across all tasks (Fig. 1). In the VCT, pictures illustrating target verbs were presented in arrays containing four pictures displayed over two rows (Fig. 1b). In the ASPT, pictures of target verbs also displayed labels for the verb and for the participants in the action (Fig. 1c). In this task, verbs with optional arguments were presented twice (Appendix III), once with the extended argument structure (i.e. two arguments for optional two-argument verbs and three arguments for optional three-argument verbs) and once with the minimal argument structure (i.e. one argument for optional two-argument verbs and two arguments for optional three-argument verbs), for a total of 32 trials.

2.2.2. Subtests assessing syntactic complexity—Materials were adapted from the sentence production and sentence comprehension subtests (SPPT and SCT) included in the original version of the NAVS. The target verbs used in sentences were the same as in the original version, with the exception of one (*pull*) that was modified because Agent and Theme (*bambino*, boy, and *bambina*, girl) were morphologically related and thus likely to generate errors in production of bound morphemes in individuals with aphasia. The picture for this verb was therefore re-drawn so that the two participants were better described as *uomo* (man) and *donna* (woman).

Target sentences were translated from the original version of the tests, with the exception of subject and object Wh-questions, which were replaced with subject- and object-cleft sentences. The test therefore included a total of 30 sentences, half canonical (actives, subject relatives, subject clefts) and half non-canonical (passives, object relatives, object clefts; see Appendix IV). For each target sentence, picture stimulus pairs from the original SPPT and SCT were used (except for sentences with the verb to *pull*). The 30 picture arrays (448×336 pixels) displayed two pictures illustrating the same action and the same participants, although thematic roles were reversed, so that the Agent in one picture was the Theme in the other picture (Fig. 2).

2.3. Procedure

Participants sat in front of a laptop (Acer Travel Mate 15.4") connected to a pair of headphones and to a response box (200A, Psychology Software Tools), which was in turn connected to a microphone. Stimuli were presented through E-Prime (Psychology Software

Tools). Verbs in the VCT and sentences in the SPPT and SCT subtests were recorded with natural prosody and intonation and were digitized at 44.1 Hz. For comprehension tasks (VCT, SCT), participants were instructed to listen to verbs or sentences and then provide a response as soon as possible by pressing a button on the keyboard. The keyboard was preferred to the response box due to difficulties in performing the VCT, where the four pictures in each array were displayed over two rows, whereas buttons on the response box were displayed on the same row (see Fig. 3). Keys “F”, “L”, “V”, and “,” on the keyboard were labeled with different colors and used as response buttons for the VCT (all four keys) and the SCT (red and blue keys). For comprehension tasks, accuracy and reaction times (RT) were used as main dependent variables in the analyses.

For production tasks (VNT, ASPT, SPPT), participants were instructed to speak into a microphone, with RTs automatically triggered by voice onset and measured as the time interlapsing between stimulus presentation and response onset. Reaction times generated by involuntary responses or environmental sounds were excluded from the analysis. Responses to production tasks were audiotaped on a portable recorder (Philips Voice Tracker LFH0615), transcribed off-line, and coded for accuracy. A brief description of each subtest procedure follows.

2.3.1. Verb Naming Test (VNT)—In this subtest, participants were instructed to look at a fixation cross (800 msec) followed by a randomly selected picture appearing at the center of the screen, and then name the action as fast as possible. Pictures remained on the screen until the participants’ response, as they did for the subtests described below.

2.3.2. Verb Comprehension Test (VCT)—Participants were instructed to look at a fixation cross (800 msec) followed by a randomly selected picture array at the center of the screen (see Fig. 1b), listen to the verb that was played through headphones and then indicate which of the four pictures in each array matched the verb by pressing a key (Fig. 3) as fast as they could.

2.3.3. Argument Structure Production Test (ASPT)—Participants were instructed to look at a fixation cross (800 msec) followed by a randomly selected picture, and then produce a complete sentence that described the picture by using the words provided for the verb and the participants in the action.

2.3.4. Sentence Production Priming Test (SPPT)—Each trial began with a fixation cross (800 msec) followed by random presentation of a picture array and a prime sentence played through headphones, which described the picture on the left side of the array. Participants were instructed to listen to the prime and then produce a similar sentence to describe the picture on the right side of the array.

2.3.5. Sentence Comprehension Test (SCT)—Stimuli in this subtest were the same as in the SPPT. Following a fixation cross (800 msec), participants were instructed to look at a randomly selected picture array, listen to the sentence and indicate the picture that correctly matched the sentence by pressing a key.

2.4. Data analysis

Analyses were performed in R, version 3.3.3. (R Core Team, 2016). Logistic mixed-effects regression analyses (Jaeger, 2008), using the *glmer()* function included in the *lme4* package (Bates, Mächler, Bolker, & Walker, 2014), were run on the accuracy data. Models containing main effects for each of the fixed effects and random effects for item and participant were compared through the *anova()* function in the *lme4* package. Linear mixed-effects regression analyses (Baayen, Davidson, & Bates, 2008) were conducted on RT data using the *lmer()* function in the *lme4* package. Regression analyses on RTs were performed using a backward automated procedure included in the *lmerTest* package version 2.0–30 (Kuznetsova, Brockhoff, & Bojesen Christensen, 2016). For each subtest, the full factorial model was implemented by including all the fixed factors and their interactions, as well as random effects for *item* and *participant*, and random slopes for each of the fixed factors. The automated backward procedure progressively removed non-significant factors to reach the best-fit model for the data. This model was then refitted after outliers were identified and excluded (using a 2.5 SD criterion above the model standardized residuals, see Baayen, 2008; Tremblay & Ransijn, 2015), to ensure that results were not driven by outliers. For both accuracy and RT analyses, statistical significance for each predictor was determined by computing p-values using Satterthwaite approximation, as provided within the *lmerTest* package. In the presence of factors with three or more levels, planned comparisons were carried out using the *multcomp* package (Hothorn, Bretz, & Westfall, 2008), applying Tukey correction for multiple comparisons. All the reported p-values indicate significance levels after performing correction for multiple comparisons (e.g., a p-value of 0.06 indicates a corrected marginally-significant p-value). In the presence of significant effects, R-squared values for the best-fit model (fixed effects only) were computed using the *r.squaredGLMM()* function within the *MuMIn* package (Nakagawa & Schielzeth, 2013). In addition, power calculations were run on the best-fit model, to assess the ability of the model to detect the fixed effects, following Brysbaert and Stevens (2018), and using the *powerSim()* function provided within the *simr* package (see Green & MacLeod, 2016), with N = 200 Monte Carlo simulations. Models yielding an observed power of at least 80% were considered as sufficiently powered.

To determine the effect of the *number of verb arguments*, optional two-argument and obligatory two-arguments verbs were grouped together, as were optional and obligatory three-arguments verbs (as in the original work by Cho-Reyes & Thompson, 2012). To assess the effect of *argument optionality*, obligatory one, two and three-argument verbs were grouped together as were optional two- and three-arguments verbs. *Verb frequency*, *verb imageability*, and *voice onset speed* (see above) were mean-centered to avoid collinearity (Iacobucci, Schneider, Popovich, & Bakamitsos, 2016; Irwin & McClelland, 2001). For the VNT and VCT, the full factorial model included the following fixed factors: number of verb arguments, argument optionality, verb imageability, and verb frequency (and all their interactions, with the exception of the interaction number of arguments*argument optionality, because the model was not full factorial). To ensure that differences between verb classes were not a byproduct of differences in the effectiveness of the first phoneme in triggering the voice key, once the best-fit model for the VNT was achieved, the *voice onset speed* was entered as a covariate in the model. Notably, this required the removal of one item

(*ridere*, laugh), for which no data regarding the voice key triggering speed were available in Kessler et al. (2002). This procedure was not adopted for the ASPT, where all sentences began with a definite article (*il* or *la*), which was of equal speed in triggering the voice key. In the ASPT, the *number of verb arguments* was tested in the same way as described for VNT/VCT subtests (i.e. by classifying verbs based on the number of possible arguments) and by classifying sentences based on the number of realized arguments. In this subtest, *sentence length* (measured in number of syllables, and mean centered) was introduced as a fixed factor in all the analyses, together with *verb frequency* and *verb imageability*.

Analyses for SPPT and SCT were conducted first by attempting to reach the best fit model with word order (canonical vs. non-canonical) as the main predictor of interest, i.e. by aggregating data from passive, object-cleft and object-relative sentences (non-canonical order) and active, subject-cleft, and subject-relative sentences (canonical order). A second analysis was then conducted using *sentence type* as the main predictor of interest (with 6 levels). Based on the idea that syntactically complex sentences may contain more pauses, and therefore affect sentence production planning (which typically requires planning at least up to the verb, see Ferreira, 2000) on the SPPT subtest, a variable reflecting the *speech rate* of the prime sentence, was computed by dividing sentence length in syllables by the total duration (in seconds) of the prime sentence. This variable was used as a proxy of the individual participants' speech production rate and introduced (after being mean-centered) in the analyses conducted on the SPPT, to ensure that differences in syntactic complexity were not a mere reflection of differences in sentence speech rate. Finally, *sentence length* (in syllables) was mean-centered and introduced as a predictor in the analyses on both the SPPT and the SCT subtests.

3. Results

3.1. Accuracy

Performance on all subtests was high, with the exception of one subject who performed more than 2.5 SD below the average on the VNT and was therefore excluded from the analyses on this subtest. Accuracy on the VNT and SPPT subtests (94% and 95% respectively) was relatively lower than on other subtests (all above 98%).

3.1.1. Verb and verb-argument structure subtests—Mean accuracy by verb type on the VNT, VCT and ASPT is shown in Table 2. Logistic regression analyses were run only on the VNT (because accuracy on the VCT and ASPT was above 96% across all verb types). Table 3 reports statistics for the model after that the random effect of item was removed, given that the full model failed to converge. Results of this analysis indicate no effect of the number of verb arguments on accuracy.

3.1.2. Sentence production and comprehension subtests—Accuracy on the SPPT and SCT is reported in Table 2. Logistic regression analyses were not run on data from the SPPT, since accuracy was 95% and above across all sentence types. On the SCT, logistic regression analyses indicate lower accuracy for sentences with non-canonical than canonical order, and a significant effect of sentence type, with object-cleft sentences being comprehended less accurately than both subject clefts and passives (see Table 3).

3.2. Reaction times

Incorrect responses, null responses due to technical failure, and disproportionately long RTs (i.e., RT greater than 5000 ms on the verb subtests, and greater than 7000 ms on the sentence subtests) were deleted prior to conducting the regression analysis on reaction times.

Disproportionately long RTs were very rare on all subtests (VNT: N = 2; VCT: N = 0; ASPT: N = 3; SPPT: N = 4; SCT: N = 2).

3.2.1. Verb and verb-argument structure subtests—The best-fit model for reaction times on the VNT included the *number of verb arguments*, *argument optionality*, *verb imageability*, *verb frequency*, and the interaction between verb imageability and verb frequency. Reaction times were longer for three-argument than for one-argument verbs, and a marginally significant difference was found between two- and one-argument verbs as well as between three- and two-argument verbs, with longer RTs as the number of verb arguments increased (Table 4). In addition, RTs to the task were predicted by argument optionality, in a different direction than expected (Shapiro et al., 1987, 1989), i.e. longer RTs were found for verbs with obligatory arguments (Table 4), although the mean difference in RTs was relatively small (27 ms). Furthermore, when looking at *argument optionality* as a continuous variable, i.e. when entering the proportion of occurrences (for each verb) in a sentence where the second or the third argument were explicitly realized based on the data derived from the NoSketch database (see Materials), the effect was not significant. Finally, imageability (both as main effect and in interaction with frequency) was also a significant predictor in the analysis, as highly imageable verbs were associated with faster response times (Table 5). Although the introduction of this variable significantly improved the model fit ($\chi^2 = 5.992$, $p = .014$), as confirmed by the increase in the variance explained by the model (R-squared: 0.167), both the *number of arguments* and the *argument optionality* effects remained significant, with the exception of the planned comparison between two- and three-argument verbs (one- vs. two-arguments: $p = .048$; three- vs. one- argument: $p = .019$; two- vs. three-arguments: $p = .233$; optional vs. obligatory: $p = .002$).

On the VCT, the best model included only the number of verb arguments among the fixed effects, and marginal differences between three- and two-argument verbs and between three- and one-argument verbs were observed, again in the direction of three-argument verbs being associated with longer response times (Tables 4 and 5).

On the ASPT, collinearity issues prevented inclusion of *verb frequency*, *verb imageability* and *syllable length* all in the same model. Therefore, separate models were built for each of these variables, indicating that none of them was a significant predictor of RTs on this task (*frequency*: $p = .174$; *imageability*: $p = .896$; *syllable length*: $p = .113$). The best-fit model for this subtest included the *number of realized arguments* as the only fixed effect. However, given that sentences with different number of realized arguments also differed in length, *syllable length* was entered as a covariate in this model (see Tables 4 & 5). As shown in Table 5, sentences where two- or three-argument verbs appeared with all their arguments (Agent and Theme for two-argument and Agent, Theme and Goal for three-argument verbs) elicited longer RTs than sentences where the only realized argument was the Agent. Similarly, sentences where a third argument was present elicited longer RTs than sentences

where the verb occurred with only two arguments. Interestingly, the *number of arguments* subcategorized by verbs (i.e., the number of possible arguments) was not a significant predictor of RTs to this task.

3.2.2. Sentence production and comprehension subtests—On the SPPT, collinearity issues prevented inclusion of *sentence length* and *speech rate* in the same model. Therefore, separate models were built for each of these variables and indicated that while *sentence length* was a significant predictor of RTs to this task (see Table 6), *speech rate* was not ($p = .649$). In addition, *word order* was eliminated by the automated backward elimination procedure, suggesting that RTs to canonical and non-canonical sentences did not differ significantly when sentence length was taken into account. Turning to the analysis of differences among sentence types, the best-fit model included both *sentence type* and *sentence length* as fixed effects (Table 6). Again, sentence length significantly predicted RTs to the task, with longer RTs observed for long versus short sentences. Planned comparisons for the effect of sentence type revealed longer RTs for object-cleft versus subject-cleft sentences, and for object-relative versus subject-relative sentences. No difference emerged between active and passive sentences. Among canonical sentences, subject-relative elicited longer RTs than subject-cleft sentences, which in turn elicited longer RTs than active sentences. No significant difference in RTs emerged between subject relatives and actives when sentence length was introduced in the model (see Tables 4 and 6). Among non-canonical sentences, longer RTs were observed for object-relative versus passive and object-cleft sentences (Table 6), while no difference was found between object-cleft and passive sentences.

On the SCT, both *word order* and *sentence length* were significant predictors, i.e. RTs to the task were longer for long than for short sentences, and longer RTs were elicited by non-canonical than by canonical sentences (Tables 4 and 6). Among sentence types, object clefts elicited longer RTs than subject clefts, and object-relative sentences were associated with longer RTs than subject relatives. No difference was found between active and passive sentences. Among canonical sentences, subject relatives elicited longer RTs than subject clefts, and no difference was found either between subject relatives and actives or between subject clefts and actives. Among non-canonical sentences, object-relative sentences elicited longer RTs than both passives and object clefts, while no difference was found between object-cleft and passive sentences (see Tables 4 and 6).

Power simulations conducted on models for accuracy and RTs indicated that all models were sufficiently powered (i.e., observed power greater than 80%), with the exception of the model for accuracy on the VNT (35%) and the models for RTs on the VCT (63.5%) and ASPT (65.5%). Results derived from these models should therefore be considered with caution.

4. Discussion

The present study investigated verb and sentence processing in healthy speakers of Italian, to generate production and comprehension patterns to use as a basis for development of an Italian version of the NAVS. Specifically, the aim was to test differences in 1) verb processing patterns, based on verb-argument structure properties and 2) sentence processing

patterns, based on syntactic complexity, as reported in the literature on healthy and aphasic individuals. Verb processing was investigated with respect to both the number of verb arguments (one, two or three) and the optionality of verb arguments. Verb processing was assessed by the first three subtests: the Verb Naming Test (VNT), the Verb Comprehension Test (VCT), and the Argument Structure Production Test (ASPT). Sentence processing was investigated with respect to the order in which arguments appeared in a sentence (canonical vs. non-canonical), and both sentences entailing NP-movement (passives) and Wh-movement (object-cleft and object-relative sentences) were included among non-canonical structures. Sentence processing was assessed by the Sentence Production Priming Test (SPPT), which requires production of sentences in response to a picture, based on a previously presented sentence-picture prime of the same type, and by the Sentence Comprehension Test (SCT), where participants were asked to match a given sentence with the correct picture.

4.1. Verb processing patterns

Argument structure complexity did not affect task accuracy on any of the three tests assessing verb production and comprehension. This is not surprising, given that all the selected verbs had been pre-tested on a separate group of healthy participants, and elicited high naming agreement. However, longer RTs were observed for three-argument (vs. one-argument) verbs and for three-argument (vs. two-argument) verbs on the VNT, the VCT and the ASPT; longer RTs for two-argument (vs. one-argument) verbs were also found on the VNT and the ASPT, but not on the VCT. These results are in line with the predictions, and contribute to the existing psycholinguistic (Ahrens & Swinney, 1995; Ahrens, 2003; Barbieri et al., 2011; Malyutina & den Ouden, 2017) and neuroimaging (Ben-Shachar, Hendler, Kahn, Ben-Bashat, & Grodzinsky, 2003; Ben-Shachar, Palti, & Grodzinsky, 2004; Den Ouden et al., 2009; Meltzer-Asscher et al., 2015; Shetreet et al., 2010; Thompson et al., 2007, 2010) literature on normal language processing that indicates that greater processing resources are needed for verbs subcategorizing for two or more arguments (vs. one-argument verbs). Results are also in line with cross-linguistic studies reporting that verb processing in aphasia is more difficult as the number of verb arguments increases (Barbieri et al., 2015; De Bleser & Kauschke, 2003; Dragoy & Bastiaanse, 2010; Luzzatti et al., 2002; Sanchez-Alonso et al., 2011; Thompson et al., 1997; Thompson, 2003) and with the data reported by Cho-Reyes and Thompson (2012) for the original English version of the NAVS. Finally, these data are also in line with the predictions made by the Argument Structure Complexity Hypothesis (Thompson, 2003), which suggests that the more complex the argument structure (both in terms of number and type of arguments), the more difficult verb production will be for individuals with aphasia. The present study therefore indicates that argument structure complexity effects found in individuals with agrammatic aphasia may reflect more general patterns of normal language processing. The greater processing cost of verbs selecting for more than one argument may result, as suggested by the Argument Structure Complexity Hypothesis, from a more complex lexical representation, plausibly at the lemma level (Levelt et al., 1999; Pickering & Branigan, 1998). The finding of argument structure complexity effects both during verb retrieval and verb comprehension can be accounted for by classical lexicalist approaches to verb-argument structure processing. According to these approaches, verb-argument structure information is stored within the

verb lemma and is inherently activated as soon as a verb lemma is accessed (Levelt et al., 1999). Since verb lemmas are accessed both during verb retrieval and auditory verb comprehension, longer RTs to three- and two-argument (vs. one-argument) verbs in these two tasks are likely due to the greater amount of information contained in the lemmas of verbs selecting for more than one argument. However, Pickering and Branigan's (1998) model states that verb-argument structure information is stored within so-called "combinatorial nodes" that become active only when the verb is used with a particular construction. Assuming that verb naming implies building a minimal syntactic structure, as suggested by some authors (Cairns, Marshall, Cairns, & Dipper, 2007), Pickering and Branigan's (1998) model may still account for the results obtained on the VNT: according to this interpretation, after the verb lemma is accessed, activation will spread to the combinatorial nodes in the same way as if a sentence was produced. Since the amount of argument structure information stored in the combinatorial nodes is greater for three- and two-argument (vs. one-argument) verbs, RTs for the former will be longer than RTs for the latter. On the VCT, however, verbs were presented in absence of any combinatorial information/sentential context: while it is safe to assume that verb comprehension would proceed from performing an auditory/perceptual analysis of the input to accessing the phonological word form and then the semantic content and the verb lemma, it seems unlikely that the combinatorial nodes would be activated under these task demands. Hence, results obtained on the VCT can be explained under the classical lexicalist approaches, i.e. by assuming that verb-argument structure information is stored within verb lemmas, and are not fully accounted for by Pickering and Branigan's approach. Furthermore, data derived from the ASPT mirror the patterns found on the VNT, although with a crucial difference: on this subtest, the effect of argument structure complexity was observed for the actual number of realized arguments rather than (as in the VNT) for the number of possible arguments subcategorized by a verb. It should be noted that the ASPT, as opposed to the VNT, does not require verb retrieval, but rather requires participants to build a sentence - after the verb and its arguments are provided - by correctly mapping the verb arguments onto the sentence structure. Therefore, it is plausible to assume that the effect of argument structure complexity on RTs to the ASPT is reflecting processes at the positional level rather than at the functional (lemma) level (see Bock & Levelt, 1994), i.e. that the longer RTs observed for sentences containing two or three (vs. one) arguments reflect a greater processing cost in mapping arguments onto the sentence structure. Nevertheless, based on the evidence that the models for data obtained from the VCT and ASPT were not sufficiently powered and that fixed effects explained a low proportion of the total variance (7.5% and 3% respectively), results derived from these subtests should be considered with caution.

Turning to argument optionality, longer RTs were observed for verbs subcategorizing obligatory (vs. optional) arguments on the VNT when the variable was used as categorical, although the mean difference in RTs was relatively small. Furthermore, the absence of a significant effect when the variable was used as continuous suggests that processing times to optional transitive verbs may be similar to obligatory transitive verbs. Either way, results are not in line with our predictions, which were based on Shapiro and coworkers' findings (Shapiro & Levine, 1990; Shapiro et al., 1987, 1989). According to Shapiro and colleagues, verbs with optional arguments allow for multiple argument structure alternations, one with

and one without the argument (“two frames theory”, see Van Valin & LaPolla, 1997), and – whenever the verb lemma is accessed – all the available argument structure alternations become active, thereby resulting in longer RTs. Our results are not in line with this assumption. The discrepancy in results between the present study and the findings of Shapiro and colleagues may stem from differences in task paradigms (verb naming vs. cross-modal lexical decision task). Previous research conducted using both tasks (i.e., naming and lexical decision), however, has reported either no difference between verbs carrying optional (vs. obligatory) arguments, or greater processing demands for verbs carrying obligatory arguments (see e.g. Ahrens & Swinney, 1995; Malyutina & den Ouden, 2017; Meltzer-Asscher et al., 2015; Shetreet, Palti, Friedmann, & Hadar, 2006, 2010). In one of these studies, Shetreet et al. (2010) proposed that the lexical representation of verbs carrying optional arguments contains only one argument structure frame, which incorporates the optional argument (Chierchia, 1995; Rizzi, 1986); however, when the verb is implemented in a sentence without the optional argument, the unassigned thematic role undergoes a lexical operation called saturation, by which the optional argument is not projected on the syntax (Bresnan, 1982). Based on this account, it would be expected that – when accessing and retrieving verbs as singletons – verbs like *guidare* (*drive*, where the second argument is optional) would not differ from verbs like *tagliare* (*cut*, where the second argument is obligatory) and both would elicit longer RTs than verbs like *camminare* (*walk*, which only have one argument). Results obtained from analyses where argument optionality was entered as a categorical variable (as done in previous studies) are not in line with this interpretation. Malyutina and den Ouden (2017) suggested a third possibility – that the lexical representation of verbs carrying optional arguments are encoded with two argument structure frames, but only one (the most prominent) is accessed, although which of them is the most prominent is unclear from their study and may depend on task demands. Results obtained from analyses where argument optionality was entered as a categorical variable are in line with this interpretation and suggest that the optional frame is putatively most prominent. However, in light of the relatively small difference in RTs between optional transitive and obligatory transitive verbs, and given that analyses conducted by entering argument optionality as a continuous variable led to different results (i.e., no difference in RTs between predominantly optional and predominantly obligatory arguments), it remains unclear which theoretical account best explains our findings. Nevertheless, results of both analyses are not in line with Shapiro’s hypothesis of a greater complexity of verbs subcategorizing for optional arguments (vs. verbs with an established set of arguments). Contrary to the VNT, no effect of argument optionality was found on either the VCT or the ASPT. However, as earlier mentioned, analyses conducted on these two subtests were underpowered, which could have prevented results from reaching statistical significance.

Finally, our findings indicate, not surprisingly, a significant effect of imageability on the VNT, in the direction of faster RTs to highly imageable verbs, even though the main effect of number of arguments and argument optionality was significant even when imageability, frequency and their interaction were introduced in the model. The imageability effect found in the present study indicates that greater processing resources are required for verbs with low (vs. high) imageability, in line with evidence on noun processing coming from neuroimaging studies conducted on healthy individuals (Binder, Westbury, McKiernan,

Posing, & Medler, 2005; Jessen et al., 2000) and from studies conducted on individuals with aphasia (Bates, Burani, D'Amico, & Barca, 2001; Berndt, Haendiges, Burton, & Mitchum, 2002; Crepaldi et al., 2006; Luzzatti et al., 2002), thereby suggesting that imageability affects lexical access for verbs as it does for nouns. In addition, our results indicate a significant effect of frequency on RTs, but only for low imageable verbs, suggesting that – when verbs are less highly imageable – verb frequency plays an important role in lexical retrieval. On the contrary, when verbs are highly imageable, frequency does not affect lexical retrieval (see Wise et al., 2000 for similar results on nouns). Notably, no effect of imageability/frequency was found on either the VCT or the ASPT. This result is in line with the literature (see Bates et al., 2001; see also Berlingeri et al., 2008 for a review) showing that frequency and imageability mostly play a role during verb retrieval, a process that was not required by either the VCT or the ASPT, where all the lexical items were provided. However, once again, results obtained from these two subtests were derived from underpowered analyses, which could have prevented frequency and imageability effects from emerging.

4.2. Sentence processing patterns

Results on the SCT and SPPT showed no differences in either accuracy or RTs between active and passive sentences. While the first finding was expected, given the relative simplicity of the task, we expected passive sentences to elicit longer RTs than actives in Italian, based on prior evidence from English psycholinguistic and neuroimaging studies (Feng et al., 2015; Ferreira, 2003; Hirotsu, Makuuchi, Ruschemeyer, & Friederici, 2011; Mack, Meltzer-Asscher, Barbieri, & Thompson, 2013; Yokoyama et al., 2006). On the one hand, task-related differences could account for the inconsistency between the present findings and the results reported by Ferreira (2003). On the other hand, the present study failed to replicate Mack et al. (2013) findings, in which a similar task was employed in comprehension. Contrary to Mack et al. (2013) study, we did not find any difference in processing times between active and passive sentences in Italian, likely due to the smaller number of observations in our study. Nevertheless, the greater complexity of passive (vs. active) sentences in comprehension is supported by the lower accuracy on passive (vs. active) sentences in Italian-speaking individuals with agrammatic aphasia (Luzzatti et al., 2001), as previously reported by studies on English and Turkish agrammatic participants (Bastiaanse & Edwards, 2004; Berndt, Mitchum, & Haendiges, 1996; Caplan & Futter, 1986; Grodzinsky et al., 1999; Linebarger et al., 1983; Yarbrough Duman et al., 2011).

Turning to relative and cleft structures, accuracy was equally high for subject- and object-relative sentences on both the SPPT and the SCT, whereas object clefts were comprehended less accurately than subject clefts. The finding of ceiling-level accuracy in comprehension of both subject relatives and subject clefts substantiates the claim that an Agent-first interpretation of these structures is preferred even in absence of morphological or pragmatic cues informing thematic role assignment, in line with previous studies (Carminati et al., 2006). This finding can be accounted for by the use of a heuristic strategy, the Minimal Chain Principle (De Vincenzi, 1991; Frazier & Flores D'Arcais, 1989), according to which healthy speakers tend to establish syntactic dependencies early on in the sentence in order to build the simplest syntactic structure. Therefore, in a sentence like *Pietro vede il gatto che*

sta inseguendo il cane (Pete sees the cat who is chasing the dog), participants will establish a dependency between *che* and the first noun phrase (NP, *il gatto*) encountered in the sentence, thereby preferring an Agent-first interpretation. The RT analysis supports this hypothesis, given that both subject-relative and subject-cleft sentences did not elicit longer RTs than their non-canonical counterparts, as it would be expected if sentences were syntactically ambiguous and both interpretations (Agent-first and Theme-first) were accessed at the same time. Rather, RTs to object-relative and object-cleft sentences were longer than to subject-relative and subject-cleft sentences on both the SPPT and the SCT. This finding is consistent with data reported by Cho-Reyes and Thompson (2012) from aphasic participants for these sentence types, and indicates that production and comprehension of object-relative and object-cleft sentences engage greater processing resources than subject-relative and subject-cleft sentences. Given that verbs and nouns included in both sentence types were identical, it is safe to assume that this difference in processing cost stems from differences at post-lexical levels of processing, namely either in the way thematic roles are mapped onto the syntactic structure, or in the syntactic structure itself. Object-relative and object-cleft sentences entail a non-canonical assignment of thematic roles that – following Chomsky (1981) – is the result of Wh-movement, and is associated with processes of syntactic re-analysis that render object-relative and object-cleft sentences more difficult to produce and comprehend for individuals with aphasia (Caramazza & Zurif, 1976; Faroqi-Shah & Thompson, 2003). The greater processing cost associated with object relatives and object clefts versus their canonical counterparts may again result, according to some authors, (Arosio et al., 2009; Guasti et al., 2007) from the use of the Minimal Chain Principle (De Vincenzi, 1991; Frazier & Flores D’Arcais, 1989). For object relatives (e.g. *Pietro vede il cane_{Obj} che il gatto_{Subj} sta guardando*, Pete saw the dog who the cat is chasing), the Minimal Chain Principle predicts that a dependency between the Italian relative pronoun *che* and the closest noun phrase (*il cane*, the dog) will be established first, and then, when a second noun phrase is encountered, a re-analysis takes place and the role of Agent is re-assigned to the upcoming noun phrase (*il gatto*, the cat), thereby determining longer RTs for object relatives (and object clefts) than for subject relatives (and subject clefts). Notably, the processes of re-analyses engaged by object clefts also affected comprehension accuracy. It should be noted that all the hypotheses illustrated so far to account for the different processing of object and subject relatives assume that subject-relative sentences are restrictive relative clauses, i.e. that their syntactic structure is like the one in Fig. 4a. Although this is a safe assumption in English, some authors (Costa, Fernandes, Vaz, & Grillo, 2016; Grillo & Costa, 2014; Guasti, 1993) have suggested that in romance languages (e.g. Italian, Spanish, Portuguese) canonical sentences where the verb in the main clause is a perception verb may also be represented as pseudo-relatives (see Fig. 4b), i.e. as “apparent” relative clauses where the complementizer phrase is the predicate of a small clause. As explained in detail in Grillo and Costa (2014), in restricted relative clauses, the complement of the verb is a determiner phrase that is modified by the relative clause. In pseudo-relative constructions, on the other hand, the verb complement is a small clause that takes the determiner phrase as a subject. According to Grillo and Costa (2014), healthy speakers of Italian prefer – whenever available – a pseudo-relative construction (vs. a relative clause construction).

Note that pseudo-relative constructions are viable options only for canonical sentences (Guasti, 1993). Therefore, it is still safe to assume that the observed differences in processing times between subject and object relatives are the result of differences in the way thematic roles are mapped onto the syntactic structure (non-canonical for object-relative and canonical for subject-relative sentences). Furthermore, as opposed to relative clauses, it is still a matter of debate whether subject cleft sentences may allow for a pseudo-relative construction (see Belletti, 2008). Therefore, both interpretations provided for the longer RTs observed for object-relative/object-cleft (versus subject-relative/subject-cleft) sentences, i.e. thematic re-analysis or syntactic re-analysis triggered by Wh-movement, seem currently plausible. Finally, results revealed longer RTs for object relatives compared to all other non-canonical sentences (i.e. object clefts and passives) on both the SPPT and the SCT. The differences in RTs between object-relative and passive sentences were not unexpected, based on the evidence coming from the aphasia literature (Thompson & Shapiro, 2005) that points to qualitatively different processing underlying sentences containing Wh- vs. NP-movement. Object-relative sentences are syntactically more complex than passive sentences for a variety of reasons: 1) they differ in the underlying syntactic movement, as constituents move from argument position to another argument position in passive sentences (NP-movement) but from argument position to a non-argument position in object-relative sentences (Wh-movement); 2) they entail sentence-embedding, and 3) they include a greater number of nodes in the syntactic tree. All these factors can account for the greater processing cost of object-relative than passive sentences. However, in absence of a clear effect of NP-movement on processing demands in this study (as indicated by the evidence of equally fast RTs to active and passive sentences), the first hypothesis seems unlikely. Furthermore, subject raising in object-relative sentences may result in a more complex thematic re-analysis that could account for the greater processing cost associated with object-relative than object-cleft sentences (Thompson et al., 2010). Finally, it should be noted that object-relative sentences were longer than both object-cleft and passive sentences. However, sentence length was included as covariate in the analysis. Therefore, it is unlikely that RT differences may be accounted for by differences in length.

5. General discussion and conclusion

The present study indicates that the number of possible verb arguments affects the speed in both retrieving a verb from the lexicon and accessing the verb lexical representation from an auditorily presented verb, while the number of realized arguments affects reaction times on a task tapping into post-lexical processes. This finding is consistent with lexicalist approaches, according to which verb-argument structure information is stored within the verb lexical representation and becomes available as soon as the verb representation is active (Levelt et al., 1999), and less in line with Pickering and Branigan's model (1998). In addition, results suggest that argument structure complexity affects RTs in healthy individuals in similar ways as it affects accuracy in agrammatic individuals, thereby extending the Argument Structure Complexity Hypothesis (Thompson, 2003) from an account of agrammatic participants' performance patterns to a more general principle regulating normal language processing. The present study also demonstrates that verbs carrying optional arguments are processed faster than verbs carrying only obligatory arguments in picture naming, but not in

a word-to-picture matching task (word comprehension) or in a facilitated production task. Again, this result substantiates the idea that verb argument structure information is stored within the verb lexical representation and becomes active whenever the verb is accessed, rather than playing a role only when a verb is used with a particular construction. In addition, data regarding the effect of argument optionality are in contrast with the “two frames theory” (Van Valin & LaPolla, 1997; Shapiro et al., 1987, 1989), although further research is necessary to better understand how the lexical representations of verbs subcategorizing optional (vs. obligatory) arguments differ.

Turning to sentence processing, data confirm the greater processing cost associated with comprehension and production of non-canonical than canonical sentences, as reported in neuroimaging studies on healthy English speakers (2004, Ben-Shachar et al., 2003; Den Ouden et al., 2012; Shetreet & Friedmann, 2014; Thompson et al., 2010). However, the present study cannot help to determine whether this processing cost stems from thematic re-analysis or from computation of syntactic dependencies that are associated with Wh-movement processing. In addition, the study failed to replicate previous findings of longer reaction times to passive (vs. active) sentences (e.g. Mack et al., 2013). Finally, the present results inform about processing of cleft sentences in healthy speakers of Italian, by showing major processing similarities with relative sentences.

In conclusion, the study provides evidence that an adapted version of the NAVS is able to detect the effect of verb-argument structure and syntactic complexity on processing times in healthy speakers of Italian, by providing results that are similar to those reported by previous psycholinguistic and neuroimaging studies conducted on healthy English speakers. The present results demonstrate the cross-linguistic reliability of the materials included in this adapted version of the NAVS, and thus justify the realization of such Italian version, which will be validated on a group of Italian individuals with aphasia and of healthy participants of similar age and education. We acknowledge that the present findings do not have a direct clinical application, due to the differences in the range of age and education between participants in this study and that of individuals with aphasia. Data on a representative, neurologically intact population are being collected, and will serve as a proper comparison for the data obtained from individuals with aphasia. The study also highlights some language-specific aspects of sentence processing that render some structures (interrogative sentences) unsuitable to be used as testing material in the Italian language, and that render the interpretation of other structures (subject-relative sentences) more challenging than in English due to the availability of pseudo-relative constructions in romance languages. Finally, the study contributes to theories of verb processing and representation of verb-argument structure information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The present study was partially supported by the R01 DC001948 grant awarded by the NIH to Dr. Cynthia K. Thompson. Preliminary results were presented at the 31st European Workshop on Cognitive Neuropsychology,

Bressanone, January 20-25, 2013, and to the 51th meeting of the Academy of Aphasia, Lucerne, October 20-22, 2013. The authors would like to thank Dr. Maria Teresa Guasti, the associate editor and two anonymous reviewers for their helpful comments and suggestions on an earlier version of the manuscript, and Kathy Xie and Candace Todd for assistance with the preparation of the manuscript.

Abbreviations:

AAT	Aachen Aphasia Test
ASPT	Argument Structure Production Test
BDAE	Boston Diagnostic Aphasia Examination
CAT	Comprehensive Aphasia Test
ENPA	Esame Neuropsicologico per l'Afasia
NAVS	Northwestern Assessment of Verbs and Sentences
NP	noun phrase
OVS	Object-Verb-Subject
PP	prepositional phrase
RTs	reaction times
SCT	Sentence Comprehension Test
SSPT	Sentence Production Priming Test
SVO	Subject-Verb-Object
VAST	Verb and Sentence Test
VCT	Verb Comprehension Test
VNT	Verb Naming Test
WAB	Western Aphasia Battery

References

- Agnew ZK, van de Koot H, McGettigan C, & Scott SK (2014). Do sentences with unaccusative verbs involve syntactic movement? Evidence from neuroimaging. *Language, Cognition and Neuroscience*, 29, 1035–1045. 10.1080/23273798.2014.887125.
- Ahrens K (2003). Verbal integration: The interaction of participant roles and sentential argument structure. *Journal of Psycholinguistics Research*, 32, 497–516. 10.1023/A:1025452814385.
- Ahrens K, & Swinney D (1995). Participant roles and the processing of verbs during sentence comprehension. *Journal of Psycholinguistic Research*, 24, 533–547. 10.1007/BF02143166. [PubMed: 8531172]
- Arosio F, Adani F, & Guasti MT (2009). Children's processing of subject and object relatives in Italian e. al. (Eds.) In Gavarro A' (Ed.). *Merging features: Computation, interpretation and acquisition*. Oxford: Oxford University Press.
- Baayen RH (2008). *Analyzing linguistic data A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

- Baayen RH, Davidson DJ, & Bates DM (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412. 10.1016/j.jml.2007.12.005.
- Barbieri E, Basso A, Frustaci M, & Luzzatti C (2010). Argument structure deficits in aphasia: New perspective on models of lexical production. *Aphasiology*, 24, 1400–1423. 10.1080/02687030903580325.
- Barbieri E, Luzzatti C, den Ouden D-B, Aggujaro S, Verga R, & Thompson CK (2011). An investigation of the Argument Structure Complexity Hypothesis (ASCH) in English and Italian speakers: Data from aphasic and healthy participants, Poster presentation at Structuring the argument: A multidisciplinary workshop on the mental representation of verbal arguments. Paris, France.
- Barbieri E, Aggujaro S, Molteni F, & Luzzatti C (2015). Does argument structure complexity affect reading? A case study of an Italian agrammatic patient with deep dyslexia. *Applied Psycholinguistics*, 36, 533–558. 10.1017/S0142716413000337.
- Baroni M, Bernardini S, Comastri F, Piccioni L, Volpi A, Volpi R, ... Mazzoleni M (2004). Introducing the La Repubblica Corpus: A large, annotated, TEI(XML)-compliant corpus of newspaper Italian, Lisbon. *Proceedings of LREC 2004*, 1771–1774.
- Bastiaanse R, & Edwards S (2004). Word order and finiteness in Dutch and English Broca's and Wernicke's aphasia. *Brain and Language*, 89, 91–107. 10.1016/S0093-934X(03)00306-7. [PubMed: 15010241]
- Bastiaanse R, & van Zonneveld R (2005). Sentence production with verbs of alternating transitivity in agrammatic Broca's aphasia. *Journal of Neurolinguistics*, 18, 57–66. 10.1016/j.jneuroling.2004.11.006.
- Bastiaanse R, Edwards S, Maas E, & Rispens J (2003). Assessing comprehension and production of verbs and sentences: The Verb and Sentence Test (VAST). *Aphasiology*, 17, 49–73. 10.1080/729254890.
- Bates E, Burani C, D'Amico S, & Barca L (2001). Word reading and picture naming in Italian. *Memory & Cognition*, 29, 986–999. 10.3758/BF03195761. [PubMed: 11820758]
- Bates D, Mächler M, Bolker B, & Walker S (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Belletti A (2008). The CP of clefts. *Rivista di Grammatica Generativa*, 33, 191–204. <https://hdl.handle.net/10278/2509>.
- Ben-Shachar M, Hendler T, Kahn I, Ben-Bashat D, & Grodzinsky Y (2003). The neural reality of syntactic transformations: Evidence from functional magnetic resonance imaging. *Psychological Science*, 14, 433–440. 10.1111/1467-9280.01459. [PubMed: 12930473]
- Ben-Shachar M, Palti D, & Grodzinsky Y (2004). Neural correlates of syntactic movement: Converging evidence from two fMRI experiments. *NeuroImage*, 21, 1320–1336. 10.1016/j.neuroimage.2003.11.027. [PubMed: 15050558]
- Berlingeri M, Crepaldi D, Roberti R, Scialfa G, Luzzatti C, & Paulesu E (2008). Nouns and verbs in the brain: Grammatical class and task specific effects as revealed by fMRI. *Cognitive Neuropsychology*, 25, 528–558. 10.1080/02643290701674943. [PubMed: 19086201]
- Berndt RS, Haendiges AN, Burton MW, & Mitchum CC (2002). Grammatical class and imageability in aphasic word production: Their effects are independent. *Journal of Neurolinguistics*, 15, 353–371. 10.1016/S0911-6044(01)00030-6.
- Berndt RS, Mitchum CC, & Haendiges AN (1996). Comprehension of reversible sentences in "agrammatism": A meta-analysis. *Cognition*, 58, 289–308. 10.1016/0010-0277(95)00682-6. [PubMed: 8871341]
- Binder JR, Westbury CF, McKiernan KA, Possing ET, & Medler DA (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, 17, 1–13. 10.1162/0898929054021102. [PubMed: 15701235]
- Bock K, & Levelt WJ (1994). Language production: Grammatical encoding In Gernsbacher MA (Ed.). *Handbook of psycholinguistics* (pp. 945–984). San Diego: Academic Press.
- Bresnan JW (2000). *Lexical-functional syntax*. Oxford, UK: Wiley-Blackwell.
- Bresnan JW (1982). *The mental representation of grammatical relations*. Cambridge, MA: The MIT Press.

- Brysbaert M, & Stevens M (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1), 10.5334/joc.10.
- Cairns D, Marshall J, Cairns P, & Dipper L (2007). Event processing through naming: Investigating event focus in two people with aphasia. *Language and Cognitive Processes*, 22, 201–233. 10.1080/01690960500489644.
- Capasso R, & Miceli G (2001). *Esame Neuropsicologico per l'Afasia*: ENPA. Springer.
- Caplan D, & Futter C (1986). Assignment of thematic roles to nouns in sentence comprehension by an agrammatic patient. *Brain and Language*, 27, 117–134. 10.1016/0093-934X(86)90008-8. [PubMed: 3947937]
- Caramazza A, & Zurif EB (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, 3, 572–582. 10.1016/0093-934X(76)90048-1. [PubMed: 974731]
- Carminati S, Guasti MT, Schadee H, & Luzzatti C (2006). Subject and object relative clauses in Italian: Normal subjects and an agrammatic patient. *Brain and Language*, 99, 164–165. 10.1016/j.bandl.2006.06.091.
- Cecchetto C, Di Domenico A, Garraffa M, & Papagno C (2012). *Comprendo: Batteria per la comprensione di frasi negli adulti*. Milan: Raffaello Cortina.
- Chierchia G (1995). The variability of impersonal subjects In Bach E, Jelinek E, Kratzer A, & Partee BH (Eds.). *Quantification in natural language* (pp. 107–143). Dordrecht: Kluwer.
- Chomsky N (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Cho-Reyes S, & Thompson CK (2012). Verb and sentence production and comprehension in aphasia: Northwestern Assessment of Verbs and Sentences (NAVS). *Aphasiology*, 26, 1250–1277. 10.1080/02687038.2012.693584. [PubMed: 26379358]
- Choy JJ, & Thompson CK (2010). Binding in agrammatic aphasia: Processing to comprehension. *Aphasiology*, 24, 551–579. 10.1080/02687030802634025. [PubMed: 20535243]
- Ciurli P, Marangolo P, & Basso A (1996). *Esame del linguaggio-II*. OS
- Costa J, Fernandes B, Vaz S, & Grillo N (2016). (Pseudo-) Relatives and prepositional infinitival constructions in the acquisition of European Portuguese. *Probus*, 28, 119–143. 10.1515/probus-2016-0006.
- Crepaldi D, Aggujaro S, Arduino LS, Zonca G, Ghirardi G, Inzaghi MG, ... Luzzatti C (2006). Noun-verb dissociation in aphasia: The role of imageability and functional locus of the lesion. *Neuropsychologia*, 44, 73–89. 10.1016/j.neuropsychologia.2005.04.006. [PubMed: 15922372]
- De Bleser R, & Kauschke C (2003). Acquisition and loss of nouns and verbs: Parallel or divergent patterns? *Journal of Neurolinguistics*, 16, 213–229. 10.1016/S0911-6044(02)00015-5.
- De Vincenzi M (1991). *Syntactic parsing strategies in Italian: The minimal chain principle*. Springer Science and Business Media.
- Den Ouden D-B, Fix SC, Parrish TB, & Thompson CK (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, 22, 196–215. 10.1016/j.jneuroling.2008.10.004. [PubMed: 20160850]
- Den Ouden D-B, Saur D, Mader W, Schelter B, Lukic S, Wali E, ... Thompson CK (2012). Network modulation during complex syntactic processing. *NeuroImage*, 59, 815–823. 10.1016/j.neuroimage.2011.07.057. [PubMed: 21820518]
- Dragoy O, & Bastiaanse R (2010). Verb production and word order in Russian agrammatic speakers. *Aphasiology*, 24, 28–55. 10.1080/02687030802586902.
- Druks J, & Masterson J (2000). *An object and action naming battery*. Hove: Psychology Press.
- Faroqi-Shah Y, & Thompson CK (2003). Effect of lexical cues on the production of active and passive sentences in Broca's and Wernicke's aphasia. *Brain and Language*, 85, 409–426. 10.1016/S0093-934X(02)00586-2. [PubMed: 12744953]
- Feng S, Legault J, Yang L, Zhu J, Shao K, & Yang Y (2015). Differences in grammatical processing strategies for active and passive sentences: An fMRI study. *Journal of Neurolinguistics*, 33, 104–117. 10.1016/j.jneuroling.2014.09.002.
- Ferreira F (2000). *Syntax in language production: An approach using tree-adjoining grammars*. Aspects of language production, 291–330.

- Ferreira F (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164–203. 10.1016/S0010-0285(03)00005-7. [PubMed: 12948517]
- Ford M (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 203–218. 10.1016/S0022-5371(83)90156-1.
- Frazier L, & Flores D'Arcais GB (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, 28, 331–344. 10.1016/0749-596X(89)90037-5.
- Friedmann N, Taranto G, Shapiro LP, & Swinney D (2008). The leaf fell (the leaf): The online processing of unaccusatives. *Linguistic Inquiry*, 39, 355–377. 10.1162/ling.2008.39.3.355. [PubMed: 22822348]
- Garraffa M, & Grillo N (2008). Canonicity effects as grammatical phenomena. *Journal of Neurolinguistics*, 21, 177–197. 10.1016/j.jneuroling.2007.09.001.
- Goldberg AE (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, 7, 219–224. 10.1016/S1364-6613(03)00080-9. [PubMed: 12757824]
- Goodglass H, & Kaplan E (1983). Boston diagnostic aphasia examination. Lea & Febiger.
- Gordon PC, Hendrick R, Johnson M, & Lee Y (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 32, 1304–1321. 10.1037/0278-7393.32.6.1304.
- Green P, & MacLeod CJ (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. 10.1111/2041-210X.12504.
- Grillo N, & Costa J (2014). A novel argument for the universality of parsing principles. *Cognition*, 133, 156–187. 10.1016/j.cognition.2014.05.019. [PubMed: 25032904]
- Grodzinsky Y, Piñango MM, Zurif E, & Drai D (1999). The critical role of group studies in neuropsychology: Comprehension regularities in Broca's aphasia. *Brain and Language*, 67, 134–147. 10.1006/brln.1999.2050. [PubMed: 10092346]
- Guasti MT (1993). Causative and perception verbs A comparative study. Torino: Rosenberg & Sellier.
- Guasti MT, Stavrakaki S, & Arosio F (2007). Number and case in the comprehension of relative clauses: Evidence from Italian and Greek In Gavarro A', & Freitas MJ (Eds.). *Language Acquisition and Development: Proceedings of GALA 2007*. Newcastle: Cambridge Scholar Press.
- Hale KL, & Keyser SJ (2002). *Prolegomenon to a theory of argument structure*, 39 MIT press.
- Hirotani M, Makuuchi M, Ruschemeyer S, & Friederici AD (2011). Who was the agent? The neural correlates of reanalysis processes during sentence comprehension. *Human Brain Mapping*, 32, 1775–1787. 10.1002/hbm.21146. [PubMed: 21391256]
- Hothorn T, Bretz F, & Westfall P (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50, 346–363. 10.1002/bimj.200810425. [PubMed: 18481363]
- Iacobucci D, Schneider MJ, Popovich DL, & Bakamitsos GA (2016). Mean centering helps alleviate 'micro' but not 'macro' multicollinearity. *Behavioral Research Methods*, 48, 1308–1317. 10.3758/s13428-015-0624-x.
- Irwin JR, & McClelland GH (2001). Misleading heuristics and moderated multiple regression models. *Journal of Marketing Research*, 38, 100–109. 10.1509/jmkr.38.1.100.18835.
- Jaeger TF (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434–446. 10.1016/j.jml.2007.11.007. [PubMed: 19884961]
- Jessen F, Heun R, Erb M, Granath D-O, Klose U, Papassotiropoulos A, ... Grodd W (2000). The concreteness effect: Evidence for dual coding and context availability. *Brain and Language*, 74, 103–112. 10.1006/brln.2000.2340. [PubMed: 10924219]
- Kegl J (1995). Levels of representation and units of access relevant to agrammatism. *Brain and Language*, 50, 151–200. 10.1006/brln.1995.1044. [PubMed: 7583186]
- Kertesz A (1982). Western aphasia battery. New York: Grune & Stratton.
- Kessler Treiman, & Mullennix (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, 47, 145–171. 10.1006/jmla.2001.2835.
- Kilgarrieff A, Baisa V, Bušta J, Jakubík M, Kovvá V, Michelfeit J, ... Suchomel V (2014). The sketch engine: Ten years on. *Lexicography*, 1, 7–36. 10.1007/s40607-014-0009-9.

- Kiss K (2000). Effect of verb complexity on agrammatic aphasics' sentence production In Bastiaanse R, & Grodzinsky Y (Eds.). *Grammatical disorders in aphasia. A neurolinguistic perspective* (pp. 152–170). London: Whurr Publishers.
- Kuznetsova A, Brockhoff PB, & Bojesen Christensen RH (2016). Tests in linear mixed effects models (Version 2.0–30). <http://cran.uib.no/web/packages/lmerTest>.
- Lee J, & Thompson CK (2011). Real-time production of unergative and unaccusative sentences in normal and agrammatic speakers: An eyetracking study. *Aphasiology*, 25, 813–825. 10.1080/02687038.2010.542563. [PubMed: 22319225]
- Levelt WJ, Roelofs A, & Meyer AS (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38. 10.1017/S0140525X99001776. [PubMed: 11301520]
- Linebarger MC, Schwartz MF, & Saffran EM (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, 13, 361–392. 10.1016/0010-0277(83)90015-X. [PubMed: 6683142]
- Love T, & Oster E (2002). On the categorization of aphasic typologies: The SOAP (A test of syntactic complexity). *Journal of Psycholinguistic Research*, 31, 503–529. 10.1023/A:1021208903394. [PubMed: 12528429]
- Luzzatti C, Raggi R, Zonca G, Pistarini C, Contardi A, & Pinna GD (2002). Verb-noun double dissociation in aphasic lexical impairments: The role of word frequency and imageability. *Brain and Language*, 81, 432–444. 10.1006/brln.2001.2536. [PubMed: 12081411]
- Luzzatti C, Toraldo A, Guasti MT, Ghirardi G, Lorenzi L, & Guarnaschelli C (2001). Comprehension of reversible active and passive sentences in agrammatism. *Aphasiology*, 15, 419–441. 10.1080/02687040143000005.
- Luzzatti C, Willmes K, & De Bleser R (1996). *Aachener Aphasie Test (AAT). Versione italiana*. Florence: Organizzazioni Speciali.
- Mack JE, Meltzer-Asscher A, Barbieri E, & Thompson CK (2013). Neural correlates of processing passive sentences. *Brain Sciences*, 3, 1198–1214. 10.3390/brainsci3031198. [PubMed: 24961525]
- Malyutina S, & den Ouden (2017). Task-dependent neural and behavioral effects of verb argument structure. *Brain and Language*, 168, 57–72. 10.1016/j.bandl.2017.01.006. [PubMed: 28160740]
- McAllister T, Bachrach A, Waters G, Michaud J, & Caplan D (2009). Production and comprehension of unaccusatives in aphasia. *Aphasiology*, 23, 989–1004. 10.1080/02687030802669518.
- Meltzer-Asscher A, Mack J, Barbieri E, & Thompson CK (2015). How the brain processes different dimensions of argument structure complexity: Evidence from fMRI. *Brain and Language*, 142, 65–75. 10.1016/j.bandl.2014.12.005. [PubMed: 25658635]
- Nakagawa S, & Schielzeth H (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142. 10.1111/j.2041-210x.2012.00261.x.
- Osterhout L, & Swinney DA (1993). On the temporal course of gap-filling during comprehension of verbal passives. *Journal of Psycholinguistic Research*, 22, 273–286. 10.1007/BF01067834. [PubMed: 8366478]
- Pickering MJ, & Branigan HP (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39, 633–651. 10.1006/jmla.1998.2592.
- Pollard C, & Sag IA (1994). *Head-driven phrase structure grammar*. Chicago, IL, USA: The University of Chicago Press.
- Porter G, & Howard D (2004). *CAT: Comprehensive aphasia test*. Psychology Press.
- R Core Team (2016). *R: A language and environment for statistical computing* URL: Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rappaport Hovav M, & Levin B (1998). In Butt M, & Geuder W (Eds.). *The projection of arguments: Lexical and compositional factors*. CSLI Publications.
- Rizzi L (1982). *Issues in Italian syntax*, 11 Walter de Gruyter.
- Rizzi L (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17, 501–557.
- Saffran EM, Schwartz MF, Linebarger M, Martin N, & Bochetto P (1988). *Philadelphian comprehension battery*. Unpublished test.

- Sanchez-Alonso S, Martinez-Ferreiro S, & Bastiaanse R (2011). Clitics in Spanish agrammatic aphasia: A study of the production of unaccusative, reflexive and object clitics In Hendrickx I, Lalitha Devi S, Branco A, & Mitkov R (Eds.). *Anaphora processing and applications*. Berlin: Springer Berlin Heidelberg.
- Shapiro LP, & Levine BA (1990). Verb processing during sentence comprehension in aphasia. *Brain and Language*, 38, 21–47. 10.1016/0093-934X(90)90100-U. [PubMed: 2302544]
- Shapiro LP, Zurif EB, & Grimshaw J (1987). Sentence processing and the mental representation of verbs. *Cognition*, 27, 219–246. 10.1016/S0010-0277(87)80010-0. [PubMed: 3691026]
- Shapiro LP, Zurif EB, & Grimshaw J (1989). Verb processing during sentence comprehension: Contextual impenetrability. *Journal of Psycholinguistic Research*, 18, 223–243. 10.1007/BF01067783. [PubMed: 2738859]
- Shetreet E, & Friedmann N (2014). The processing of different syntactic structures: fMRI investigation of the linguistic distinction between wh-movement and verb movement. *Journal of Neurolinguistics*, 27, 1–14. 10.1016/j.jneuroling.2013.06.003.
- Shetreet E, Friedmann N, & Hadar U (2010). Cortical representation of verbs with optional arguments: The theoretical contribution of fMRI. *Human Brain Mapping*, 31, 770–785. 10.1002/hbm.20904. [PubMed: 19890846]
- Shetreet E, Palti D, Friedmann N, & Hadar U (2006). Cortical representation of verb processing in sentence comprehension: Number of complements, sub-categorization, and thematic frames. *Cerebral Cortex*, 17, 1958–1969. 10.1093/cercor/bhl105. [PubMed: 17101687]
- Stadie N, Schroder A, Postler J, Lorenz A, Swoboda-Moll M, Burchert F, ... De Bleser R (2008). Unambiguous generalization effects after treatment of non-canonical sentence production in German agrammatism. *Brain and Language*, 104, 211–229. 10.1016/j.bandl.2007.08.006. [PubMed: 17928044]
- Stromswold K, Caplan D, Alpert N, & Rauch S (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language*, 52, 452–473. 10.1006/brln.1996.0024. [PubMed: 8653390]
- Thompson CK (2003). Unaccusative verb production in agrammatic aphasia: The argument structure complexity hypothesis. *Journal of Neurolinguistics*, 16, 151–167. 10.1016/S0911-6044(02)00014-3. [PubMed: 21274410]
- Thompson CK (2011). *Northwestern assessment of verbs and sentences*. Evanston, IL,.
- Thompson CK, & Shapiro LP (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of underlying forms. *Aphasiology*, 19, 1021–1036. 10.1080/02687030544000227. [PubMed: 17410280]
- Thompson CK, Bonakdarpour B, Fix SC, Blumenfeld HK, Parrish TB, Gitelman DR, ... Mesulam M-M (2007). Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience*, 19, 1753–1767. 10.1162/jocn.2007.19.11.1753. [PubMed: 17958479]
- Thompson CK, Den Ouden D-B, Bonakdarpour B, Garibaldi K, & Parrish TB (2010). Neural plasticity and treatment-induced recovery of sentence processing in agrammatism. *Neuropsychologia*, 48, 3211–3227. 10.1016/j.neuropsychologia.2010.06.036. [PubMed: 20603138]
- Thompson CK, Lange KL, Schneider SL, & Shapiro LP (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology*, 11, 473–490. 10.1080/02687039708248485.
- Thompson CK, Tait ME, Ballard KJ, & Fix SC (1999). Agrammatic aphasic subjects' comprehension of subject and object extracted wh-questions. *Brain and Language*, 67, 169–187. 10.1006/brln.1999.2052. [PubMed: 10210629]
- Tremblay A, & Ransijn J (2015). *LMERConvenienceFunctions: Model selection and post-hoc analysis for (G) LMER models*. R package version, 2 (10).
- Van Valin RD, & LaPolla RJ (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.
- Wise RJ, Howard D, Mummery CJ, Fletcher P, Leff A, Büchel C, ... Scott SK (2000). Noun imageability and the temporal lobes. *Neuropsychologia*, 38, 985–994. 10.1016/S0028-3932(99)00152-9. [PubMed: 10775709]

- Yarbay Duman T, Altinok N, Özgirgin N, & Bastiaanse R (2011). Sentence comprehension in Turkish Broca's aphasia: An integration problem. *Aphasiology*, 25, 908–926. 10.1080/02687038.2010.550629.
- Yokoyama S, Miyamoto T, Riera J, Kim J, Akitsuki Y, Iwata K, ... Kawashima R (2006). Cortical mechanisms involved in the processing of verbs: An fMRI study. *Journal of Cognitive Neuroscience*, 18, 1304–1313. 10.1162/jocn.2006.18.8.1304. [PubMed: 16859416]

Fig. 1.

Example item from the VNT (a), VCT (b), and ASPT (c). In all pictures, the target verb is *dare* (to give). In (a), the participant was instructed to name the action using one word; in (b) the participant had to indicate which one out of the four pictures illustrated *dare* (top right), with distractors bearing the same argument structure (as for *inserire*, to insert, bottom right) or different argument structure (as for *correre*, to run, in the top left corner, or *sbadigliare*, to yawn, in the bottom left corner); finally, in (c), participants were instructed to describe the picture using a sentence that employed all the provided words (i.e. *La donna sta dando la mela al ragazzo*, The woman is giving the apple to the boy).

Fig. 2.

Example item from the SPPT and SCT. The figures depict the action *tirare* (to pull). In the SPPT, participants heard a prime sentence, *la donna è tirata dall'uomo* (the woman is pulled by the man), matching the picture on the left, and were instructed to produce a similar sentence for the picture on the right (i.e. *l'uomo è tirato dalla donna*, the man is pulled by the woman). In the SCT, the sentence *la donna è tirata dall'uomo* (the woman is pulled by the man) was provided and participants had to indicate which picture matched the sentence (i.e. the picture on the left).

Fig. 3.

Response keys used in the VCT. Red, blue, yellow and green indicate respectively pictures on the top left, top right, bottom left and bottom right of the array. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

Fig. 4. Syntactic representation of the subject-relative sentence *Pietro vede il cane che sta inseguendo il gatto*, Pete saw the dog who is chasing the cat, as a) restrictive relative clause and b) pseudo-relative construction.

Table 1

Summary of predictions for accuracy and RTs. Opt = verbs taking optional arguments; Ob = verbs taking only obligatory arguments. Act = active; Psv = passive; S.Rel = subject relative; O.Rel = object relative; S.Cleft = subject cleft; O.Cleft = object cleft.

	number of arguments	argument optionality	canonicity
<i>accuracy</i>	1-arg = 2-arg = 3-arg	opt = ob	Psv = Act O.Rel = S.Rel O.Cleft = S.Cleft
<i>RTs</i>	3-arg > 2-arg > 1-arg	opt > ob	Psv > Act O.Rel = S.Rel O.Cleft > S.Cleft

Table 2

Percentage mean (M) accuracy and standard deviation (SD) by subtest. Mean and SD are provided by verb type for subtests assessing verb argument structure (VAS) processing, and by sentence type and canonical/non-canonical word order for subtests assessing sentence processing.

VAS processing (% accurate responses)						
	VNT		VCT		ASPT	
	M	SD	M	SD	M	SD
Ob1	98	14.1	96	20	100	0
Ob2	93	25.6	100	0	99	10
Op2	97	17.1	97	17	100	0
Ob3	97.5	15.8	100	0	100	0
Op3	91	28.8	98	14	99.5	7
Total	95	21.8	98	14	99.7	5.5

Sentence processing (% accurate responses)					
	SPPT		SCT		
	M	SD	M	SD	
Act	100	0	99	9.7	
Psv	100	0	98.1	13.7	
S.Cleft	100	0	99	9.7	
O.Cleft	99	9.7	81.9	38.7	
S.Rel	95.2	21.4	97.1	16.7	
O.Rel	99	9.8	91.4	28.1	
All Canonical	98.3	12.8	94.5	12.5	
All Non-Canonical	99.5	6.9	90.5	29.4	

Table 3

Results of the logistic regressions conducted on the Verb Naming Test (VNT) and Sentence Comprehension Test (SCT), for accuracy. Analyses for accuracy on other subtests were not conducted because performance was at ceiling. Asterisks indicate statistically significant results: $p < .05$ (*), $p < .01$ (**) and $p < .001$ (***). R-squared values, reflecting the proportion of global variance accounted for by fixed effects, are provided for each regression model. Act = active; Psv = Passive; S.Cleft = subject cleft; O.Cleft = object cleft; S.Rel = subject relative; O.Rel = object relative.

Verb argument structure processing (R-squared = .068)						
subtest	fixed effect	Estimate	Std.Error	z	p-value	sig.
VNT	2 vs. 1 arguments	-0.947	0.784	-1.208	.439	
	3 vs. 1 arguments	-1.327	0.787	-1.688	.202	
	3 vs. 2 arguments	-0.379	0.461	-0.822	.682	
Sentence processing						
subtest	fixed effect	Estimate	Std.Error	z	p-value	sig.
SCT	<i>Model 1: word order (R-squared = .166)</i>					
	Word Order	-1.935	0.629	-3.079	.002	**
	<i>Model 2: sentence type (R-squared = .261)</i>					
	Psv vs. Act	-0.712	1.284	-0.554	.993	
	O.Cleft vs. S.Cleft	-3.259	1.100	-2.963	.032	*
	O.Rel vs. S.Rel	-1.194	0.771	-1.549	.611	
	O.Cleft vs. Psv	-2.548	0.842	-3.026	.027	*
	O.Rel vs. Psv	-1.620	0.873	-1.857	.406	
	O.Rel vs. O.Cleft	0.928	0.562	1.65	.543	
	S.Cleft vs. Act	0.000	1.465	0	1.00	
	S.Rel vs. Act	-1.138	1.217	-0.935	.931	
	S.Rel vs. S.Cleft	-1.138	1.217	-0.935	.931	

Table 4

Mean (M) reaction times (RT) and standard deviation (SD) by subtest. Mean and SD are provided by verb type for subtests assessing verb argument structure (VAS), and by sentence type and canonical/non-canonical word order for subtests assessing sentence processing.

VAS processing (RT)						
	VNT		VCT		ASPT	
	M	SD	M	SD	M	SD
1-arg	1347	441	1355	295	1624	584
2-arg	1497	519	1403	346	1720	620
3-arg	1746	640	1586	335	1743	671
obligatory	1547	567	1413	322	1672	555
optional	1520	554	1501	365	1709	660
Total	1535	560	1455	346	1694	622

Sentence processing (RT)				
	SPPT		SCT	
	M	SD	M	SD
Active	3249	480	2329	496
Passive	3254	497	2394	561
Subject Cleft	3126	410	2398	505
Object Cleft	3388	551	2677	547
Subject Relative	4019	520	3110	608
Object Relative	4299	456	3373	662
All Canonical	3451	609	2611	643
All Non-canonical	3641	683	2806	725

Table 5

Results of the regression analyses conducted on RTs obtained from participants on the VNT and VCT. Analyses conducted on the ASPT are not reported, as all fixed effects were eliminated by the backward automated procedure. Asterisks indicate statistically significant comparisons: $p < .05$ (*), $p < .01$ (**), $p < .001$ (***). Marginally significant results ($p < .1$) are also listed (~). R-squared values reflecting the proportion of variance accounted for by fixed effects are provided for each regression model.

Fixed Effect	Estimate	Std.Error	z	p-value	sig.
Verb Naming Test (R-squared = .167)					
2 vs. 1 arguments	0.057	0.028	2.027	.098	~
3 vs. 1 arguments	0.112	0.040	2.809	.013	*
3 vs. 2 arguments	0.056	0.027	2.033	.097	~
Arg. Optionality	-0.066	0.023	-2.837	.012	*
Imageability	-0.110	0.038	-2.896	.011	*
Frequency	-0.002	0.005	-0.413	.685	
Imageability*Frequency	0.039	0.016	2.473	.026	*
Verb Comprehension Test (R-squared = .075)					
2 vs. 1 arguments	0.013	0.030	0.444	.896	
3 vs. 1 arguments	0.070	0.032	2.175	.074	~
3 vs. 2 arguments	0.057	0.025	2.236	.064	~
Argument Structure Production Test (R-squared = .03)					
2 vs 1 arguments	0.084	0.031	2.708	.011	*
3 vs. 1 arguments	0.152	0.064	2.364	.029	*
3 vs. 2 arguments	0.068	0.036	1.897	.087	~
Sentence Length	-0.015	0.009	-1.643	.104	

Table 6

Results of the regression analysis conducted on RTs obtained from participants on the SPPT and SCT. Asterisks indicate statistically significant comparisons: $p < .05$ (*), $p < .01$ (**), and $p < .001$ (***). R-squared values reflecting the proportion of variance accounted for by fixed effects are provided for each regression model. Sentence types are abbreviated as follows: Act = active; Psv = Passive; S.Cleft = subject cleft; O.Cleft = object cleft; S.Rel = subject relative; O.Rel = object relative.

Sentence Production Priming Test					
Fixed Effect	Estimate	Std.Error	t	p-value	sig.
<i>Model 1: word order (R-squared = .313)</i>					
Sentence Length	0.022	0.002	9.284	<.001	***
<i>Model 2: sentence type (R-squared = .396)</i>					
Psv vs. Act	-0.006	0.007	-0.792	.961	
O.Cleft vs. S.Cleft	0.034	0.007	4.644	<.001	***
O.Rel vs. S.Rel	0.029	0.007	3.895	.001	**
O.Cleft vs. Psv	0.000	0.008	-0.022	1.00	
O.Rel vs. Psv	0.070	0.013	5.239	<.001	***
O.Rel vs. O.Cleft	0.071	0.010	6.886	<.001	***
S.Cleft vs. Act	-0.040	0.009	-4.511	<.001	***
S.Rel vs. Act	0.036	0.015	2.421	.127	
S.Rel vs. S.Cleft	0.076	0.011	7.128	<.001	***
Sentence Length	0.011	0.003	4.516	<.001	***
Sentence Comprehension Test					
Fixed Effect	Estimate	Std.Error	t	p-value	sig.
<i>Model 1: word order (R-squared = .34)</i>					
Word Order	0.023	0.009	2.466	.020	*
Sentence Length	0.029	0.002	13.156	<.001	***
<i>Model 2: sentence type (R-squared = .365)</i>					
Psv vs. Act	0.000	0.012	0.031	1.00	
O.Cleft vs. S.Cleft	0.041	0.012	3.321	.010	**
O.Rel vs. S.Rel	0.037	0.012	3.098	.020	*
O.Cleft vs. Psv	0.019	0.014	1.407	.683	
O.Rel vs. Psv	0.075	0.022	3.406	.007	**
O.Rel vs. O.Cleft	0.056	0.017	3.256	.012	*
S.Cleft vs. Act	-0.021	0.014	-1.446	.657	
S.Rel vs. Act	0.039	0.024	1.606	.550	
S.Rel vs. S.Cleft	0.060	0.017	3.425	.007	**
Sentence Length	0.018	0.004	4.270	<.001	***