Factives at hand: When presupposition mode affects motor response

Robert Reinecke[1,2], Tatjana A. Nazir[1,3], Sarah Carvallo[2,4], and Jacques Jayez[1,2]

[1]Institut des Sciences Cognitives - Marc Jeannerod, CNRS/UCBL

[2]Ecole Normale Supérieure de Lyon

[3]Université Claude Bernard Lyon 1

[4]Ecole Centrale de Lyon

Author Note

This work was supported by the French National Agency for Research (CogHuLiCe ANR 16-CE28-0004) attributed to TN and SC, by a DAAD scholarship to RR, and by the French National Center for Scientific Research (CNRS).

Corresponding author: Robert Reinecke – rreineck@isc.cnrs.fr

Institut des Sciences Cognitives - Marc Jeannerod, UMR 5304, CNRS/UCBL, Université Claude Bernard Lyon1, 67 Bd. Pinel, Bron, France

Tel : +33 758037189

ABSTRACT

It is well-established that the processing of hand, mouth, and foot-related action terms can activate areas of the motor cortex that are involved in the planning and execution of the described actions. In the present study, the sensitivity of these motor-structures to language processes is exploited to test linguistic theories on information-layering. Human languages possess a variety of linguistic devices, so-called presupposition triggers, which allow us to convey background information without asserting it. A statement such as "*Marie stopped smoking*" presupposes, without asserting it, that Marie used to smoke. How such presupposed information is represented in the brain is not yet understood. Using a grip force sensor that allows capturing motor brain activity during language processing, we investigate effects of information-layering by comparing asserted information (*In the living room, Peter* <u>*irons*</u> *his shirt*) with information embedded under a presuppositional factive verb construction (*Louis* ***knows*** *that Peter* <u>*irons*</u> *his shirt*; Experiment 1) and a non-factive verb construction (*Louis* ***believes*** *that Peter irons his shirt;* Experiment 2). Furthermore, we examine whether the projection behavior of a factive verb construction modulates grip force under negation (*Louis* ***does not know*** *that Peter* <u>*irons*</u> *his shirt;* Experiment 3). The data show that only the Presupposed Action verb in affirmative contexts (Experiment 1) triggers an increase in grip force comparable to the one of Asserted Action verbs, whereas the non-factive complement shows a weaker response (Experiment 2) and an even weaker response is observed for projective action verbs (Experiment 3). While the first two experiments seem to confirm the sensitivity of the grip force response to the construction of a plausible *event model*, in which the motor action is represented as taking place, the third one raises the question of how robust this hypothesis is and how it can take the specificity of projection into account.

Keywords: language processing, presupposition, negation, language-induced motor activity

1. INTRODUCTION

Human languages possess a variety of linguistic devices, so-called *presupposition triggers*, which allow us to convey background information without asserting it. Among these, we find factive verbs, like the verb *to know*. When someone says *Paul knows that Mary writes a letter*, the factive verb *know* presupposes the truth of the complement clause that Mary writes a letter and asserts that Paul is certain that Mary writes a letter (Egre, 2008; Kiparsky & Kiparsky, 1970). The former corresponds to the background information, the latter to the foreground or assertive information. Presupposition is, thus, information which is old, previous, or given, or at least presented as such (Stalnaker, 1974). After having been extensively scrutinized from a theoretical perspective (e.g. Beaver, 2001; Gazdar, 1979; Geurts, 1999; Heim, 1983; Karttunen, 1974; Stalnaker, 1974; Schlenker, 2008), presuppositions have been more recently investigated experimentally. The current experimental literature on presuppositions provides insights into the time course of their interpretation as well as into the cognitive costs associated to presupposition processing (e.g. Domaneschi, 2016; Schwarz, 2015).

This paper aims at deepening the understanding of the cognitive underpinnings of presupposition processing by focusing on its sensori-motor correlates. The relationship between language processing and motor activation has received great attention within the field of cognitive neuroscience (see for a review Pulvermüller, 2005; Willems & Casasanto, 2011; Kiefer & Pulvermüller, 2012) but its implications for linguistic theories have just started to be explored. In this paper, we address the question of whether presuppositional contexts modulate the sensori-motor activation elicited by action verbs. While hand–related action verbs in simple affirmative assertive sentences trigger a response in sensori-motor structures of the brain (e.g. Aziz-Zadeh, Wilson, Rizzolatti, & Iacobini, 2006; Hauk, Johnsrude & Pulvermüller, 2004; Tettamanti et al., 2005), it has been shown that it is not always the case in other linguistic environments, such as negation or volitional verbs (*want, desire*), (e.g. Aravena et al., 2012, 2014; Zwaan, Taylor, & de Boer, 2010; Papeo, Hochmann, & Batelli, 2016; Tettamanti et al., 2008). The question then naturally arises whether hand-related action verbs trigger a similar response when they are part of the presupposition. Answering this question is interesting from two perspectives. First, it would augment our knowledge of the array of contexts in which a sensori-motor response can be evoked. Second, it would contribute to a better understanding of the cognitive status of presuppositional information. In particular, if, as the descriptive and theoretical literature suggests, presuppositions are not the main piece of information in the linguistic message, it is possible that this secondary or peripheral status is reflected in a difference of impact on the sensori-motor system.

In order to convey a more concrete sense of what is at stake, we proceed as follows in this introduction: First, we present a short overview of recent studies indicating the context-sensitivity of sensori-motor activation during language processing. Second, we introduce the linguistic phenomenon of presuppositions and discuss the properties of factive verbs. Finally, we give an overview of the experimental studies presented in the paper.

1.1. Variations in sensori-motor cortex activation during language processing

A large body of evidence shows that sensori-motor cortices are recruited during the processing of action-related language. Early studies highlighted that arm-, mouth- and leg-related words can activate areas of the motor cortex that are involved in the planning and execution of the described actions (e.g. Aziz-Zadeh et al., 2006; Hauk et al., 2004; Tettamanti et al., 2005) in a rapid and automatic manner (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005). Furthermore, they revealed that the processing of action verbs can also occur when an action content is not explicitly attended (Gallese & Lakoff, 2005; Pulvermüller et al., 2005). However, these early findings on such a *motor resonance* have more recently been challenged by a series of studies which have questioned the automaticity of word-related sensori-motor activation. More precisely, it has been shown that modality-specific brain activity during action word processing is context-sensitive. For instance, Moody and Gennari (2010) show that this activation is modulated as a function of the degree of effort that is implied by the relevant action. The authors found that premotor cortex activation was strongest in a high effort condition (*The athlete is throwing the javelin*), middle in a low effort condition (*The athlete is throwing the frisbee*) and lowest in a no effort condition (*The athlete noticed the frisbee*). So, the description of distinct actions can differentially activate the same brain region in accordance with the effort which is typically associated with the performance of each action.

Crucially, though, language-related sensori-motor activation is not only modulated by extra-linguistic context but appears also to be affected by the *linguistic* context which embeds the relevant action word. One of the first evidences that the excitability of motor brain structures depends on the linguistic context comes from studies which focus on the distinction between literal and non-literal uses of language. Aziz-Zadeh et al.'s (2006) results indicate that literal hand-, foot-, and mouth-related action verbs activate similar motor brain structures as observations of these respective actions, whereas their non-literal counterpart, in expressions such as *chewing over the details*, *grasping the idea*, and *kicking off the year*, did not elicit the same response. In accordance with this finding, Raposo and colleagues (2009) also highlight the context sensitivity of action verbs, that is, the fact that context is a crucial determiner of how an action verb is processed. Their data indicate that isolated action verbs (e.g. *grab*) activate motor regions to a higher degree than action verbs in literal sentential contexts (e.g. *The fruit cake was the last one so Claire grabbed it*). In addition, the motor and premotor cortices are not activated when action verbs are presented in an idiomatic context (e.g. *The job offer was a great chance so Claire grabbed it*). Much evidence confirms the substantial difference of activation between the literal and idiomatic use of action verbs (Cacciari et al., 2011; Desai, Conant, Binder, Park, & Seidenberg, 2013; Lauro, Mattavelli, Papagno, & Tettamanti, 2013; but see Boulenger, Hauk, & Pulvermüller, 2008; Boulenger, Shtyrov, & Pulvermüller, 2012 – for a detailed interpretation of their different findings see Willems & Casasanto, 2011). Lauro et al. (2013) examine the difference between the literal use of action verbs and three different figurative meanings, i.e. metaphors, fictive-motion, and idioms. Their results point out that literal and idiomatic uses of action verbs appear to be endpoints of a motor brain activation continuum, that is, literal action verbs activate premotor brain areas, whereas action verbs that are part of an idiom do not. Metaphors like *Paul throws his sadness away* range between these two poles of the continuum. This intermediate status is typically explained by arguing that even if metaphors depict action simulations that are impossible to perform, the comprehension process relies on past body related experiences in order to correctly infer the metaphorical meaning (for a theoretical view see Gibbs, Lima, & Francozo, 2004; Gibbs, 2006).

The linguistic modulation of sensori-motor activity is not limited to the literal/non-literal distinction as other linguistic factors have also been shown to be critical. Firstly, the same action word embedded in a negated sentence (*I do not push the button*) does not activate the brains' motor structures in the same way as in affirmative sentences (e.g. Aravena et al., 2012; Tettamanti et al., 2008). Furthermore, no language-induced motor activity is present when action words are embedded within a volitional context (*Fiona wants* *to sign the contract*) (Zwaan et al. 2010; Aravena et al., 2014). However, motor structures can be activated by a gapped verb, that is when the context sets up an expectation of an upcoming motor-related action such as in "John closes a juice bottle and Jim [ ] a lemonade bottle" (Claus, 2015). Altogether, these studies highlight that contextual manipulations of lexical properties – interpretation of metaphors and idioms, the presence of a negation operator, a volitional or a gapped verb – have an impact on the involved brain structures.

Contextual manipulations that involve discourse properties also have an impact on the elicited sensori-motor activation. For instance, van Ackeren and colleagues (2012) showed that a sentence such as *It is very hot here* can be processed in different ways. In a context where this utterance can be interpreted as an indirect request of action – seeing a picture of a window (which triggers the indirect request *open the window*) – cortical motor areas are activated, whereas this is not the case when the utterance is not interpreted as an indirect request – seeing a picture of a desert (for related findings see also van Ackeren, Smaragdi, & Rueschemeyer, 2016; Egorova, Pulvermüller, & Shtyrov, 2014; Egorova, Shtyrov, & Pulvermüller, 2016). These findings provide the first evidence that sensori-motor activation is elicited even when the information is conveyed implicitly and must be inferentially derived in order to understand what is meant (an implicature in the sense of Grice, 1975). Van Ackeren et al.'s (2012) results highlight that discourse properties – such as the layering of information realized by the distinction between what is literally said and what is implicated – also drive sensori-motor activation. Thus, it appears worth extending these findings and investigating how further interactions of language and context play an active role in modulating the sensori-motor activation elicited by action verbs. We now turn to the phenomenon of linguistic presupposition, which is assumed information beyond what is said and what is implicated.

1.2. **Presuppositions**

Linguistic presupposition is a type of information which is triggered by the presence of certain linguistic expressions (*presupposition triggers*) and is conveyed in discourse as part of the background of the conversation. A variety of distinct linguistic forms such as definite descriptions as in (1), change of state verbs as in (2), iterative adverbs as in (3), wh-question as in (4) and constructions like temporal clauses as in (5) trigger presuppositions (for an extensive list see Levinson, 1983).

1 I have to pick up my sister at the airport.

a. I have a sister. (Presupposition)

2 Peter stopped smoking.

a. Peter used to smoke. (Presupposition)

3 Barack Obama was elected again.

a. Barack Obama was elected before. (Presupposition)

4 When did Michael leave the house?

a. Michael left the house. (Presupposition)

5 Before Strawson was even born, Frege noticed presuppositions.

a. Somebody named Strawson was born. (Presupposition)

As can be seen from these examples, presupposition triggers are often used without even noticing it and are ubiquitous in discourse. In contrast to implicatures, which heavily depend on inference, the presuppositional layering is *coded* as an intrinsic property of constructions. The presupposition triggers of interest in our study are *factive* verbs, which presuppose that their complement clause expresses a *true* proposition (Egre, 2008; Kiparsky & Kiparsky, 1970). For instance, in (6) the factive verb *know* presupposes that Mary writes a letter (6b) and asserts that Paul is certain that Mary writes a letter (6a). It is important to note that the conveyed presupposed information is considered as *true information* which the speaker is committed to (Peters, 2016). Similarly, several authors, including for instance Geurts (1999) and Reboul (2017) note that presuppositions are not cancellable or defeasible, hence the oddness of a sentence like in (7), where the presupposition is negated. Presuppositions are typically part of the common ground, that is, the set of beliefs shared by the participants. In this line, presuppositions are not considered as the questions under discussion – not *at issue* in the current linguistic terminology – since they are taken for granted and (presented as) admitted by the participants (Ducrot, 1972; Stalnaker, 1974). Along these lines, presuppositions are considered to condition the appropriateness of an utterance. Using the sentence in (6) in which the presupposition of (6b) is not yet part of the common ground would lead to presupposition failure (Stalnaker, 2002). Presupposition failure may result into accommodation, that is, the process by which the hearer accepts the presuppositional content as true and includes it into her set of beliefs (Heim, 1983; Lewis, 1979).

6 Paul knows that Mary writes a letter.

(6a) Paul is certain that Mary writes a letter. (Assertion)

(6b) Mary writes a letter. (Presupposition)

7 Mary didn't write a letter and Paul knows that she wrote a letter.

Recent experimental investigations comparing contextually satisfied versus accommodated presuppositions indicate that accommodated presuppositions are integrated rapidly (for a review, see Schwarz, 2015). However, in the accommodation condition, higher processing costs are involved with respect to the triggering point, that is, at the position of the presupposition trigger and with the word that immediately follows the presupposition trigger (EEG study by Domaneschi, Canal, Masia, Vallauri, & Bambini, 2018; self-paced reading study by Tiemann et al., 2011; Domaneschi & Di Paola, 2017; eye-tracking study by Tiemann & Schwarz, 2012). More precisely, in a sentence, for instance, like *Peter stopped smoking* the verb *smoking* elicited a biphasic N400/P600 in the accommodation condition that is no knowledge of smoking has been provided a priori in comparison to the satisfaction condition, where prior information of smoking has been provided (Domaneschi et al., 2018). Nevertheless, these higher early sentential processing costs fade away towards the end of the sentence and do not have an impact on the accuracy ratings and on the response time to a question regarding the accommodated presupposed content (see for instance the behavourial results of Domaneschi et al., 2018; Domaneschi & Di Paola, 2019). Further evidence that accommodated presuppositions come with an early extra processing cost but are integrated rapidly is provided by Masia and colleagues (2017). The authors directly compare event-related potentials recorded during the processing of assertive and presupposed content, in which indefinite descriptions like *a migration* vs definite descriptions like *the migration* were used. For the latter, the presupposition – a specific migration – had to be accommodated and was not satisfied a priori. The authors observed a larger N400 for the presupposition condition, which is compatible with the hypothesis of an extra processing cost during processing the accommodated presupposition. Such a result fits well in the nature of the N400 component, in which the N400 is usually linked to semantic or thematic relations (for a review, see Kutas & Federmeier, 2011). In the case of a definite, the hearer does not a have a rich expectation about the upcoming word given its prior discourse absence, consequently an N400 is elicited. However, this effect, at least for definite description, is only transitory since no associated P600 effect was observed by the authors. The P600 usually reflects discourse updating, that is the resolution of a prior incongruence (Friederici, 2011). Masia et al.'s finding shows that in the case of definite descriptions the information of the presupposition does not generate an incongruence that is the absence of a P600. Such an immediate integration "may be most naturally compatible with accounts that all assume presupposed content [to] be encoded conventionally" (Schwarz, 2016; p. 286). Further evidence regarding the immediate availability of the presupposed content comes from eye-tracking studies using the visual world paradigm. These studies show that, after the onset of the presupposition trigger, fixations shift immediately to the picture containing both asserted and presupposed information, that is, depicting the presupposed and asserted content in one picture (see Romoli et al., 2015; Schwarz, 2014 for *also* and Schwarz, 2014 for *stop*). In conclusion, processing presuppositions comes with an additional cost when the presupposition has not yet been part of the common ground; however, such a cost, if present, only appears to be detectable on-line and fades away rapidly.

1.3. Presupposition Projection

Further evidence of the difference between asserted and presupposed content comes from linguistic tests and from criteria such as the so-called *projection* property of presuppositions: When an operator that suspends or shifts the truth value is applied to a sentence containing a presupposition trigger, it affects the asserted piece of information but, in general, not the presupposition. For instance, the negation of the factive verb in (8), denies that Paul is certain that X, that is, it alters the meaning of the asserted content in comparison to the sentence in (6). However, the presupposition remains untouched, that is, the negated sentence still presupposes that *Mary writes a letter*, exactly as the positive sentence in (6) does. This and similar observations on questions as in (9) and on modal verbs as in (10) correspond to what linguists have called *presupposition projection* (see for instance Chierchia & McConnell-Ginet, 1990; Ducrot, 1972; Geurts, 1999; Heim, 1983).

8 Paul does not know that Mary writes a letter.

9 Does Paul know that Mary writes a letter?

10 Paul might know that Mary writes a letter.

After extensive scrutiny in the descriptive and theoretical literature, the projection phenomenon has also received some experimental examination. For instance, Chemla and Bott (2013) investigated the strength of the projection effect in factive verb constructions and found evidence that the projective interpretation was derived faster than the non-projective interpretation. Similarly, using a picture selection task, Romoli and Schwarz (2015) show that the projective interpretation is preferred and chosen significantly faster. Such results suggest that presuppositions have a behavioral correlate when processing the presupposed content of a negated assertion.

In contrast to factive verbs, non-factive verb constructions as in (11) impose no constraint on the truth-value of the embedded *that*-clause (for an interesting overview on factive and non-factive mental states see Nagel, 2017). As noted earlier, the sentences in (12) and (14) entail the truth of the complement, whereas its truthfulness in (11) and (13) may depend, among other factors, on the reliability of Paul (Nagel, 2017).

11 Paul thinks that Mary writes a letter.

12 Paul knows that Mary writes a letter.

13 Paul does not think that Mary writes a letter.

14 Paul does not know that Mary writes a letter.

1.4. The Present Research

Our study is aimed at (i) investigating the cognitive correlates of presupposition processing and, at the same time, (ii) extending our current understanding of which linguistic contexts modulate motor brain structures. The phenomenon of presupposition is special in three respects.

(1) As mentioned above, motor resonance during the processing of action verbs is not observed within sentential environments that involve negation (Aravena et al., 2012; Tettamanti et al., 2008; Papeo

et al., 2016) or volitional verbs (Aravena et al., 2014; Zwaan, et al., 2010). A straightforward interpretation of these observations is that the discourse or situation model (cf. Zwaan and Radvansky, 1998) constructed by listeners in such cases does not include the *event*, which, consequently, does not generate a motor response (see Aravena et al., 2014). Sensori-motor activation is triggered when the action of the corresponding verbal group actually takes place. Taylor and Zwaan (2008) called this the *Linguistic Focus Hypothesis*. In this line, saying that *Paul does not throw the ball*, for instance, leaves no room for an event of throwing a ball in the situation model. If presuppositions are considered as true by default, they are part of the depicted situation and should therefore trigger a motor response whenever a motor action is involved (e.g. the word *writes* in a factive verb construction as in (12)). Contrariwise, if presuppositions are peripheral information, one could also expect that they will not elicit the same response as assertions. Our study will shed light on how presuppositions are processed in the motor brain structure.

(2) Most of the currently available observations and experiments on presuppositions concern linguistic operators or contexts (e.g. for aspectual verbs and definite descriptions Domaneschi et al., 2018; for *wieder (again)* Schwarz & Tiemann, 2012, for processing factives versus non-factives Shetreet et al., 2019). Operators like negation or interrogation reveal projection properties, while contexts are relevant to the discourse-based properties of presupposition triggers, like satisfaction or accommodation. In contrast, we have only scarce information about the *intrinsic* properties of triggers. Does a presupposition in a simple decontextualized assertive sentence have cognitive properties that distinguish it from an assertion or are such properties visible only in richer environments (embedding operators, or the presence of an explicit context)? Given that a large part of the literature on motor resonance focuses on isolated words or simple sentences, it is necessary to design experiments that allow testing these intrinsic properties of presupposition triggers.

(3) Studying presupposition triggers raises the question of which trigger(s) to use in experiments. Factive verb constructions are a particularly interesting starting point from different aspects. First, with factives, the assertive and presupposed contents are explicitly expressed at the sentential level,

which is a unique characteristic when compared to other presupposition triggers such as so-called *aspectual* verbs[1]. This explicitness allows one to directly compare the presupposed content of *Paul knows that Mary writes a letter* to an assertion such as *Mary writes a letter* without making an effort to infer the implicit presupposed content as with other presupposition triggers. Second, it was shown that, in otherwise totally parallel clausal complement constructions, factive verbs (*know*-type) presuppose the truth of its complement clause, whereas non-factive verbs (*believe*-type) do not commit one to the veracity of their complement clause. Therefore, it makes sense to investigate whether this distinction also has a cognitive motor resonance counterpart.

On these grounds, we present three experiments, in which a grip force sensor is used to monitor variations of grip force between thumb and index finger after the onset of a critical word (a hand-related action verb in our case). A word-induced increase of grip-force can be interpreted as an incomplete inhibition that arises from primary motor cortex activity (Aravena et al., 2012; 2014; Jeannerod, 1994). Previous research has shown that the grip force plays an essential role with respect to the predictive and reactive control of the capacity to hold and lift objects (for a review see Delevoye-Turrell & Wing, 2005). Healthy adults, whose anticipatory predictive and reactive control is intact, easily adjust grip force to the mass and texture of an object (Johansson, Westling, & Roles, 1984). Crucially, previous studies have demonstrated the link between grip force and motor brain activity: the primary, premotor, supplementary and cingulate cortical motor areas play a crucial role when information is sent via spinal motor neurons to the finger muscles (e.g., Dum & Strick, 1991; Lemon, 1993). Moreover, recent neurophysiological evidence using the fMRI technique showed that when gently holding an object, grip forces activate the left primary sensorimotor cortex, the ventral premotor cortex and the left posterior parietal cortex (Kuhtz-Buschbeck, Ehrsson, & Forssberg, 2001). By investigating grip force in a healthy adult population, it has been demonstrated that subtle grip force variations have also been observed during language processing depending on the action status. When comparing action verbs and Non-Action related nouns, Frak et al.

[1] The term *aspectual* in this context denotes *change of state* or *transition* verbs like *begin*, *stop*, *resume*, *interrupt*, *continue*, etc. Their presupposed content is inferred by default. The sentence *Mary stopped smoking* asserts explicitly that *Mary does not smoke* and presupposes that *Mary used to smoke*.

(2010) found that grip force variation increased for the former but not for the latter. This finding was extended to verbs by Aravena et al. (2012, 2014) and Nazir et al. (2017), with a direct comparison of action and Non-Action verbs.

Grip-force variation is typically measured while participants actively listen to auditory stimuli. Using this tool, the following three research questions will be addressed:

(RQ1): Does the presupposed (action-related) content of factive verb constructions elicit an increase in grip force?

> The first experiment addresses the issue whether a piece of information which is (i) true but (ii) syntactically marked as backgrounded activates motor brain structures. As noted above, the hybrid status of presuppositions (they are true but secondary) fosters doubt about which prediction is a priori the most plausible. Presuppositions are part of the situation model; however, their marginal status compared to an assertion may impact the processing in motor brain structures.

(RQ2): Does the entailed (action-related) content of non-factive verb constructions elicit an increase in grip force?

> The second experiment directly compares the action-related content of an assertion to the non-factive complement. The a priori plausible prediction depends on the results of the first experiment. If the presupposition of factive verbs has a sensori-motor impact in virtue of being true, it is expected that this impact is weaker or absent with non-factive verbs, since the truth of the embedded clause is not guaranteed.

(RQ3): Does the presupposed (action-related) content of negative factive verb constructions elicit an increase in grip force?

> The third experiment addresses the projection phenomenon of factive verb constructions. Again, if Experiment 1 provides evidence for some sensori-motor activation due to the truth of the presupposition, it is expected that a projective reading (i.e. where the presupposition is not negated)

is preferred in projective environments; thus, giving rise to an increase of motor activation comparable to that observed in Experiment 1.

2. METHOD

In the following we will describe the general method that applies to all experiments. Further details are provided in the respective method sections of the individual experiments.

The present study was approved by the Ethical Committee CPP (Comité de Protection des Personnes) Sud-Est II in Lyon, France.

Participants

Participants were French undergraduate students and native speakers of French. They had normal or corrected to normal vision, no history of neurological or psychiatric disorders, and were right-handed as attested by the Edinburgh Handedness questionnaire. All participants gave an informed written consent and were informed that they could end the experiment at any moment. They were paid for their participation.

Stimuli

All stimuli sentences contained hand-related action verbs involving grip actions, except the control sentences, which contained Non-Action verbs. Sentence specific characteristics are detailed in the method sections of the individual experiments. All stimuli are provided in the Supplementary Material.

Measures and pre-tests

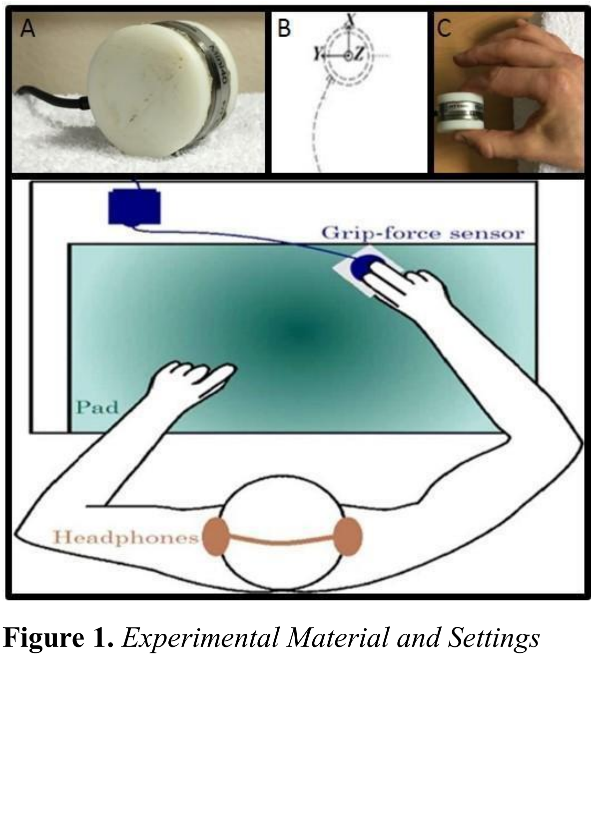
The hand-related action verbs were selected in two steps. First, 20 participants rated a list of 120 hand-related action verbs as to the likelihood that the verb refers to a manual action using a 5-point Likert scale: (1) never, (2) rarely, (3) sometimes, (4) often, and (5) always. Second, for the 66 verbs that received a rating of at least (4), another group of 58 participants was requested to complete a list of sentences containing the selected verbs (e.g *Ines ties* _______). We randomly divided the 66 verbs into two lists, each containing 33 verbs. Participants were randomly assigned to one of the two lists. Sentences which met the following criteria were included in the study: (1) Sentence completion was related to a manual activity and (2) the mean cloze probability for the chosen continuation was at least 25%. A final list of 37 sentences served as stimuli for the grip force study.

Recording

The stimuli for all the experiments were recorded in a sound booth by the same female speaker with a Roland Edirol R-09, at a 48KHz sampling rate with 24-bit digitalization. Special care was taken that the speaker maintain a relatively flat prosody and avoid any loudness or pitch variation on the critical words (i.e. the verb and the noun).

Equipment and data acquisition

Two distinct computers were used for data recording and stimulus presentation, in order to ensure synchronization between audio files and grip-force measurements (estimated error < 5 milliseconds). The first computer read the play-list of the pseudo-randomized stimuli. The second computer recorded the incoming force signals from the load cell at a high sampling rate of 1000 Hz. To measure the activity of the hand muscles, a 6-axis load cell of 68 g was used (ATI Industrial Automation, USA, see Figure 1). Like in previous studies (e.g. Frak et al., 2010; Nazir et al., 2017), only the three main forces were recorded: the longitudinal (Fx,), radial (Fy) and compression forces (Fz), respectively (Figure 1B).



(A) A standalone 6-axis load cell of 68 g was used (ATI Industrial Automation, USA). **(B)** The directions of the recorded forces: longitudinal (Fx), radial (Fy), and compression (Fz). **(C)** Participants held the grip-force sensor with their right thumb and index. Their wrist was placed on a 15 cm high box. **Bottom panel:** Participants wore headphones and were comfortably seated behind a desk on which a pad was placed. They were asked to rest their arms on the pad when holding the sensor.

Figure 1. *Experimental Material and Settings*

Procedure

Participants wore headphones and were comfortably seated behind a desk on which a 15 cm high box was placed. They were asked to rest their right wrist on the box. Their hand was detached from the box, that is, it was free-standing and not in contact with the table when participants held the grip-force sensor with their right hand (see Figure 1C). The experimenter demonstrated the correct way to hold the grip force sensor and participants were requested to hold the cell with a constant force, measured as 1.5 Newton (N). The thumb and index finger remained on the load cell during each block.

The experiment started with a training session of two blocks (in total, 21 stimuli), in which instructions about the experiment were given. In this session, the participants got familiarized with the task and had the opportunity to ask any question they found relevant. When they felt ready, the experiment started. Participants had to listen to 111 stimuli, distributed into 10 blocks, 9 blocks of 11 stimuli and a final block of 12. In order to avoid muscular fatigue, a 30-second pause occurred between two consecutive blocks, but the participants could ask for more if they judged that they needed more time to relax. At the beginning of each block, they had to control their initial grip force and adjust it to 1.5 N, using the screen to monitor their performance. The experimenter informed the participant and started the auditory presentation as soon as the mentioned grip force level was met and no fluctuations occurred. Participants kept their eyes closed for the duration of each block. At the end of each block, they put down the cell and a question with respect to the Action/Non-Action related verb appeared on the screen in front of them, which had to be answered by using the left ("*yes*") or right ("*no*") button of the mouse. The total length of each experiment was approximately 25 min. Participants were debriefed at the end of the experiment.

Data Analysis

Data processing and visualization were carried out in *R* (R Core Team, 2019) using a number of specialized libraries, most notably *stats* (R Core Team, 2019), *lme4* (Bates, Maechler, Bolker & Walker, 2015), *forcats* (Wickham, 2019) and *ggplot2* (Wickham, 2016).

Prior to data analysis, each signal component was pre-treated in order to eliminate the electro-magnetic oscillations of the cell. We used the function *loess* implemented in basic R. This function derives its name from the contraction of *LOcally weighted Scatterplot Smoother*. It replaces every point of a scatter plot by the average of a weighted linear regression calculated on neighbor points. More precisely, starting from any value y at position x, the algorithm (i) selects points in an interval provided by the user, (ii) calculates a regression line over these points, giving more weight to the points closer to x, (iii) returns the value of the regression equation at x. The oscillating waveform is thus replaced by a smooth curve. A visual trial-and-error procedure led to an interval of 0.15, meaning that 15% of all the points were kept around each value to estimate the regression line. The result of this preprocessing is illustrated in Figure 2, where the initial series of peaks is replaced by the white curve.

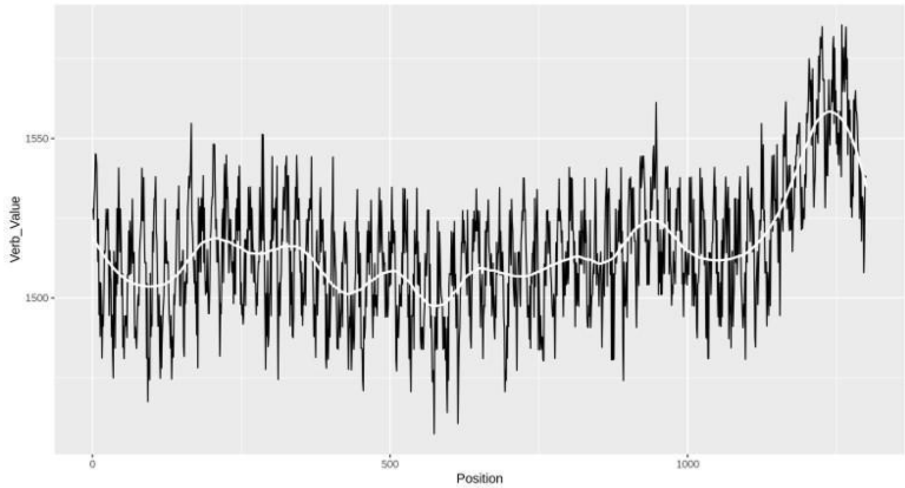


Figure 2. Replacing the oscillating waveform with the loess function in R

Finally, a baseline correction was performed from -300 to 0 ms prior to target onset. This correction was implemented because of a possible global change in grip-force during the session (~25 min per participant), and because we were only interested in grip-force changes. Thus, we adjusted the post-stimulus values by subtracting the baseline values from all of the values in the epoch. Given that the participants were asked to hold the grip-force sensor throughout the experiment, a "negative" grip-force refers to a lesser grip-force and not to the absence of grip-force, which is impossible in this context (the cell would just fall).

Only *Fz* (compression force) was included in the analysis because this parameter was determined to be the most accurate indicator of prehensile grip-force (e.g. Frak et al., 2010). Since the expressions describing hand-related actions using a verb *and* a noun phrase, as in *tie* (verb) *her shoes* (noun phrase), we analyzed possible effects not only after the onset of the verb (either a hand-related action verb or a Non-Action verb), but also after the onset of the noun. Using the *Praat* software (Boersma & Weenink, 2019), the *Fz* signals were segmented off line. The temporal distance between the verb and the noun phrase varies across stimuli.

A preliminary visual investigation had shown that, on average, the noticeable effects occur in the 300-1000 region after verb onset (for very similar time windows, see also Aravena et al., 2012, 2014). However, in order to detect possible more fine-grained variations at later stages, we included a larger region, extending until 1000 ms after the 'latest' noun, that is the noun with the largest distance from the onset of its verb. This resulted into a 2305 ms time span for the first and second experiments (maximum interval between noun onset and verb onset = 1005 ms), and a 1958 ms time span for the third experiment (maximum interval between noun onset and verb onset = 658 ms). We also studied an even later temporal window, a point to which we return when presenting the third experiment in Section 3.3.

Before starting the statistical analysis, we inspected the average time-Fz plots for each participant in order to detect *negative drifts*, that is global and systematic decreasing curve slopes during the first 1000 ms after verb onset. This might indicate that the participant did not hold the cell with sufficient pressure, due to inability, stress or misunderstanding. An example of negative drift is given in Figure 3 (participant 16 in experiment 3).

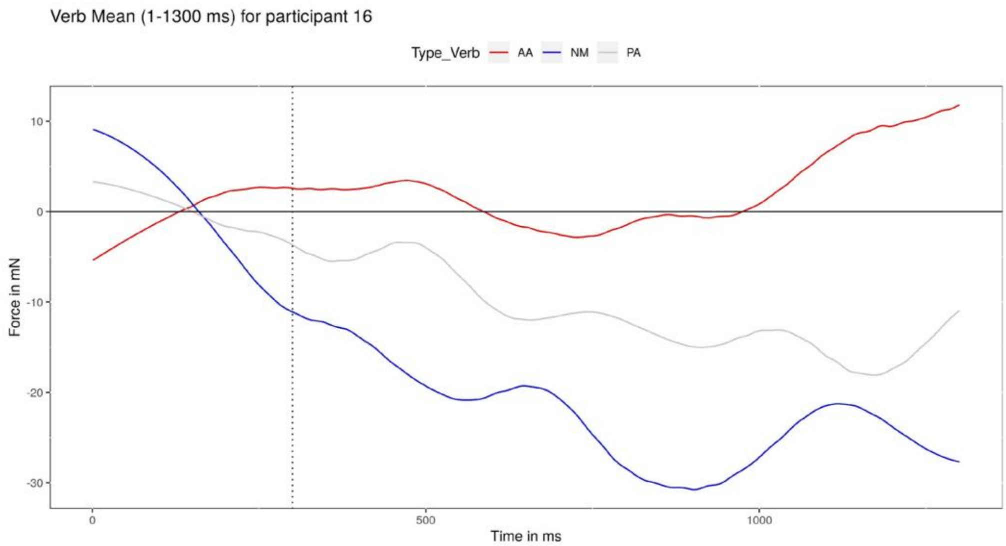


Figure 3. An example of negative drift

Although there is large variation across participants, a rapid (during the first 1000 ms) and uniform decrease in intensity is unusual. We preferred to ignore the contribution of participants with negative drifts because there was the risk of incorporating data which did not correspond to the experimental conditions. Moreover, following the filtering options of Aravena et al. (2012, 2014), we also eliminated trials that showed points above 200 mN or below -150mN. To detect such points, we chose a relatively large time window starting from verb onset to 500 ms after noun onset. All the final data sets are provided in the Supplementary Material.

The statistical analysis is a bit technical. To keep the focus on the results and their interpretations, we provide only a compact summary in the rest of this section, leaving a more detailed exposition to the appendix. We analyzed the grip force variations using two strategies. First, we ran linear mixed-effects models in a (constant shift/constant span) moving window setting. Specifically, the variations for the different conditions were statistically compared over 300 ms intervals (constant span). The endpoints of the 300 ms time interval were gradually shifted to the right by 100 ms. So, 1-300 ms, 101-400 ms, 201-500 ms, 301-600 ms, etc. time-windows were investigated in succession. The chosen models were *maximal*, in the sense of Barr, Levy, Scheepers and Tily, 2013[2]. In the context of our experiments, this means running models with the structure described in Table 1.

Table 1.
Structure of the maximal mixed-effect models

Fixed effect	Random effect 1	Random effect 2
Intensity of grip force (dependent variable) in function of condition (independent variable)	Participant (intercept and slope)	Item (intercept)

The fixed effect measures the dependence of grip force on condition (like in any standard linear model with categorical independent variable(s)). The first random effect takes into account a possible individual sensitivity of participants to (i) the experimental device, for instance in relation to their particular grip force strength and (ii) the various conditions. The second random effect takes into account possible differences between the various items (sentences) presented to participants. While the fixed and first random effects are only marginally different from what a standard ANOVA with repeated measures estimates, the second random effect makes a genuine difference and contributes to seriously limit the type I error rate, making this type of maximal model the currently recommended choice for most behavioral experiments in psychology (Brauer & Curtin, 2018; Singmann & Kellen, 2020).

[2] Given that we have no interaction structure, the problem of negotiating type I error against power loss does not occur here (Bates, Kliegl, Vasishth & Baayen, 2015; Matuschek, Kliegl, Vasishth, Baayen & Bates, 2017).

It turns out that our data are not linear. As a result, although linear mixed-effects models provide us with an approximation, it is safer to complement them with non-parametric measures and to accept the existence of an effect only if all the tests agree on its direction. Specifically, we started from the grip force values observed for *items and participants*. For each 300 ms time window, each participant and each item in each condition, we recorded the mean of the item. We ranked the items by means in ascending order and for each pair of conditions, compared their grip force values for each time point. An item was considered as 'winning' over its competitor if the former had at least 1.5 more higher values than the latter. Otherwise the competition resulted in a 'tie'. To illustrate, in the factive experiment (Experiment 1), the first two rows of the comparison table are as follows.

Table 2.
An example of time window ′ participant ′ contrast pair comparison. First two rows of the global count table for Experiment 1. (*NA = Non-action, PA = Presupposed action*)

Window	Comparison	Part.	Items	Winners	Losers	W. Counts	L. Counts
1-300	NA-PA	1	ruine (*ruins*) vs. voit (*sees*) - caresse (*caresses*)	tie	tie	159	141
1-300	NA-PA	1	impose (*imposes*) vs. sait (*knows*)-peigne (*combs*)	PA	NA	259	41

The first three columns indicate that we are in the first time window (1-300 ms), comparing items of type Non-Action (NA) and Presupposed-Action (PA) for participant 1. The *Items* column contains abbreviated names for the item pairs. Remember that, for each pair, the item of type Non-Action (e.g. an item containing a verb phrase of the form X *Ruine*$_{\text{NON-ACTION}}$ Y) occupies the same rank as the item of type Presupposed-Action (e.g. an item containing a verb phrase of the form X *Voit*$_{\text{FACTIVE}}$ *que* Y *Caresse*$_{\text{ACTION}}$ Z). The *Winners* (*Losers*) column shows the winners (losers) and the last two columns the corresponding figures. In the first row, there is a tie because the count difference is small (18 values). In the second row, the PA items has more than six times more higher values than the NA item.

We extract various information from this initial table. The most important are (i) the total counts and (ii) counts by participant. The total counts are the sums of counts across participants for each condition, *excluding* the counts of ties. We used count sums to compare differences between conditions by means of

Fisher tests. For instance, in the factive experiment, we compared the count sum contrast between Action and Non-Action with the count sum contrast between Action and Presupposed Action. Is one of these contrasts significantly bigger than the other or are they in the same order of magnitude? The counts by participant are, for each temporal window and pairs of conditions, the numbers of winners in each condition for each participant, again excluding ties. For instance, in the 1-300 ms window of the count data for Experiment 1, we have, when comparing Non-Action (NA) and Presupposed Action (PA): 20 vs. 3 for participant 1, 2 vs. 18 for participant 2, 4 vs. 17 for participant 4, etc. We ran Wilcoxon paired tests on such vector pairs. For instance; the Wilcoxon paired test does not detect a significant difference between the two mentioned vector (p = .13). Running through the contrasted scores by participant, the test tells us whether a condition produces significantly more winners than the compared condition.

We are interested in comparing the results of the mixed-effects model, the Fisher test and the Wilcoxon test. Suppose that, for some time-window, the mixed model delivers a significant p value when comparing conditions C1 and C2 and a non-significant p value when comparing C1 and C3. If, in addition, (i) the Wilcoxon test also delivers a significant p value when comparing C1 and C2 but no significant p value when comparing C1 and C3 and (ii) the Fisher test tells us that the contrasts C1 vs. C2 and C1 vs. C3 are significantly different, we can be reasonably sure that some effect takes place which separates C1 and C2 but not C1 and C3. This and similar configurations will be our main targets in the statistical analysis.

3. EXPERIMENTS

3.1. Experiment 1: Factivity

Method

Participants

30 participants (25 women, 18 – 32 years old; M_{age} = 21.7, SD_{age} = 1.55) participated in this study. All were right-handed ($M_{laterality}$ = .83; $SD_{laterality}$ = .165; cf. Oldfield, 1971).

Stimuli

A total of 111 French sentences served as stimuli. 37 target hand-related action verbs were used. 8 distinct French factive verbs were used with respect to the factive stimuli: *voir* (*to see*, 5 times), *s'apercevoir* (*to realize*, 3 times), *entendre* (*to hear*, 5 times), *réaliser* (*to realize*, 6 times), *remarquer* (*to notice*, 6 times), *observer* (*to observe*, 5 times), *se rendre compte* (*to realize*, 2 times), and *savoir* (*to know*, 5 times). In addition, 37 sentences containing asserted Non-Action verbs served as control sentences (see Aravena et al., 2012; 2014; Frak et al., 2010; Nazir et al., 2017). The action verbs and asserted Non-Action verbs were controlled for number of letters and number of syllables (New, Pallier, Ferrand, & Matsos, 2001). Three examples of experimental stimuli are provided in Table 3.

All critical verbs were in the present tense and in neutral third person. Verbs always appeared on the fifth position (±1) of the sentence. The onset of the target verb and the total duration of the sentence was determined using *PRAAT*. The onset of the critical verb and noun for the assertive action condition (*Ines ties her shoes*) were on average 1406ms (*SD* = 202ms) and 1882ms (*SD* = 236ms) after the beginning of the sentence; for the Presupposed Action (*Paul knows that Ines ties her shoes*), they were on average 1255ms (*SD* = 157ms) and 1676ms (*SD* = 191ms) ; for the Non-Action condition (*Philippe hires a player*) they were 1257ms (*SD* = 181ms) and 1734ms (*SD* = 215ms).

Table 3.
Example of stimuli used in Experiment 1 and their approximate English translation

Condition	**Sample stimulus**	**English approximate translation**
Asserted Action	Avant de partir, Ines **lace** ses chaussures.	Before leaving, Ines **ties** her shoes.
Presupposed Action	Daniel voit qu'Anne **lace** ses chaussures.	Daniel sees that Anne **ties** her shoes.
Non-Action	Pour le dîner, Pierre **souhaite** du poulet.	For dinner, Peter **would** like chicken.

The order of the three conditions was pseudo-randomized. No more than two items of the same category appeared consecutively. To control for order effects, we independently generated random orderings for each participant. The 111 sentences were divided into 10 blocks. The first 9 blocks contained 11 sentences, the last one 12 sentences. After each block a yes/no comprehension question concerning the action/non-action part of the sentences was asked in order to keep participants attentive during listening to the auditory stimuli. The amount of yes/no questions was balanced, that is, a participant answered either 5 yes and 6 no question or 6 yes and 5 no questions. We did not measure accuracy because the task involved memory, not motor response, and the goal was only to keep participants more attentive.

Data Analysis

As explained above, we first examined the differences between the different conditions using maximal mixed-effects models on successive time intervals, shifted by 100 ms. In contrast to other experiments of the same type (Aravena et al., 2012; 2014), we took into account a large global time span ranging from verb onset to 1000ms after noun onset. This allows one to observe possible effects of noun phrases. With a sentence like *Ines ties her shoes*, one can expect to detect a motor response 'after' *ties*, but it is not a priori clear whether the noun phrase *her shoes* plays a role in triggering the response, or, in other terms whether the semantic content of the verb alone is sufficient or whether the full verb phrase *ties her shoes* adds significantly to the motor response.

5 participants were removed because of negative drift (3, 10, 12, 15, 19). 24 items with an intensity below - 150mN or above 200 mN were suppressed (11 for Asserted Action, 9 for Presupposed Action and 4 for Non-Action).

Results

The averaged results of grip force activation for all three conditions are depicted in Figure 4.

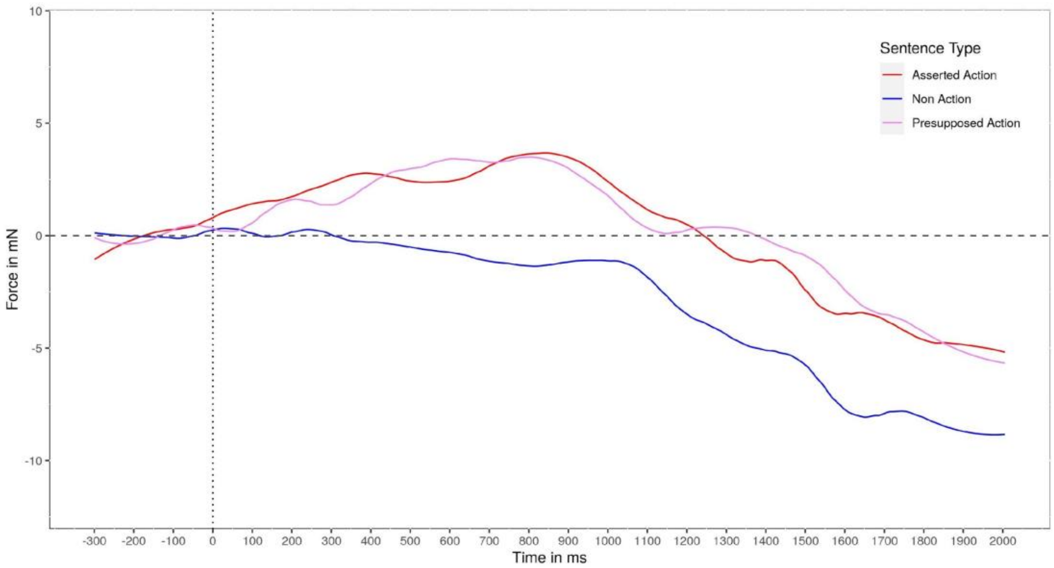


Figure 4. Modulation of grip force amplitude across conditions in Experiment 1.

The statistical results indicate that significant differences exist in grip force variation for the 500 – 800 ms, 600-900 ms and 700-1000 ms time intervals between the Asserted Action and Non-Action condition as well as between the Presupposed Action and Non-Action one. The Fisher tests[3] are consonant with the contrasts calculated by the mixed-effects models and Wilcoxon tests. Asserted Action (respectively Presupposed Action) is more different from Non-Action than from Presupposed Action (respectively Asserted Action) (see for example [1] (respectively [2]) in Table 5). The Asserted Action vs. Non-Action and Presupposed Action vs. Non-Action contrasts are not evaluated as significantly different. except in the 500-800 window (Table 5, [3]). This is due to the fact that the Asserted Action condition is comparatively less distinct from Non-Action than Presupposed Action.[4]

Table 5.
P values for the mixed model, Wilcoxon test, and Fisher exact test of Experiment 1. * $p < .05$, ** $p < .001$

Windows	Mixed model	Wilcoxon test	Contrasts	Count scores	Fisher
501-800	.065 .8	.04* .15	AA vs. NA AA vs. PresA	445 vs. 234 338 vs. 376	<.001**
501-800	.065 .045*	.04* .02*	NA vs. AA NA vs. PresA	234 vs. 445 209 vs. 544	.007* [3]
501-800	.8 .045*	.15 .02*	PresA vs. AA PresA vs. NA	376 vs. 338 544 vs. 209	<.001**
601-900	.052 .98	.02* .7	AA vs. NA AA vs. PresA	461 vs. 239 353 vs. 363	<.001** [1]
601-900	.052 .037*	.02* .005*	NA vs. AA NA vs. PresA	239 vs. 461 221 vs. 505	0.14
601-900	.98 .037*	.7 .005*	PresA vs. AA PresA vs. NA	363 vs. 353 505 vs. 221	<.001** [2]
701-1000	.064 .8	.02* 1	AA vs. NA AA vs. PresA	475 vs.237 357 vs. 353	<.001**
701-1000	.064 .06	.02* .008*	NA vs. AA NA vs. PresA	237 vs. 475 238 vs. 496	0.73
701-1000	.8 .06	1 .008*	PresA vs. AA PresA vs. NA	353 vs. 357 496 vs. 238	<.001**

We note two additional points. First, Figure 4 shows a steady decrease of intensity, relatively uniform across conditions, starting at about 850-1000 ms after verb onset. This is a general phenomenon, which can be observed in the three experiments. It probably reflects a two-stage automatic process: participants focus on the stimulus and, then, relax before the next stimulus. Second, the *p* values obtained are moderate, in particular when compared to those reported in previous similar studies (e.g. Aravena et al., 2012, 2014). This is not surprising given that we used maximal models. To illustrate the difference with more standard models, one can run the 'equivalent' of a RM ANOVA for mixed-effects models, that is a model where the

[3] To recall, count scores sum winner items across participants for each condition in the Contrasts column (see Table 5).
[4] Barplots for significance along time can be found in the Supplementary Material.

random intercept for items is suppressed. The *p* values are then as follows: for the 500-800 window, p = .017 for the Asserted-Action vs. Non-Action comparison, p = 0.01 for Presupposed-Action vs. Non-Action and p = 0.79 for Asserted Action vs. Presupposed Action (for the 600-900 window, the *p* values are .015, .01 and .97, respectively to the previous mentioned comparisons). Clearly, these significance figures are bigger only because one ignores the item-based variation, which, to repeat, is quite important.

Discussion

The first experiment addresses the question of whether true but backgrounded action-related information activates motor brain structures. If presuppositions are considered as true by default, it seems that they should trigger a motor response. However, since presuppositions are backgrounded information, they might not elicit the same response compared to simple assertions. Our results reveal that Presupposed Action verbs elicit an increase in grip force. More precisely, grip force in the Presupposed Action condition is significantly higher than in the Non-Action condition and does not differ from the grip force in the Asserted Action condition.

Previous research has shown that language-induced motor activation is not triggered by the presence of an action verb *per se* but depends on contextual factors – cf. the interpretation of metaphors and idioms, the presence of a negation operator or of a volitional verb. More precisely, negative operators (Aravena et al., 2012; Tettamanti et al., 2008) and volitional contexts neutralize such an activation (Aravena et al., 2014; Zwaan et al., 2010). Furthermore, discourse properties – such as the layering of information of what is said and what is implicated – also have an impact on the involved motor structures (van Ackeren et al., 2012). The present study extends these findings to the presupposition triggered by factive verbs.

We used factive verb constructions to manipulate the layering of the described action. Under some approaches to presupposition (e.g. Stalnaker, 1974), the action *Mary throws the ball* is considered as novel information in an assertion as in *Mary throws the ball*, whereas the same information is considered as backgrounded in a factive verb construction as in *Paul knows that Mary throws the ball*. Our results indicate that the novel and the backgrounded information trigger a comparable increase in grip force. Such a result appears at a first sight counter-intuitive with respect to other studies on presupposition processing, which show that processing accommodated presuppositions comes with a transient processing cost (EEG study by

Domaneschi et al., 2018; Masia et al., 2017; self-paced reading study by Tiemann et al., 2011; Domaneschi & Di Paolo, 2017; eye-tracking study by Tiemann & Schwarz, 2012). Since we used decontextualized sentences, accommodating the presupposition could thus have either weakened or delayed the onset of the grip force effects. However, the results show that the increase of grip force for Presupposed Action verbs starts at least as early as this is the case for Asserted Action verbs. Moreover, the trajectory of the grip-force curve does not differ significantly in the two conditions. This suggests that the action denoted by the factive complement (for instance throwing a ball or tying one's shoes) is immediately integrated in the situation model, as it is for an assertive sentence.

Sensori-motor activation is triggered by linguistic contexts where the action denoted by the corresponding verbal group actually takes place. According to the *Linguistic Focus Hypothesis*[5] (Zwaan and Taylor, 2008), a motor resonance is triggered when the action presents the focus of an utterance. Hence, negative (*Marie does not throw the ball*) and volitional sentences (*Paul wants to throw the ball*) do not give rise to the phenomenon observed for assertive sentences (*Marie throws the ball*), simply because the action under a negative or volitional operator is not (yet) true in a model of the current situation. Consequently, the linguistic surrounding can switch off motor semantic features when they appear to be irrelevant within the situation model. If we admit that a sentence like *Paul knows that Mary throws the ball* communicates the truth of the complement clause, the situation model includes the proposition that *Mary throws the ball* as the sentence corresponding to a simple assertion does. From this point of view, it is not surprising that the grip force activation of the Presupposed Action verb has a comparable trajectory as that of the Asserted Action verb.

More generally, the results indicate that the *truth-conditional status* of the presupposed information (about a hand-related action) elicits an increase in grip force. This supports the idea that presuppositions engage the speaker's *commitment* (see Peters, 2016), or, in other terms, that the speaker who uses a presupposition presents himself as believing in his truth. When the addressee has no particular reason to question the beliefs of the speaker, she takes them for granted if she considers the speaker as sufficiently

[5]It is important to note that Zwaan and Taylor's (2008) use of the term *focus* is different from what linguists call 'focus'. For Zwaan and Taylor, focus is linked to an action that takes place at the current time point.

reliable in terms of honesty and competence. To ensure that the observed increase in grip force relates to the fact that the hand-related action verb occurs in the complement of a factive verb that guarantees its truth (e.g. *know*), we designed a second experiment where we replaced factive verbs with non-factive verbs such as *believe* or *think*. If our hypothesis is correct, this manipulation should weaken or neutralize the motor effect.

3.2. Experiment 2: Non-Factivity

Method

Participants

34 participants (10 men; 19 – 35 years old; M_{age} = 22.71, SD_{age} = 4.03) participated in this study. All were right-handed ($M_{laterality}$ = .95; $SD_{laterality}$ = .15).

Stimuli

A total of 111 French sentences served as stimuli (see Supplementary Material). Thirty-seven target hand-related action verbs were embedded into Asserted Action and Non-Presupposed sentences. In addition, thirty-seven sentences containing asserted Non-Action verbs were used. In contrast to experiment 1, the sentences for Presupposed Action were replaced by Non-Presupposed ones (see Table 6). We avoided to have both factive and non-factive sentences in the same experiment in order to prevent a contrastive reading (*know* vs. *believe*), which might have induced the participants to interpret the complement of a non-factive verb as (probably) false. 8 distinct French non-factive verbs were used with respect to the factive stimuli: *imaginer* (*to imagine*, 5 times), *dire* (*to say*, 5 times), *soupçonner* (*to suspect*, 4 times), *suspecter* (*to suspect*, 5 times), *penser* (*to think*, 5 times), *croire* (*to believe*, 5 times), *supposer* (*to suppose*, 4 times), and *soutenir* (*to claim*, 4 times). All other selection and condition criteria used for experiment 1 also applied for this experiment.

Table 6
Example of Stimuli Used in Experiment 3 and their Approximate English Translation

Condition	Sample stimulus	English approximate translation
Asserted Action	Avant de partir, Ines lace ses chaussures.	Before leaving, Ines **ties** her shoes.
Non-Presupposed Action (Non-factive construction)	Daniel imagine qu'Anne lace ses chaussures.	Daniel imagines that Anne **ties** her shoes.

Non-Action	Pour son repas, Pierre **souhaite** du poulet.	For his meal, Peter **would** like chicken.

Equipment, data acquisition**,** and ***procedure*** were exactly the same as presented in Experiment 1.

Data Analysis

The data analysis was exactly the same as in Experiment 1.

4 participants were ignored because of negative drift (participants 2, 16, 19, 31). 102 items with a grip force below -150mN or above 200 mN were suppressed (39 for Asserted Action, 34 for Non-Presupposed Action and 29 for Non-Action).

Results

The averaged results of grip force activation for all three conditions are depicted in Figure 5.

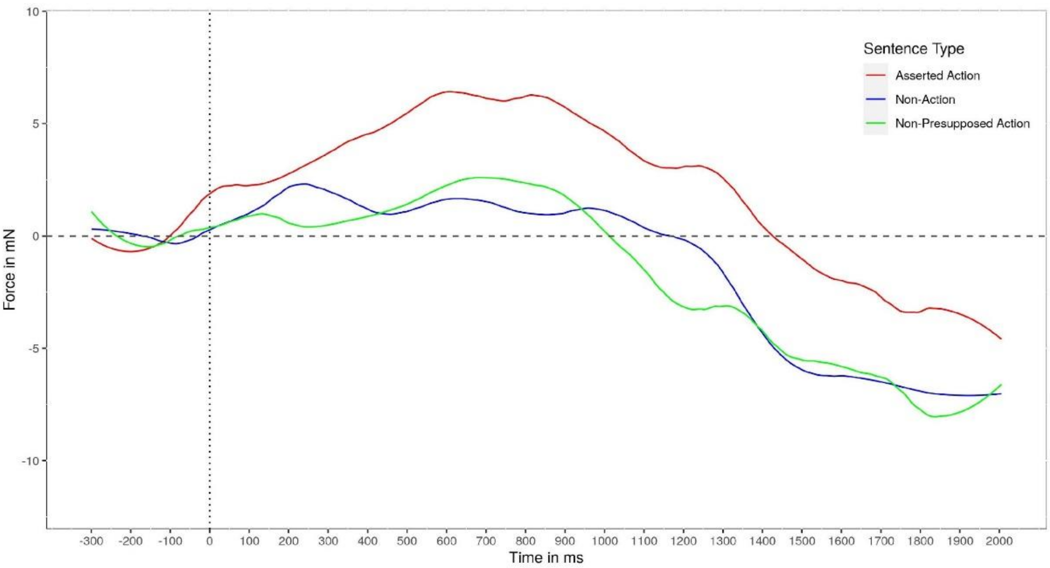


Figure 5**.** Modulation of grip force amplitude across conditions in Experiment 2.

The results indicate that significant differences emerge in the time interval between 300 – 700 ms. More precisely, the results reveal that Asserted Action differs significantly in comparison to the Non-Presupposed Action condition and Non-Action one. When compared to the Non-Presupposed Action condition, the Asserted Action condition shows a moderate significance, which is in sharp contrast to the Asserted Action vs. Non-Action comparison, which also spans to a wider time interval, i.e. until 1000 ms. Combining the *p* values, the count scores and Fisher results, the results indicate that Non-Presupposed Action and Non-Action conditions do not differ significantly. Whenever a Non-Presupposed Action vs. Non-Action contrast is compared to another contrast, the Fisher tests are significant. Smaller but significant *p* values are also observed when Asserted Action is compared to the other two conditions (Table 7, see [1]-[5])[6].

Table 7

*P Values for the Mixed Model, Wilcoxon Test, and Fisher Exact Test of Experiment 2. * p < .05, ** p < .001*

Windows	Mixed model	Wilcoxon test	Contrasts	Count scores	Fisher
301-600	.05* .07	.008* .098	AA vs. NA AA vs. NPresA	506 vs. 237 485 vs. 305	.006* [1]
301-600	.05* .93	.008* .8	NA vs. AA NA vs. NPresA	237 vs. 506 378 vs. 399	<.001**
301-600	.07* .93	.098 .8	NPresA vs. AA NPresA vs. NA	305 vs. 485 399 vs. 378	<.001**
401-700	.03* .08	.03* .09	AA vs. NA AA vs. NPresA	537 vs. 241 477 vs. 311	<.001** [2]
401-700	.03* .86	.03* .5	NA vs. AA NA vs. NPresA	241 vs. 537 371 vs. 430	<.001**
401-700	.08 .86	.09 .5	NPresA vs. AA NPresA vs. NA	311 vs. 477 430 vs. 371	<.001**
501-800	.024* .1	.008* .1	AA vs. NA AA vs. NPresA	524 vs. 273 488 vs. 316	.038 *[3]
501-800	.024* .76	.008 .3	NA vs. AA NA vs. NPresA	273 vs. 524 345 vs. 441	<.001**
501-800	.1 .76	.1 .3	NPresA vs. AA NPresA vs. NA	316 vs. 488 441 vs. 345	<.001**
601-900	.025* .137	.007* .1	AA vs. NA AA vs. NPresA	532 vs. 273 490 vs. 309	.048 *[4]
601-900	.025* .69	.007* .5	NA vs. AA NA vs. NPresA	273 vs. 532 345 vs. 419	<.001**
601-900	.13 .69	.1 .5	NPresA vs. AA NPresA vs. NA	309 vs. 490 419 vs. 345	<.001**
701-1000	.045* .147	.009* .16	AA vs. NA AA vs. NPresA	534 vs. 275 484 vs. 343	<.001** [5]
701-1000	.045* .79	.009* .66	NA vs. AA NA vs. NPresA	275 vs. 534 374 vs. 427	<.001**
701-1000	.147 .79	.16 .66	NPresA vs. AA NPresA vs. NA	343 vs. 484 427 vs. 374	<.001**

Discussion

The second experiment directly compares the action-related content of an assertion to the non-factive complement. Without prior context, the truth of the non-presupposed complement is unknown, that is, the information is neither true nor false. If, as we assume, the truth of the complement is a prerequisite for the recruitment of motor structures during the processing of action verbs, a weaker or null increase of the grip

6 Barplots for significance along time can be found in the Supplementary Material.

force should be expected with non-factive complements. Our results show that the Asserted-Action condition shows a significant increase in grip force when compared to the Non-Action or Non-Presupposed Action conditions, whereas the difference between the Non-Presupposed Action and Non-Action conditions is not significant. This contrast suggests that Non-Presupposed Action verbs (Experiment 2) and Presupposed Action verbs (Experiment 1) trigger different grip force activations. However, the *p* values and the results of the Fisher tests are compatible with a more nuanced hypothesis, namely that Non-Presupposed Action is slightly more susceptible to motor response than Non-Action. Admittedly, the observed differences are small but this is not a priori unlikely, given that the sentences in the Non-Presupposed Action condition describe a hand-related action occurring in a situation which, though not presented as the *actual* situation, is still a *possible* situation, whose truth is assumed by an agent different from the speaker.

Taken together, the results of experiments 1 and 2 suggest that the driving force behind the observed grip force modulations is the truth-conditional status of the action-related verb. In experiment 1, the presupposition of a factive verb (e.g. *know*) is presented as true and the observed motor activation is not different from that of action-related verbs in simple assertive sentences. In experiment 2, the presupposition is not presented as true, since it is embedded under a non-factive verb (e.g. *believe*), which does not presuppose the truth of the complement clause. In that case, the grip force does not reach the activation of Asserted-Action verbs and, in fact, does also not differ significantly from Non-Action verbs. Contrariwise, the grip force activation of the Presupposed-Action condition, as observed in experience 1, is significantly from that of the Non-Action condition. Overall, the results of the two experiments confirm that action-related verbs in themselves are not always sufficient to generate a motor response and that the linguistic environment plays a crucial role (e.g. Willems & Casasanto, 2011).

Our results suggest that the presuppositional status in itself is not different from the asserted status for factive constructions, although differences have been observed when presuppositions are put into a discourse context (see for instance Masia et al., 2017, for definite versus indefinite descriptions and Simons, Beaver, Roberts, & Tonhauser, 2017, for factive constructions). A part of the theoretical literature on presuppositions assumes that, by default, presuppositions project, that is, are considered as true under certain operators like negation or interrogation. Accordingly, one might argue that they should trigger a motor activation under these operators. But, even though the truth-conditional status plays an important role, it is perhaps not sufficient to counteract the effect of operators which express opposition (negation) or uncertainty (interrogation). Admittedly, negation or interrogation do not bear directly on the presupposition. A sentence like *Paul doesn't know that Mary writes the letter* negates a certain knowledge of the agent Paul, but not the proposition that Mary writes the letter. Still, it might be the case that the negation affects the force of the presupposition. This can be done in at least two ways. First, negation could be *parasitic* on the presupposition, meaning that, although it does not combine with the presupposition, it could somehow 'taint' it. For instance, Aravena et al. (2012) suggest that negation could block the motor semantic representation of the negation target (for candidate neurophysiological grounds for this idea see de Vega et al., 2016; Papeo et al., 2016; Tettamanti et al., 2008). Second, it has been argued that, in some cases, negated factive verbs do not give rise to projection (Beaver, Roberts, Simons, & Tonhauser, 2017; Simons et al., 2017). For instance, a sentence like *Paul didn't observe that Mary was in the office* can mean either that Mary was in the office and Paul did not notice her (the projection interpretation) or that Paul had no evidence that Mary was in the office (the non-projection interpretation). Adopting a projective reading, the perspective of the speaker outweighs the perspective of Paul (the agent). In contrast, a non-projective reading focuses on the perspective of the agent. Consequently, the latter interpretation should not elicit a grip force activation, whereas the former one should elicit one. The goal of our third experiment is to determine whether the negation operator influences the motor response in projective environments.

3.3. Experiment 3: Projection

Method

Participants

29 participants (14 men; 18 – 30 years old; M_{age} = 21.06, SD_{age} = 3.22) participated in this study. All were right-handed ($M_{laterality}$ = .91; $SD_{laterality}$ = .19).

Stimuli

A total of 111 French sentences served as stimuli. We decided to have a slightly more complex context clause for projective environments (a full sentence instead of a prepositional clause). This is due to the fact that, in some cases, having only a prepositional clause made the full target sentence somewhat unclear. For instance, *In the launderette, Michael does not know that Cédric irons his shirt* does not a priori make much sense if Cédric is not himself in the launderette. To solve this referential problem, we replaced *in the launderette* by the sentence *Cédric is in the launderette*. Sentences for other conditions were modified accordingly. Hand-related action verbs always appeared on the twelfth position (±2) of the sentence.

9 distinct French factive verbs were used under negation in the projective construction: *voir* (*to see*, 6 times), *s'apercevoir* (*to realize*, 4 times), *entendre* (*to hear*, 4 times), *réaliser* (*to realize*, 4 times), *remarquer* (*to notice*, 4 times), *observer* (*to observe*, 5 times), *se rendre compte* (*to realize*, once), *savoir* (*to know*, 5 times) and *constater* (*witness*, 4 times). A sample of stimuli is provided in table 8. All previous selection and condition criteria used for experiments 1 and 2 also applied for this experiment.

Table 8

Example of Stimuli Used in Experiment 3 and their Approximate English Translation

Condition	Sample stimulus	English approximate translation
Asserted action	Ines va partir pour aller travailler. Avant de sortir, elle lace ses chaussures.	Ines is leaving for work. Before going out, she **ties** her shoes.
Projected action	Robert est occupé dans le salon. Il ne voit pas que Ghislaine lace ses chaussures.	Robert's busy in the living room. He does not see that Ghislaine **ties** her shoes.
Non-Action	Samuel préfère de beaucoup la volaille, Pour le dîner il **souhaite** du poulet.	Samuel greatly prefers poultry. For the dinner he **would** like chicken.

Measures and pre-tests

To ascertain that the negation of a factive verb does, indeed, leave the factive complement unaffected, we first tested the projection of the factive complement in an online pilot study. Twenty-four French native speakers, aged from 21 to 48 years participated in this study (M = 31.66, SD = 9.82). None of them followed a program in linguistics. Each participant saw five (randomly selected) of the thirty-seven presupposed action-projection sentences and ten filler sentences. After having read the sentence, the participant had to indicate whether the factive complement was true or false. In 84.2% of all questions, the factive complement was rated as true, whereas in 15.8% the factive complement was rated as false. This difference is significant (z = 10.59, p < .001, CI for correct answers = 70.10% - 90.70%). In addition, the correct results also differ significantly from chance (z = 7.62, p < .0001). The results can be seen as evidence that, by default, the factive complement projects, that is, it remains unaffected under a negative operator.

Equipment, data acquisition, and ***procedure*** were exactly the same as presented in experiment 1.

Data Analysis

The data analysis was exactly the same as in Experiment 1.

Participant 12 was removed because of recording problems. Furthermore, 6 participants were removed since they their grip force recordings showed a negative drift (3, 15, 16, 19, 20, 25). 163 items with a grip force below -150mN or above 200 mN were suppressed (51 for Asserted Action, 53 for Projection and 59 for Non-Action).

Exploratory Analysis

It has been noted that projection is not an automatic or effortless process. In order to investigate whether projection elicits a grip force at a later stage, we investigated grip force activations in the temporal window starting at the onset of the noun until the end of the sentence.

Results

The averaged results of grip force activation for all three conditions are depicted in Figure 6.

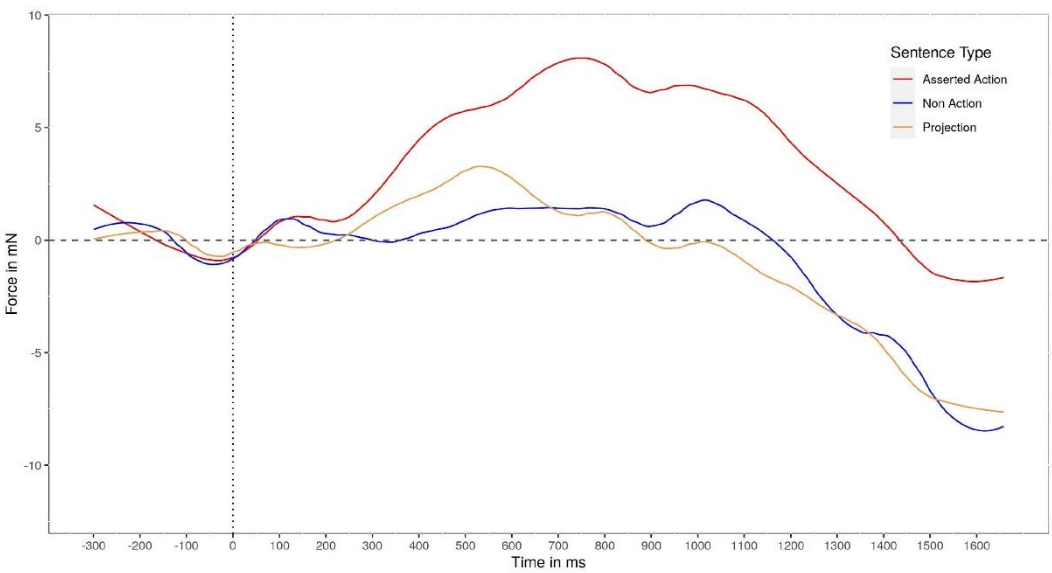


Figure 6. Modulation of grip force amplitude across conditions in Experiment 3.

The results point out that significant differences emerge between 600 – 1000 ms after the onset of the action verb between the Asserted Action and Projection condition. Moreover, the findings indicate that only the Asserted Action condition elicits a grip force activation, whereas this is not the case in the Projection and Non-Action condition. The Asserted Action condition elicits a significant higher grip force activation than the Projection and Non-Action condition.

The Fisher tests in Table 9 show that the contrasts between Asserted Action on one side and Non-Action and Projection on the other side are quite comparable. This agrees with the results of the mixed-effects models, which indicate that Projection is close to Non-Action[7].

Table 9

*P Values for the Mixed Model, Wilcoxon Test, and Fisher Exact Test of Experiment 3. * p < .05, ** p < .001*

Windows	Mixed model	Wilcoxon test	Contrasts	Count scores	Fisher
501-800	.066 .074	.008* .005*	AA vs. NA AA vs. ProjA	445 vs. 175 420 vs. 162	.9 [1]
501-800	.066 .81	.008* .34	NA vs. AA NA vs. ProjA	175 vs. 445 255 vs. 336	<.001*
501-800	.074 .81	.005* .34	ProjA vs. AA ProjA vs. NA	162 vs. 420 336 vs. 255	<.001*
601-900	.058 .04*	.006* .005*	AA vs. NA AA vs. ProjA	457 vs. 166 422 vs. 162	.7 [2]
601-900	.058 1	.006* .54	NA vs. AA NA vs. ProjA	166 vs. 457 268 vs. 321	<.001*.
601-900	.04* 1	.005* .54	ProjA vs. AA ProjA vs. NA	162 vs. 422 321 vs. 268	<.001*
701-1000	.07 .033*	.01* .004*	AA vs. NA AA vs. ProjA	456 vs. 165 426 vs. 180	.23 [3]
701-1000	.07 .86	.01* .56	NA vs. AA NA vs. ProjA	165 vs. 456 279 vs. 339	<.001*
701-1000	.033* .86	.004* .56	ProjA vs. AA ProjA vs. NA	180 vs. 426 339 vs. 279	<.001*

Exploratory analysis

It has been noted that projection is not an automatic or effortless process, a point to which we return in the next discussion section. Taking this possibility into account, we decided to investigate whether the limits of our temporal windows (1000 ms after noun onset) had possibly prevented us from detecting some relevant phenomenon. The intuition was that we might have missed some late episode in the response to the sentences, between the noun onset and the beginning of the next auditory stimulus. Figure 9 shows the last part of the average grip-force activations across participants and items for our three experiments. While the two plots concerning the experiments on factivity and non-factivity do not show anything different from a simple pressure relaxing, before the participants refocus on the next stimulus, the plot for the projection

[7] Barplots for significance along time can be found in the Supplementary Material.

experiment suggests that the Projection condition is associated with a rise starting at about 1300 ms after noun onset.

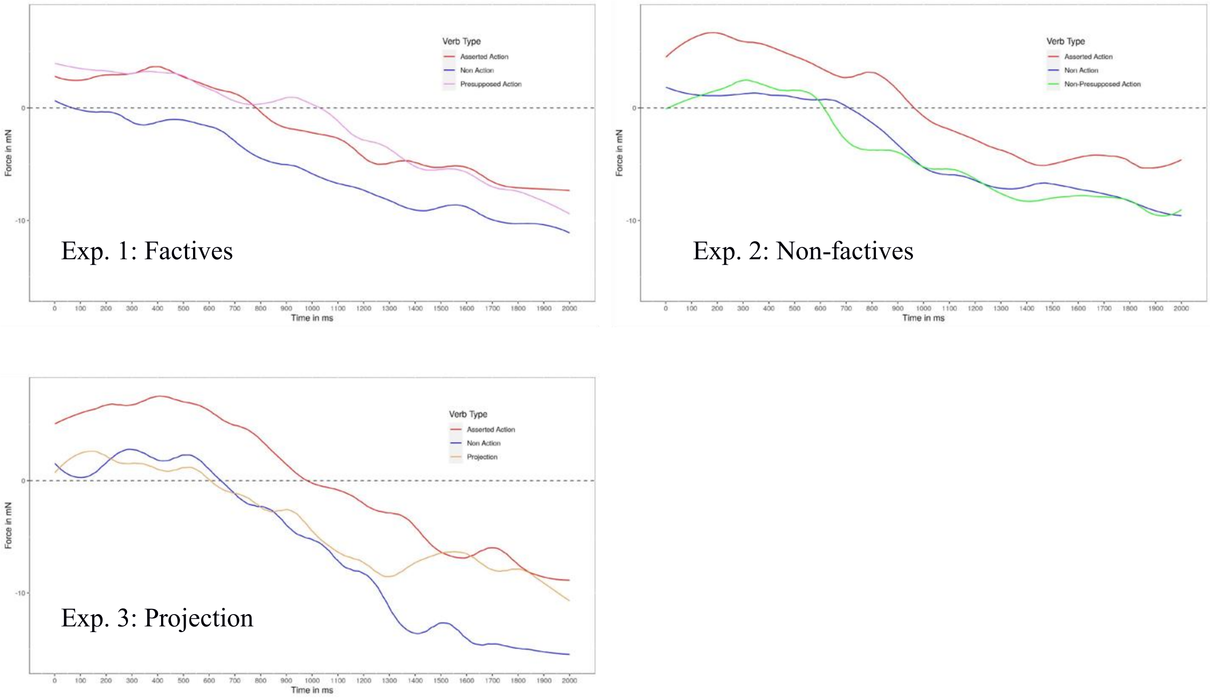


Figure 7. Plots for the Last Part of the Time Point Series: Factive (top left), Non-Factive (top right) and Projection (bottom left)

The results are summarized in Table 10. The Mixed model column does not contain any significant or approximately significant figure. But the *p* values for the Asserted Action vs. Projection contrast are markedly superior to those for the Asserted Action vs. Non-Action contrast in all the regions mentioned in table 10. This is not the case for the contrasts Non-Action vs. Asserted Action and Non-Action vs. Projection, which are always similar. The Wilcoxon tests also deliver larger values for Asserted Action vs. Projection than for Asserted Action vs. Non-Action. They deliver *inferior* values for Non-Action vs. Asserted Action when compared to Non-Action vs. Projection, except for the last interval (1700-2000ms post noun onset) where the figures are comparable. The Fisher tests are always significant − although on different scale, except for the last two intervals. The *p* values obtained through the mixed-effects models and the Wilcoxon tests suggest that Projection is closer to Asserted Action than to Non-Action. According to the mixed-effects models, Asserted Action and Projection are equidistant from Non-Action whereas, according to the Wilcoxon tests, Projection is closer to Non-Action. The Fisher tests indicate similar distributions of counts for the Non-Action vs. Assertion/Projection in the 1600-1900 and 1700-2000 windows[8].

Table 10

*P Values for the Mixed Model, Wilcoxon test, and Fisher Exact Test of the Last 2000 ms of Experiment 3. * p < .05, ** p < .001*

Windows	Mixed model	Wilcoxon test	Contrasts	Count scores	Fisher
1201-1500	.098 .32	.02* .09	AA vs. NA AA vs. ProjA	444 vs. 207 382 vs. 233	.025* [1]
1201-1500	.098 .38	.02* .24	NA vs. AA NA vs. ProjA	207 vs. 444 253 vs. 371	<.001**
1201-1500	.32 .38	.09 .24	ProjA vs. AA ProjA vs. NA	233 vs. 382 371 vs. 253	<.001**
1301-1600	.13 .68	.01* .42	AA vs. NA AA vs. ProjA	454 vs. 203 349 vs. 273	<.001** [2]
1301-1600	.13 .22	.01* .12	NA vs. AA NA vs. ProjA	203 vs. 454 244 vs. 397	.007
1301-1600	.68 .22	.42 .12	ProjA vs. AA ProjA vs. NA	273 vs. 349 397 vs. 244	<.001**
1401-1700	.16 .93	.02 .8	AA vs. NA AA vs. ProjA	461 vs. 197 320 vs. 284	<.001** [3]
1401-1700	.16 .18	.02* .09	NA vs. AA NA vs. ProjA	197 vs. 461 232 vs. 400	.01
1401-1700	.93 .18	.8 .09	ProjA vs. AA ProjA vs. NA	284 vs. 320 400 vs. 232	<.001**
1501-1800	.15 .93	.02* .38	AA vs. NA AA vs. ProjA	457 vs. 196 357 vs. 269	<.001**
1501-1800	.15 .18	.02* .15	NA vs. AA NA vs. ProjA	196 vs. 457 242 vs. 400	.004*
1501-1800	.93 .18	.38 .15	ProjA vs. AA ProjA vs. NA	269 vs. 357 400 vs. 242	<.001**
1601-1900	.14 .89	.057 .59	AA vs. NA AA vs. ProjA	439 vs. 218 334 vs. 282	<.001**
1601-1900	.14 .19	.057 .15	NA vs. AA NA vs. ProjA	218 vs. 439 241 vs. 410	.15
1601-1900	.89 .19	.59 .15	ProjA vs. AA ProjA vs. NA	282 vs. 334 410 vs. 241	<.001**
1701-2000	.18 .87	.08 .54	AA vs. NA AA vs. ProjA	428 vs. 232 339 vs. 290	<.001**
1701-2000	.18 .23	.08 .05*	NA vs. AA NA vs. ProjA	232 vs. 428 233 vs. 409	.68
1701-2000	.87 .23	.54 .05*	ProjA vs. AA ProjA vs. NA	290 vs. 339 409 vs. 233	<.001**

Taken together, these various measures suggest a small or moderate rise for Projection in the 1600-2000 window. This is due to the late reaction of 11 participants (out of 22). The individual plots for those participants evidence a rise or a high plateau in the 1000 – 2000ms temporal region after noun onset (for averaged grip force activation of the late reactors versus non-reactors, see Figure 8, and for a more detailed participant-by-participant depiction, see Supplementary Material).

[8] Bar plots for significance along time can be found in the Supplementary Material.

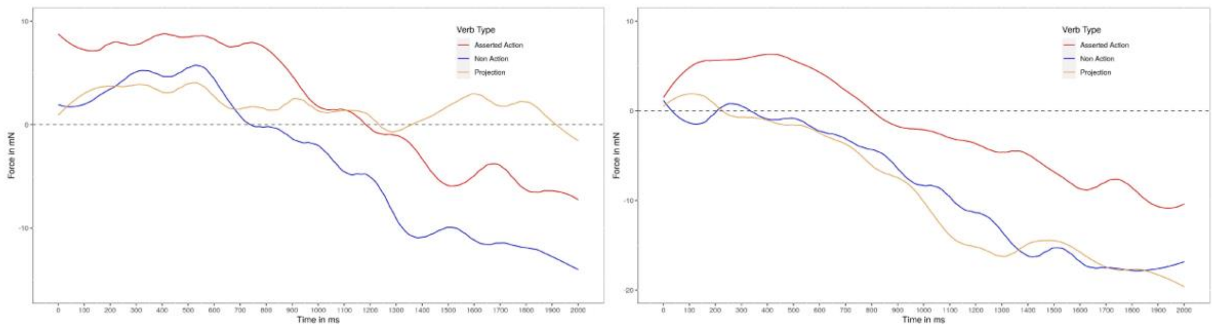


Figure 8. Plots for the Last Part of the Projection Time Point Series: 11 Participants with a final rise (left), Other 11 Participants (right)

The significant results confirm the visual observations of Figure 8. For the subset of participants without any final rise/plateau, the only mixed model values which reach significance concern the contrast between Projection and Asserted Action in the 900 – 1500 time windows. These results correspond to the trough of the Projection curve in Figure 8 (lower). For the subset of participants with a final rise/plateau, we have the results in Table 11. The mixed model and Wilcoxon *p* values indicate that Projection departs from Non-Action. The Fishers tests provide in general significant *p* values. The lowest values are those of the contrast between Asserted Action vs. Projection and Asserted Action vs. Non-Action. This is due to the fact that Asserted Action occupies an intermediate position, see Figure 8 (left) and the figures in the Count scores column of Table 11. So, the Fisher exact test reflects the symmetric status of Asserted Action, superior to Non-Action but inferior to Projection, even though none of these differences is significant[9].

Table 11

*P Values for the Mixed Model, Wilcoxon test, and Fisher Exact Test of the Last 2000 ms of Experiment 3 restricted to Participants Showing a Final Rise/Plateau * p < .05, ** p < .001*

Windows	Mixed model	Wilcoxon test	Contrasts	Count scores	Fisher
1301-1600	.37 .3	.32 .08	AA vs. NA AA vs. ProjA	210 vs. 105 113 vs. 178	<.001**
1301-1600	.37 .07	.32 .02*	NA vs. AA NA vs. ProjA	105 vs. 210 81 vs. 229	.05*
1301-1600	.3 .07	.08 .02*	ProjA vs. AA ProjA vs. NA	113 vs. 178 229 vs. 81	<.001*
1401-1700	.4 .2	.09 .1	AA vs. NA AA vs. ProjA	221 vs. 97 97 vs. 192	<001**
1401-1700	.4 .06	.09 .03*	NA vs. AA NA vs. ProjA	97 vs. 221 75 vs. 231	.1
1401-1700	.2 .06	.2 .03*	ProjA vs. AA ProjA vs. NA	192 vs. 97 231vs. 75	.018*
1501-1800	.4 .2	.3 .09	AA vs. NA AA vs. ProjA	216 vs. 97 122 vs. 178	<.001**

[9] Bar plots for significance along time can be found in the Supplementary Material.

1501-1800	.4 .047*	.3 .02*	NA vs. AA NA vs. ProjA	97 vs. 216 81 vs. 233	.15
1501-1800	.2 .047*	.09 .02*	ProjA vs. AA ProjA vs. NA	178 vs. 122 233 vs. 81	<.001**
1601-1900	.4 .24	.15 .15	AA vs. NA AA vs. ProjA	213 vs. 112 102 vs. 198	<.001**
1601-1900	.4 .043*	.15 .02*	NA vs. AA NA vs. ProjA	112 vs. 213 80 vs. 236	.01*
1601-1900	.24 .043*	.15 .02*	ProjA vs. AA ProjA vs. NA	198 vs. 102 236 vs. 80	.02*
1701-2000	.43 .26	.1 .2	AA vs. NA AA vs. ProjA	207 vs. 116 101 vs. 202	<.001**
1701-2000	.43 .052*	.1 .02*	NA vs. AA NA vs. ProjA	116 vs. 207 83 vs. 226	.01*
1701-2000	.26 .052*	.2 .02*	ProjA vs. AA ProjA vs. NA	202 vs. 101 226 vs. 83	.09

Discussion

The main finding of our last experiment is that, in contrast to the other two experiments on factive and non-factive constructions, grip force activation is not uniform in the case of projection. Projection differs significantly from Asserted Action in the first 1600 ms after verb onset, whereas no significant difference is observed to non-action verbs in this time window. The situation is different in the last time window (2000 ms after noun onset), where we observe that half of the participants show a rise or relatively high plateau for Projection.

The absence of a grip force activation in the initial window could a priori be attributed to at least two possible scenarios. First, the negation of the factive verb may have tainted a grip force activation of the true presupposed content. Such an explanation would be in line with research that suggests that negation does not give rise to a motor representation (e.g. Aravena et al., 2012; de Vega et al., 2016; Papeo et al., 2016; Tettamanti et al., 2008). Second, it is also possible, as has been argued by Beaver et al. (2017) and Simons et al. (2017) that, in some cases, a projective interpretation of the presupposition of a negated factive verb construction does not arise.

Given that some grip force activation occurs in a later time window, it is highly unlikely that negation affects the grip force activation of projective actions in such a way that it completely blocks the motor semantic representation. In addition, considering the results of the previous two experiments on factive and non-factive constructions, it is more plausible that the projective interpretation of the presupposition of a negated factive verb constructions in decontextualized sentences is less uniform, delayed and/or weaker than the factive interpretation (Experiment 1). As we will see in the next section, this is consonant with certain empirical and experimental observations about projection.

4. GENERAL DISCUSSION

Using the grip force sensor technique (Aravena et al. 2012, 2014; Frak et al. 2010; Nazir et al. 2017), the present study is the first – to our knowledge – to investigate the involvement of the sensori-motor system in coded information layering. In Experiment 1, we compared asserted information with information embedded under a presuppositional factive verb construction. In Experiment 2, we extended our investigation to a non-factive verb construction. Lastly, we examined whether the projection behavior of a factive verb construction modulates sensori-motor activation under negation (Experiment 3). Our results indicate the following:

1. The presupposed factive complement triggers an increase in grip force. The presupposed content of factive verb constructions elicits a significantly higher grip force response than Non-Action verbs. The grip force response between the asserted and presupposed content does not differ significantly.
2. The grip force activations recorded for the Non-Presupposed Action verb are not significantly different from those of the Non-Action condition. It is important to note that the *p* values for the mixed models and the Fisher tests are compatible with a more nuanced hypothesis, namely that Non-Presupposed Action is slightly more susceptible to motor response than Non-Action.
3. Our results show that the Asserted-Action condition shows a significant increase in grip force when compared to the Non-Action or Non-Presupposed Action conditions, whereas the difference between the Non-Presupposed Action and Non-Action conditions is not significant. This contrast suggests that Non-Presupposed Action verbs (Experiment 2) and Presupposed Action verbs (Experiment 1) trigger different grip force activations.
4. When the factive verb is negated, the construction does not elicit a grip force response in the reference window (roughly, the first 1600 ms after verb onset). The grip force response of the Projection verb of negated factive verb constructions differs significantly from Asserted Action verbs but not from Non-Action verbs. In the late window before the next stimulus, that is 2000 ms after noun onset, a

small/moderate positive deviation is observed. A more fine-grained analysis confirms this tendency for half of the participants.

With sensori-motor activation as criterion, in Experiment 1 we tested the assumption that the backgrounded status of the factive complement engages motor brain structures differently compared to asserted content. This assumption was not confirmed. It is worth noting here that, based on descriptive linguistic analyses, Beaver (2010) and Simons et al. (2017) recently challenged the backgrounded status of factive complements altogether. Moreover, the corpus analysis by Spenader (2002) indicates that in more than fifty percent of the cases, the factive complement is introduced as new information. In this line, it is thus not such a surprise that the factive complement also triggers a sensori-motor response. It is important to note that the activation differs significantly from the Non-Action condition. Combined with the result of our second experiment, which showed that the complement of a non-factive verb construction does not trigger an activation, it appears that the sensori-motor activation is modulated by the *truth-conditional status* of the action verb, not by the 'novelty' of the conveyed information. The difference between factives and non-factives (*know* vs *believe*) is not the fact that the complement clause describes some novel event or not, but the fact that the complement is presented as true or not. In this respect, it should be noted that the factive complement was not accented in the auditory material, which minimizes the possibility that this information represented the focus of the sentence. A follow-up study should investigate whether a focus manipulation, that is accenting the asserted content when simultaneously de-accenting the presupposed content affects sensori-motor correlates of the presupposed action. In conclusion, our findings extend the current knowledge about the contextual factors that modulate sensori-motor activity and demonstrate once more that language induced sensori-motor activation depends, in subtle ways, on contextual manipulations of lexical and discourse properties (e.g. van Ackeren et al. 2012; 2016; Egorova et al., 2014; 2016).

Moreover, the results of Experiment 1 and 2 are also relevant to the classic problem of *compositionality*, that is, the idea that the meaning of a sentence is a function of its grammatical structure and the meaning of its parts (e.g. Janssen, 1997). This discussion is often centered on the question of whether the meaning of single words is computed first and then combined into a global interpretation, or whether a global interpretation is derived immediately or at some intermediate stage (see Degen, 2013 for a discussion of implicatures). Our results provide evidence against an account that considers that the dominant factor of motor activation is the lexical content of the action verb because the critical action verb does not provoke a grip force response in all conditions (e.g. no grip force increase in the non-presupposed complement in Experiment 2 nor in the first 1600 ms for the negated factive verb constructions of Experiment 3). In this respect, the first two experiments confirm the sensitivity of the grip force response to the construction of a plausible situation model based on the representation of events and all three experiments confirm the crucial impact of linguistic constructions on the motor response.

Regarding the symmetric findings in Experiments 1 and 3, it could be argued that negation blocked or delayed a possible motor representation in Experiment 3. Given that (i) the results of our pre-test of the third experiment indicate that the factive complement was considered as true in 84.2% of all questions, (ii) in the experimental material, we took care of adding an introductory clause facilitating projection and (iii) we observed a late activation of motor response in Experiment 3, we can safely assume that negation does not just suppress any representation of the event as true in the event model. In other terms, in a micro-text like *Robert is busy in the drawing-room, he does not see that Ghislaine is tying her shoes*, the negation of the second sentence can hardly be considered as *preventing* hearers to derive the proposition that Ghislaine is indeed tying her shoes and adding it to the current event model.

On the other hand, in view of the difference between Experiments 1 and 3, there is no question that negation affects the motor response. But *how*? Simons et al. (2017) and Beaver et al. (2017) have recently put forward a framework that challenges the conventional view of projection. According to the conventional view (e.g. Gazdar, 1979a, b; Heim, 1983, 1992), presuppositional behavior is considered as context independent, that is, it does not systematically interact with contextually available information. In this line, factive complements *always* project, irrespective of the presence of entailment-canceling operators, such as negation and interrogation, or of different contexts. In contrast, Beaver et al. (2017) clearly show that the projective readings of factive complements can be contextually suppressed as illustrated by two of their examples as in (15) and (16), where the critical sentence is underlined. In example (15), the presupposed content referred to by *that* (i.e. the proposition that the newer designs being proposed are much safer) projects since A does not contradict B, whereas the same presupposition by A in (16) does not project. Beaver et al. claim that in cases where the presupposition is not under discussion as in (15), the content projects, whereas non-projective interpretations arise when the speaker is not committed to the truth of the complement, as in (16).

(15) A: People are worried. We have a major nuclear event going on in Japan, and it's far too early to claim that things are under control.
B: Well, again, these are older designs. The newer designs being proposed are much safer.
A: Our citizens don't know that, so they remain concerned. More has to be done to educate and reassure them.

(16) A: We have a major nuclear event going on in Japan, and it's far too early to claim that things are under control.
B: Well, again, these are older designs. The government assures us that the newer designs being proposed are much safer.
A: They don't know that. These were claimed to be the same—actually, the AP1000 that you were talking about building down in Vogtle, there are concerns right now about how well the containment will work.

According to Beaver et al.'s account, the projection criterion is not conventionally encoded per se, but interacts with the speaker's commitment to the truth of utterance. Adapting this framework to our results, if only the perspective of the speaker was taken into account, assuming that the speaker is judged as trustworthy and reliable, then a grip force activation should have also been observed in the projection experiment. However, given our decontextualized sentences, the speaker's perspective may not be the only one responsible for the recruitment of the motor system. Our results suggest that besides the speaker's perspective, the perspective of the agent may also be considered during the on-line recruitment of motor structures. When the speaker and the agent are committed to the truth of the utterance, then a grip force activation is observed as it is the case with factive verbs. If one of the truth conditions is not fulfilled, that is if either the speaker or the agent is not aware of the truth of the described event, then a different pattern of activation is observed as it is the case in Experiment 2 and Experiment 3.

In summary, our results indicate that the perspective of the agent also has an effect and so, that the status of the event in the representation of the agent is part of the interpretation process. What remains to be determined is whether the very moderate activation observed for belief verbs and the initial absence for projection is only triggered by the presence of a hand-related action verb, i.e. an effect of the lexicon or rather to the fact that there is at least one point of view in which the motion event takes place, the perspective of the agent for belief verbs and the perspective of the speaker for projection. In the current experiment, the trustworthiness of the speaker has not yet been manipulated. Future research could fill this vacuum by manipulating the reliability of the speaker for the projective action condition. When the presupposed content is considered as true, then a projective reading should arise, which, as a consequence should trigger a grip force activation.

Thanks to the on-line nature of the grip force measure, the three experiments broaden our understanding of which linguistic environments elicit a grip force activation. More specifically, the use of this on-line measure allows to enhance our understanding of which linguistic environments recruit motor brain structures. In addition, it also provides new insights, which are not captured using an off-line measure as our results on the pre-test of the third experiment reveal.

Conclusion and Future Directions

Over the last decades, the question of the role of the sensorimotor system in meaning representation has been vigorously debated by philosophers and neuroscientists. In a recent review, Meteyard et al. (2012) places the answers to this question on a continuum ranging from strong embodied positions (e.g. Gallese & Lakoff, 2005; Glenberg & Kaschak, 2003) to disembodied accounts (e.g. Mahon & Caramazza, 2008).

On the one hand, strong embodied accounts maintain the existence of a close link between linguistic meaning and sensorimotor structures and suggest that language processing depends on the recruitment of distributed networks of sensorimotor structures. On the other hand, disembodied accounts defend the independence of linguistic meaning from sensorimotor structures by arguing that their recruitment is no evidence of an explanatory and causal link between language processing and sensorimotor structures.

There has been ample evidence that the truth may lie between these two opposite positions. A review by Willems and Casasanto (2011) points out that language-induced motor recruitment appears to be highly flexible and is moderated by situational context, be it linguistic or extra-linguistic (with regard to the linguistic context see, e.g., Aravena et al., 2012; 2014; Tettamanti et al., 2008, van Ackeren et al., 2012; with regard to the extra-linguistic context, see, e.g., Hoenig et al, 2008). In our studies, we investigated whether the factivity of a complement clause modulates the recruitment of sensorimotor areas. Our findings suggest that it does: factive action-related complements trigger a grip force activation whereas non-factive complements elicit a weaker response. This challenges the claim that action word meaning automatically recruit motor semantic features and that sensorimotor processing is necessary for conceptual or language processing. Our results are thus compatible with an account that assumes context dependency of language-induced motor activity. The activation of relevant action schemas, recruiting the same neural mechanisms as those active in overt behavior, selectively contributes to meaning representation as a function of the role the action plays in the overall discourse representation.

The current studies open up interesting directions for future research. While they demonstrate the selective involvement of the motor brain in the processing of hand-related action verbs, they leave open the question of what the role of such an activation is. Recent studies by Miller, Brookie, Wales, Wallace and Kaup (2018) investigated the sensitivity of ERP measures to hand and foot movements, as well as hand- and foot-associated words. While they consistently found ERP differences for hand versus foot movements, they showed no evidence of a difference for hand- versus foot-associated words. The grip-force method has the potential to further contribute to this line of research by investigating the extent to which the sensorimotor activation it captures is univocally linked to the processing of hand-related semantic meanings (rather than reflecting a more general motor activation).

The implications of the studies presented in this paper go beyond the debate on embodiment, and directly address questions that are relevant to linguistics theories. In the present experiments, we focused on the distinction between factive (*know*) and non-factive verbs (*believe*). By contrasting these latter two conditions our data suggest that the truth-conditional status of a clause (as determined by a factive verb) is a precondition for the recruitment of motor structures in language processes. These findings thus support a linguistic theoretical frame that considers the speaker's commitment to the truth of presupposed information as a central property of presuppositions (2016), but they are also compatible with the idea that the agent's perspective has some impact during the recruitment of motor brain structures (Experiment 2).

At this stage, the question naturally arises whether our observations are an effect of the particular structure of factive constructions or whether the conclusions they suggest extend to other presuppositional constructions. Recall that we chose factive constructions as a starting point for our experimental investigation because in factive constructions the asserted content and the presupposition are expressed explicitly. This is not the case with other constructions. The present investigations must thus be extended to other presupposition triggers and we will briefly discuss some reasonable follow-ups in this direction. One important issue is that, with factives, the action-related verb occurs only in the presupposed part (the clausal complement), which prevents any direct comparison between asserted content and presupposed content in terms of motor response. As a result, what we have shown is that, when there is an action-related verb, the fact that it occurs in the presupposed part does not block or weaken the motor response. But what happens if asserted content and presupposed content have an opposite motor polarity (action vs. no action)?

Change-of-state verbs like *begin* or *stop* illustrate precisely this point. They assert the most recent event and presuppose a less recent state of affairs with an opposite polarity. For instance, *Paul stops ironing his shirt* asserts that Paul does not iron his shirt and presupposes that he has been doing so before. If the situation model contains all events referred to by the sentence, irrespective of their recency, it is possible that the two events (ironing vs. not ironing) cancel out and that no significant motor response is recorded. If the event of not-ironing is more salient, one would predict no motor response, and, correlatively, a motor response for *Paul begins ironing his shirt*.

While change-of-state verbs are an interesting empirical family because they combine layers of information with opposite polarities, they are not the only ones with distinct presuppositional patterns within the heterogeneous class of presupposition triggers. Another major issue is the role of *focus*[10], that is, this part of the sentence information which might be taken to address a question. For instance, with clefts such as *It's Paul who irons his shirt*, the presupposition is that someone irons a shirt and the asserted content is that it is Paul who does that. The sentence is most naturally viewed as a possible answer to a question like *Who irons his shirt*? So, the focus is on *Paul,* not on the presupposition, and one may wonder whether there is some effect on the motor response. If the latter is not significant, this would probably indicate that our hypothesis that the truth-conditional status of an action-related event is sufficient to trigger a motor response has to be amended. A similar question arises for *exclusives* (*only*, *just*), whose focus structure is a matter of debate (Beaver & Clark, 2008). So, more work is needed to construct a more complete picture of the relations between motor response and coded semantic layering.

Acknowledgments

[10] The term *focus* is understood here as in linguistics (a sensible answer to a potential question) and is not to be confused with the meaning it has in Zwaan and Taylor's (2008) paper, mentioned in the introduction.

We thank Loriana Rizzo, Emilie Ganes, Hélène Boutaric, Caroline Masson and Julie Damour for assistance with stimuli preparation and stimuli recordings, and Anne Cheylus for programming the counterbalanced measures design and providing support for the analysis tool. We also thank Diana Mazzarella for comments on the first manuscript.

This research was supported by the French National Agency for Research (CogHuLiCe ANR 16-CE28-0004) attributed to TN and SC, by a DAAD scholarship to RR, and by the French National Center for Scientific Research (CNRS).

Context Paragraph

The present work brings together two different lines of research that are central to the work of the authors. On the one hand, the linguistic and philosophical analysis of the phenomenon of presupposition. On the other hand, the neurobiological investigation of language processing, with an emphasis on motor brain structures. These two lines of research converge around the issue of the context-sensitivity of language-induced sensori-motor activation. The question addressed by this paper is whether presuppositional contexts affect the activation of motor brain structures. The methodology employed is the grip-force method, which has been extensively developed at the *Institut des Sciences Cognitives Marc Jeannerod* in the last five years.

References

Aghabozorgi, S., Shirkhorshidi, A. S. & Wah, T. Y. (2015). Time-series clustering – A decade review. *Information Systems*, 53, 16-38.

Aravena, P., Courson, M., Frak, V., Cheylus, A., Paulignan, Y., Deprez, V., & Nazir, T. (2014). Action relevance in linguistic context drives word-induced motor activity. *Frontiers in Human Neuroscience*, *8*, 163.

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., & Nazir, T. (2012). Grip force reveals the context sensitivity of language-induced motor activity during "action words" processing: evidence from sentential negation. *PLoS One*, *7*(12), e50287.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., and Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Curr. Biol.* 16, 1818–1823.

Arnau, J., Bono, R., Blanca M. J. & Bendayan, R. (2012). Using the linear mixed model to analyze nonnormal data distributions in longitudinal designs. *Behavior Research Methods*, *44*, 1224-1238.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. Available from arXiv:1506.04967.

Bates, D., Maechler, M., Bolker, B., & Walker, S.(2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48.

Beaver, D. I. (2001). *Presupposition and assertion in dynamic semantics* (Vol. 29). Stanford: CSLI publications.

Beaver, D. I. & Clark, B. Z. (2008). *Sense and Sensitivity*. *How Focus Determines Meaning.* Chichester: Wiley-Blackwell.

Beaver, D. I., Roberts, C., Simons, M., & Tonhauser, J. (2017). Questions Under Discussion: Where information structure meets projective content. *Annual Review of Linguistics*, *3*, 265–284.

Boersma, Paul & Weenink, David (2019). Praat: doing phonetics by computer [Computer program]. Version Praat 6.0.49, retrieved 19 January 2019 from http://www.praat.org/

Boulenger, V., Hauk, O., & Pulvermüller, F. (2008). Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral cortex*, *19*(8), 1905-1914.

Boulenger, V., Shtyrov, Y., & Pulvermüller, F. (2012). When do you grasp the idea? MEG evidence for instantaneous idiom understanding. *Neuroimage*, *59*(4), 3502-3513.

Brauer, M. & Curtin, J.J. (2018). Linear mixed-effects models and the analysis of nonindependent Data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389-411.

Cacciari, C., Bolognini, N., Senna, I., Pellicciari, M. C., Miniussi, C., & Papagno, C. (2011). Literal, fictive and metaphorical motion sentences preserve the motion component of the verb: a TMS study. *Brain and language*, *119*(3), 149-157.

Chemla, E., & Bott, L. (2013). Processing presuppositions: Dynamic semantics vs pragmatic enrichment. *Language and cognitive processes*, *28*(3), 241-260.

Chierchia, G., & McConnell-Ginet, S. (1990). *Meaning and Grammar: An Introduction to Semantics*. Cambridge (MA): MIT Press.

Claus, B. (2015). Verb gapping: An action-gap compatibility study. *Acta Psychologica*, *156*, 104–113.

Degen, J. (2013). *Alternatives in Pragmatic Reasoning*. Ph.D. dissertation, University of Rochester.

Desai, R. H., Conant, L. L., Binder, J. R., Park, H., & Seidenberg, M. S. (2013). A piece of the action: modulation of sensory-motor regions by action idioms and metaphors. *NeuroImage*, *83*, 862-869.

de Vega, M., Morera, Y., León, I., Beltrán, D., Casado, P., & Martín-Loeches, M. (2016). Sentential negation might share neurophysiological mechanisms with action inhibition. Evidence from frontal theta rhythm. *Journal of Neuroscience*, *36*(22), 6002-6010.

Domaneschi, F. (2016). *Presuppositions and Cognitive Processes: Understanding the Information Taken for Granted*. Springer.

Domaneschi, F., Canal, P., Masia, V., Vallauri, E. L., & Bambini, V. (2018). N400 and P600 modulation in presupposition accommodation: The effect of different trigger types. *Journal of Neurolinguistics*, *45*, 13-35.

Domaneschi, F., & Di Paola, S. (2017). The processing costs of presupposition accommodation. *Journal of Psycholinguistic Research*, 1-21.

Ducrot, O. (1972). *Dire et ne pas Dire*. Paris: Hermann.

Dudley, R., Orita, N., Hacquard, V., & Lidz, J. (2015). Three-year-olds' understanding of know and think. In *Experimental perspectives on presuppositions* (pp. 241-262). Springer, Cham.

Eager, C., & Roy, J. (2017). Mixed effects are sometimes terrible. Linguistics Society of America 2017 Annual Meeting. Austin, Tx (January 5-8, 2017). Available as: *https://arxiv.org/pdf/1701.04858.*

Egorova, N., Pulvermüller, F., & Shtyrov, Y. (2014). Neural dynamics of speech act comprehension: an MEG study of naming and requesting. Brain Topography, 27(3), 375-392.

Egorova, N., Shtyrov, Y., & Pulvermüller, F. (2016). Brain basis of communicative actions in language. NeuroImage, 125, 857-867.

Egré, P. (2008). Question-embedding and factivity. In Lihoreau, F. (Ed.), *Knowledge and Questions*, Grazer Philosophische Studien, vol. 77, Amsterdam: Rodopi, 85-125.

Field, A. P. & Wilcox R. R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy*, *98*, 19-38.

Frak, V., Nazir, T., Goyette, M., Cohen, H., & Jeannerod, M. (2010). Grip force is part of the semantic representation of manual action verbs. *PLoS One*, *5*(3), e9728.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, *6*(2), 78-84.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology*, *22*(3-4), 455-479.

Galwey, N. W. (2006). *Introduction to Mixed Modelling. Beyond Regression and Analysis of Variance.* Wiley.

Gazdar, G. (1979). A solution to the projection problem. *Syntax and semantics*, *11*, 57-89.

Geurts, B. (1999). *Presuppositions and Pronouns*. Elsevier.

Gibbs, R. W. (2006). Metaphor interpretation as embodied simulation. *Mind & Language*, *21*(3), 434-458.

Gibbs R. W., Lima, P. L. C., & Francozo, E. (2004). Metaphor is grounded in embodied experience. *Journal of pragmatics*, *36*(7), 1189-1210.

Grice, P. (1975). Logic and conversation. In Davidson, D. and Harman, G. (Eds), *The Logic of Grammar*, 64–75, Encino (CA): Dickenson.

Guijo-Rubio, D. Durán-Rosal, A. M., Gutiérrez, P. A., Troncoso, A. & Hervás-Martínez, C. (2020). Time series clustering based on the characterisation of segment typologies. *IEEE Transactions on Cybernetics*, 1-14.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301-307.

Heim, I. (1983). On the projection problem for presuppositions, *WCCFL 2*, 114-125.

Heim, I. (1992). Presupposition projection and the semantics of attitude verb. *Journal of Semantics*, *9*, 183-221.

Hosmer, D. W., Lemeshow, S. & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.

Janssen, T. M. (1997). Compositionality, with an appendix by B. Partee. Handbook of logic and linguistics, ed. by JFAK van Benthem and A. ter Meulen.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain sciences*, *17*(2), 187-202.

Huette, S. (2016). Putting context into context: sources of context and a proposed mechanism for linguistic negation. *Language, Cognition and Neuroscience*, *31*(8), 1000-1014.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *cortex*, *48*(7), 805-825.

Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical linguistics*, *1*(1-3), 181-194.

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *Quarterly journal of experimental psychology*, *60*(7), 976-990.

Kiparsky, P. & Kiparsky, C. (1970). Fact. In M. Bierwisch & K. Heidolph (Eds.), *Progress in Linguistics*, The Hague: Mouton, 143-173.

Kuznetsova A. , Brockhoff P. B. & Christensen R. H. B. (2017). lmerTest package: Tests in linear mixed-effects models. *Journal of Statistical Software*, *82*(13), 1-26.

Lauro, L. J. R., Mattavelli, G., Papagno, C., & Tettamanti, M. (2013). She runs, the road runs, my mind runs, bad blood runs between us: Literal and figurative motion verbs: An fMRI study. *NeuroImage*, *83*, 361-371.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of philosophical logic*, *8*(1), 339-359.

Masia, V., Canal, P., Ricci, I., Vallauri, E. L., & Bambini, V. (2017). Presupposition of new information as a pragmatic garden path: Evidence from Event-Related Brain Potentials. *Journal of Neurolinguistics*, *42*, 31-48.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H. & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305-315.

Moody, C. L., & Gennari, S. P. (2010). Effects of implied physical effort in sensory-motor and pre-frontal cortex during language comprehension. *Neuroimage*, *49*(1), 782-793.

Nagel, J. (2017). Factive and nonfactive mental state attribution. *Mind & Language*, *32*(5), 525-544.

Nazir, T. A., Hrycyk, L., Moreau, Q., Frak, V., Cheylus, A., Ott, L., … & Delevoye-Turrell, Y. (2017). A simple technique to study embodied language processes: the grip force sensor. *Behavior research methods*, *49*(1), 61-73.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary french: LEXIQUE™. *L'année psychologique*, *101*(3), 447-462.

Oldfield, R.C. (1971). *The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia, 9(1), 97-113.*

Peters, S. (2016). Speaker commitments: Presupposition. In *Semantics and Linguistic Theory* (Vol. 26, pp. 1083-1098).

Papeo, L., Hochmann, J. R., & Battelli, L. (2016). The default computation of negated meanings. *Journal of cognitive neuroscience*, *28*(12), 1980-1986.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576.

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*(3), 793-797.

Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, *47*(2), 388-396.

R Core Team (2019). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rao, T. S., Rao, S. S. & Rao, C. R. (Eds.). (2012). *Handbook of Statistics* (Vol. 30). *Time Series Analysis: Methods and Applications.* Elsevier.

Reboul, A. (2017). Is implicit communication a way to escape epistemic vigilance?. *Pragmatics at its Interfaces*, *17*, 91.

Romoli, J., & Schwarz, F. (2015). An experimental comparison between presuppositions and indirect scalar implicatures. In *Experimental perspectives on presuppositions* (pp. 215-240). Springer, Cham.

Rossini, P. M., Rossi, S., Pasqualetti, P., & Tecchio, F. (1999). Corticospinal excitability modulation to hand muscles during movement imagery. *Cerebral cortex*, *9*(2), 161-167.

Sardá-Espinosa, A. (2019). dtwclust: Time series clustering along with optimizations for the Dynamic Time Warping distance. R package version 5.5.6. https://CRAN.R-project.org/package=dtwclust

Schlenker, P. (2008). Be articulate: A pragmatic theory of presupposition projection. *Theoretical Linguistics*, *34*(3), 157-212.

Schwarz, F. (2014, August). Presuppositions are fast, whether hard or soft-evidence from the visual world. In *Semantics and Linguistic Theory* (Vol. 24, pp. 1-22).

Schwarz, F. (2015). Introduction: Presuppositions in Context—Theoretical Issues and Experimental Perspectives. In *Experimental perspectives on presuppositions* (pp. 1-37). Springer, Cham.

Schwarz, F. (2016). Experimental work in presupposition and presupposition projection. *Annual review of linguistics*, *2*, 273-292.

Schwarz, F., & Tiemann, S. (2012). Presupposition Processing–The Case of German wieder. In *Logic, Language and Meaning* (pp. 200-209). Springer, Berlin, Heidelberg.

Simons, M., Beaver, D., Roberts, C., & Tonhauser, J. (2017). The best question: Explaining the projection behavior of factives. *Discourse Processes*, *54*(3), 187-206.

Singmann, H & Kelen, D. (2020). An introduction to mixed models for experimental psychology. In D. Spieler & E. Schumacher (Eds.), *New Methods in Cognitive Psychology* (pp. 4-31). Routledge.

Spenader, J. (2001, August). Presupposition or Abstract Object Anaphora?: Constraints on Choice of Factive Complements in Spoken Discourse. In *Information Structure, Discourse Structure AND Discourse Semantics Workshop Proceedings* (p. 161).

Stalnaker, R. (1974). Pragmatic presuppositions. Semantics and philosophy, ed. by Milton K. Munitz and Peter Unger, 197–213.

Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, *25*(5-6), 701-721.

Stegmann, G., Jacobucci, R., Harring, J. R., & Grimm, K. J. (2017). Non-linear mixed-effects modeling programs in R. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(1), 160-165.

Stroup, W. W. (2012). *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications*. CRC Press, Boca Raton.

Tavenard, R. Faouzi, J. & Vandewiele, G. (2017). tslearn: A machine learning toolkit dedicated to time-series data. https://github. com/rtavenar/tslearn.

Taylor, L. J., & Zwaan, R. A. (2008). Motor resonance and linguistic focus. *Quarterly Journal of Experimental Psychology*, *61*(6), 896-904.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., ... & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of cognitive neuroscience*, *17*(2), 273-281.

Tettamanti, M., Manenti, R., Della Rosa, P. A., Falini, A., Perani, D., Cappa, S. F., & Moro, A. (2008). Negation in the brain: Modulating action representations. *Neuroimage*, *43*(2), 358-367.

Tian, Y., Ferguson, H., & Breheny, R. (2016). Processing negation without context–why and when we represent the positive argument. *Language, Cognition and Neuroscience*, *31*(5), 683-698.

Tiemann, S., Schmid, M., Bade, N., Rolke, B., Hertrich, I., Ackermann, H., ... & Beck, S. (2011). Psycholinguistic evidence for presuppositions: On-line and off-line data. In *Proceedings of sinn & bedeutung* (Vol. 15, pp. 581-597). Saarland University Press Saarbrücken.

van Ackeren, M. J., Casasanto, D., Bekkering, H., Hagoort, P., & Rueschemeyer, S. A. (2012). Pragmatics in action: indirect requests engage theory of mind areas and the cortical motor network. *Journal of Cognitive Neuroscience*, *24*(11), 2237-2247.

van Ackeren, M. J., Smaragdi, A., & Rueschemeyer, S. A. (2016). Neuronal interactions between mentalising and action systems during indirect request processing. *Social cognitive and affective neuroscience*, *11*(9), 1402-1410.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer Verlag.

Wickham H. (2019). forcats: tools for working with categorical variables (factors). R package version 0.4.0 https://CRAN.R-project.org/package=forcats.

Willems, R. M., & Casasanto, D. (2011). Flexibility in embodied language understanding. *Frontiers in Psychology*, *2*, 116.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, *123*(2), 162.

Zwaan, R. A., Taylor, L. J., & De Boer, M. (2010). Motor resonance as a function of narrative time: Further tests of the linguistic focus hypothesis. *Brain and Language*, *112*(3), 143-149.

APPENDIX

Detailed Explanation concerning the analysis

In this paper, we have used linear mixed-effects models as our starting point. This raises two questions. First, in some cases the default algorithm does not converge. This might be taken as an indication that the maximal model is not appropriate, being for instance too complex in regard of the number of observations (Eager & Roy, 2017). However, the availability of the *allFit* function in the *lme4* R package (Bates et al., 2015) allows one to check the results of a series of alternative algorithms. We observed that most of them converge and, more importantly, end up with practically identical *t* and *p* values (including the values calculated by the initial non-converging algorithm). This suggests that the complexity of the model is not problematic and, accordingly, we kept the random maximal structure, using the mean of all *p* values for out final estimation.

The second question is much more problematic, the Q-Q plots and Shapiro-Wilks tests on the residuals of the various models show that these residuals are not normal (see for instance Figure X, which shows the Q-Q plot for the contrast between the Asserted Action (like *Ines ties her shoes*) and Non-Action (like *Peter prefers chicken*) stimuli.

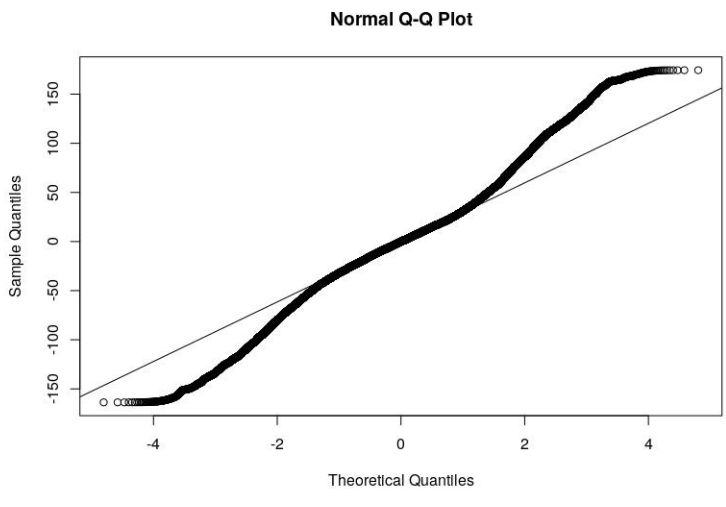


Figure 9. Q-Q Plot for the Asserted Action vs. Non-Action 601-900 ms Window in Experiment 3

No transformation of the response was found to have a positive effect on the non-normality of residuals. It is sometimes assumed that mixed-effects models are robust to deviations from normality[11]. However, recent literature points out that this may be a serious problem[12], which calls for specific solutions (for some examples, see Arnau, J., Bono, R., Blanca, M.J. & Bendayan, R., 2012; Field & Wilcox, 2017). The two traditional strategies are: (1) Considering other types of *parametric* models and (2) abandoning parametricity altogether. Identifying the appropriate type of parametric model, if any, seems extremely difficult. The high participant and item-based variability is reflected in plots which do not correspond to any standard probability distribution. Moreover, *generalized* binomial linear models (known as *logistic* regression models) fit by transforming the grip force intensity values into binary (1 vs 0) values gave poor results. We binarized the results in several ways by counting as 1 (respectively 0) the intensities above (respectively equal to or below) the mean, the median, or various given numeric thresholds. In each and every case, the explanatory power of the corresponding models remained low, as evidenced by ROC curves. To give a concrete example, the logistic regression maximal mixed model[13] for the 601-900 temporal window of the third (projection) experiment delivers the following results for the Asserted Action vs non-action contrast based on the position with respect to the mean: $p = 0.044$, A(rea) U(nder) C(urve) = 0.65. The AUC represents the discriminating power of the logistic regression. It estimates the probability that the model will guess the correct value (0 for 0, 1 for 1). We followed the rule of thumb described (Hosmer, Lemeshow & Sturdivant, 2013, p. 177) and considered that the 0.65 value for the AUC is not enough to guarantee that the model has a good discriminating power.

Does it follow that the linear mixed-effects maximal models are devoid of interest? Two question must be distinguished. The first one is whether the figures (likelihood, t values, p values, etc.) calculated by the models are good estimates. In the absence of linearity, the answer is a clear *we cannot know*. The second question is whether the effect or non-effect that the model could suggest are 'real', that is, corresponds to

[11] Or homoscedasticity, for that matter. We are not concerned with homoscedasticity here because our results do not show any obvious sign of heteroscedasticity.

[12] The non-normality of residuals does not *entail* that a linear maximal model is invalid in a strong sense, that is, that it would distort the data to the point where the hierarchy between conditions is reversed. In general, the effect detected by such models is really there but its numerical estimation (in terms of value) is not reliable.

[13] We used the *glmer* function of the *lme4* R package with the same structure as in table 1. We note that models with fewer parameters (non-maximal models) make the discriminating power of the model *decrease*. So, there is no hope of getting better results by simplifying the models.

some underlying causal mechanism. Here, the answer is relatively simple: mixed-effect models have been devised to separate the statistical impact of particular objects (very often participants and items), whose internal variation is *random (*i.e. not the focus of an experiment), from that of the main factor(s) (sentence type in the current experiments). When such models are maximal, they exhaust the set of extraneous (random) factors and the effects (impact on mean) they report is what remains once the impact of random factors has been 'subtracted' as far as possible. More precisely, the impact of particular objects in the domain of random factors (e.g. particular participants and items) is adjusted.[14] This has the consequence that maximal mixed-effect models are *stricter* than their non-maximal alternatives. Their *t* an *p* values estimations are almost always inferior and frequently under the significance threshold chosen for the non-maximal variants. Whenever a maximal model reports no effect whereas a non-maximal model reports one, it is because the maximal model has detected a strong variation in one or several random objects and it is good practice to follow its lead and to conclude that one has no evidence of an effect. This extends to cases where simple techniques (for example mean or count comparisons) hint at an effect that the maximal model does not 'see'. In the case at hand, the difference between a standard RM ANOVA and a maximal model is not so dramatically clear cut. An RM ANOVA is in practice equivalent to a mixed-effect model with a random term for participants (intercept *and* slope) but not for items. The maximal models add a random intercept for items. They make the *p* value increase above 3.5 % on average. This is not sufficient to reverse the conclusions suggested by the ANOVA, even though it provides a different estimation of the effect. Accepting the ANOVA as a more approximate measure would then be perhaps reasonable, but this would not be the best choice when it comes to precision. The likelihood ratio tests show that maximal models fit the data much more tightly. To illustrate, in the 600-900 ms window of Experiment 1, the obtained *p* value for the comparison between Asserted Action and Non-Action is .015 for the ANOVA and .052 for the maximal model. But the likelihood ratio test delivers a value of 17969 for the maximal model against 7928 for the ANOVA[15], which shows that the former is much more sensitive to sources of variation in the data.

[14] See for instance Galwey (2006, chapter 5) for a relatively clear discussion of the adjustment process.
[15] We used the *ranova* function from the *R package lmerTest* (Kuznetsova, Brockhoff & Christensen, 2017).

To conclude on this point, we can and must use maximal models as safeguards against false positives and, more generally, to provide better correspondence to the data, unless we have more powerful non-parametric techniques, a point that we examine now.

Concerning non-parametricity, the two methods of choice are (i) applying learning algorithms and (ii) sticking to strictly non-parametric inferential tests. Under the learning perspective and given the high variability already mentioned, it is unlikely that the result of a learning process would resemble a standard mathematical function. The current state of the art in learning procedures rather orients us toward (deep) learning techniques for classification. In this perspective, a crucial point is to decide which type of data we have to analyze. Grip force intensity evolves over time and the observed curves are *time series*, in the sense of temporally ordered sequences of value-time point pairs. Time series have been the subject of substantial investigation in recent years (for a representative example, see Rao, Rao & Rao, 2012), especially in the field of financial modeling where trying to anticipate the future of financial products is a central concern. However, our main research question is the *comparison* of times series, a concept which is somewhat elusive (Aghabozorgi, Shirkhorshidi & Wah, 2015). Time series can have similar shapes but differ by their values. In that case one series is a vertical translation of the other. They can also have similar values and different shapes, for example if the shapes are approximately symmetrical (i.e. typically, two linear shapes with opposite slopes) or if the similar values are not temporally aligned (for example when a segment of one series is similar to a preceding or subsequent segment of the other series). Such aspects make the classification of irregular time series a difficult and largely exploratory enterprise. More technically, the main problem with grip force time series is to determine whether a segment-based approach (Guijo-Rubio, Durán-Rosal, Gutiérrez, Troncoso & Hervás-Martínez, 2020) is appropriate. There exist good software packages in python (*tslearn,* Tavenard, Faouzi & Vandewiele, 2017) and R (*dtwclust*, Sardá-Espinosa, 2019) exploiting iterative or deep learning algorithms for time series comparison. However, at the time being, we lack a robust evaluation of their relevance and efficiency for analyzing the type of data we deal with in this paper.

Non-parametric tests seem to be more promising in the short term. The idea we develop is to use Fisher exact tests and Wilcoxon tests on counts. We proceed in 4 steps.

1. For each 300 ms time window, each participant and each item in each condition, we record the mean of the item. We rank the items by means in ascending order.

2. For each 300 ms time window, each participant and each pair of conditions, for instance Asserted Action versus Non-Action, we compare the items with the same rank, meaning that for each time point (1-300) we note whether item A of rank R in the Asserted Action condition has a higher (lower, or equal) value than item B of rank R in the Non-Action condition. The item which has at least 60 higher values (20% of all the values) is determined to be the "winner" of the comparison[16]. In any other case the items are considered as equal ("tie"). We tested whether larger thresholds modified significance for the Fisher and Wilcoxon tests. We tried 70, 80, 90 and 100 (23%, 27%, 30% and 33%). This made the p values vary but they remained significant. At this stage, we obtain sequences of comparisons results, for each time window, participant and pair of conditions.

3. Summing by condition across participants gives us the total scores for each time window and each comparison of conditions. For instance, in Table 5 of Section 3.1 (the factive experiment), the first row is:

Table 11.
*P Values for the Mixed Model, Wilcoxon test, and Fisher Exact Test. * p < .05, ** p < .001*

Windows	Mixed model	Wilcoxon test	Contrasts	Count scores	Participant scores (Winners)	Fisher
501-800	.065	.04*	AA vs. NA	445 vs. 234	14 vs. 7, 4 ties	<.0001**
	.8	.15	AA vs. PresA	338 vs. 376	9 vs. 12, 4 ties	

The Count Scores column contains the sum of all winner items for each category. For instance the table tells us that, in the 500-800 window, Asserted Action items win 445 times and Non-Action items 234 times. The Fisher tests reported in the paper are based over count comparisons. In the above table, Asserted Action (AA) is successively compared to Non-Action (NA) and to Presupposed Action (PresA). It may not be apparent how this fits in with the condition of a Fisher test. The figures in the Count scores column are considered as winning *events*, not as items. We partition the initial set of winning events (1393 = 445 + 234 + 338 + 376 in total) into Asserted Action winning events (783 = 445 + 338) and other winning events (610 = 234 + 376). We again partition the set in those winning events involved in an Asserted Action vs. Non-

[16] In other terms, if L1 (respectively L2) is the number of values of item 1 (respectively item 2) which are superior to the corresponding values of item 2 (respectively item 1), the absolute value of the difference (L1 - L2) must be equal or superior to 60 for there to be a winner. Or, equivalently, the winner has 1.5 times more higher values than the loser.

Action competition (679 = 445 + 234) and those winning events involved in an Asserted Action vs. Presupposed Action competition (714 = 338 + 376). This gives us a Fisher contingency table about which we are interested in determining whether Asserted Action winning events occur more (less, equally) often in the first competition than in the second. This is precisely one of the things a Fisher exact test is able to tell us.

4. Finally, we can exploit another well-known non-parametric test, the Wilcoxon paired test. Consider a participant and a competition between conditions observed at the level of this particular participant in a specific time window. We count the number of winners for each condition for this participant. We do the same for all participants. This gives us two vectors. For instance the Asserted Action vs. Non-Action in the 500-800 window gives [13,16,6,16,21, ...] vs. [9,10,23,7,9,...]. The i-th value in the first (second) vector is the number of Asserted Action (Non-Action) winners. Running a Wilcoxon paired test on such vectors delivers a measure of how different are the two conditions in the chosen time window.

Summarizing, we have seen that standard models cannot offer robust numerical estimations but that the combination of such models with non-parametric simple methods can help us to interpret the observations and to detect the effects differences between conditions, even though, for technical reasons, we cannot express them reliably in the language of probability measures. The crucial point is the convergence or non-convergence of the different measures. Whenever the results of the linear mixed models and the count-based tests coincide, we have reasonable evidence that the presence or absence of an effect of the conditions is grounded in reality and not just an artifact of some strange statistical algorithm. It is then unnecessary to embark on a perilous − and probably endless − discussion about what the exact 'measure' of this effect could or should be. However, it happens frequently that the different indicators disagree. In that case, we give priority to the mixed-effects model since, as explained above, it is able to detect unbalanced variation in the random factors, that is, participant and item variables. Accordingly, our method in the paper was to examine the different windows of interest, selecting those which are not clearly too far from the usual significance threshold of .05 and investigate whether the count-based measures are consonant with the model estimation. When selecting windows in the first stage, we decided to select only those which corresponded to an increase or a plateau of the grip force intensity, leaving apart those regions where the intensity drops even when there is a significant difference between conditions. In our present state of knowledge, it is unclear whether such differences in decrease rate reveal something relevant to our concerns.