

Neological intuition in French: A study of formal novelty and lexical regularity as predictors

Alizée Lombard^{*}, Richard Huyghe, Pascal Gygax

University of Fribourg, Avenue de Beauregard 13, 1700 Fribourg, Switzerland

Received 8 May 2020; received in revised form 4 February 2021; accepted 5 February 2021

Available online 27 February 2021



Abstract

This paper investigates how neological intuition (NI), i.e., the metalinguistic ability to evaluate lexical novelty, is influenced by the linguistic properties of novel words. We focus on two properties: (i) formal novelty, depending on whether neologisms result from a morphological operation or are already existing forms that take on new meanings, and (ii) lexical regularity, depending on whether neologisms are created through regular linguistic processes or not. We hypothesize that morphological neologisms are more salient than semantic ones, and that irregular neologisms are more salient than regular ones. We designed a behavioral experiment to test these hypotheses with French native speakers, measuring detection rates and response times for various types of neologisms. The results support our main hypotheses, additionally showing that lexical regularity is a stronger predictor than formal novelty, and that an interaction effect exists between the two factors.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords: Neologism; Neological intuition; Word-formation process; Derivation; Regular polysemy

1. Introduction

Neological intuition (NI) is the metalinguistic ability to evaluate lexical novelty. It relates to how speakers process and understand novel words encountered in everyday life, without necessarily having encyclopedic knowledge about language. Although NI is based on general linguistic competence, it has been shown to be subject to inter-individual variation, due to both linguistic and extralinguistic factors.

In this paper, we focus on linguistic factors, and especially investigate two properties that may influence NI: (i) formal novelty, depending on whether neologisms are new morphological items or already existing forms used with a new meaning, and (ii) lexical regularity, depending on whether neologisms are created through regular or irregular lexical formation processes. We hypothesize that these two properties are predictors of NI variation and account for the variable detection of neologisms. In order to test this hypothesis, we conducted a behavioral experiment with French native speakers, asking them to identify neologisms in given sentences. We observed which neologisms were detected and measured participants' response times. The objective of the study is to contribute to a better understanding of how lexical novelty is assessed, through a psychological approach of linguistic processing.

^{*} Corresponding author.

E-mail address: alizee.lombard@unifr.ch (A. Lombard).

The article is organized as follows. In Section 2, we present the theoretical background and motivation for investigating NI. In Section 3, we describe our experimental methodology. The results of the experiment are reported in Section 4, and discussed in Section 5.

2. Neological intuition in question

In this section, we present some general issues related to NI and report on the few experimental studies that have been conducted on the topic. We also discuss the fundamental properties of NI and introduce the typology of neologisms that is used in the study.

2.1. Issues related to neological intuition

Only a few researchers have investigated NI, and the notion itself is rarely explicitly defined. NI is usually associated with the impression that some lexical items are new in a given language, which can be expressed through metalinguistic judgments. Although scarcely studied, NI relates to general questions about novel word formation, linguistic competence, and the structure of the mental lexicon.

NI is an important element in neologism theory. There is an ongoing debate as to whether it should be considered a defining property of neologisms or not. Some authors (Bouzidi, 2010; Cabré and Estopà, 2009; Sablayrolles, 2010) consider the capacity of triggering NI a necessary condition for neologisms, whereas others (Cartier, 2018; Corbin, 1987) have dismissed it as such. It can be asked if novel words necessarily or uniformly trigger NI, and if not, what can explain the differences observed in NI, and whether neologisms should be characterized with degrees of 'neologicity' depending on their salience as novel items. More generally, NI as a special case of linguistic intuition is of epistemological interest for linguistic theory, given the role of metalinguistic judgments in the development of linguistics. The evidential relevance of speakers' intuitions in linguistic theory is the subject of many discussions (see, e.g., Devitt, 2006, 2012; Machery et al., 2012). Investigating NI can provide information about the properties of neologisms and speakers' judgments in a folk linguistics perspective, but also result in meta-analyses of the linguistic judgments themselves.

NI relates not only to basic linguistic theory, but also to applied fields such as sociolinguistics and psycholinguistics. Variations in NI can be due to sociological factors (Favreau, 2018; Podhorná-Polická and Fiévet, 2018, 1918; Xu, 2001), and analyzing NI with respect to parameters such as age, geographical origin, socio-professional category, etc. can give us information about how neologisms are apprehended by speakers with different backgrounds. For example, one could investigate possible correlations between the recognition of specific types of neologisms (e.g., loanwords) and speakers' individual characteristics (e.g., age, language proficiency, living location).

Psycholinguistic factors are involved as well, especially with respect to lexical representations, access to the mental lexicon, and the cognitive processing of lexical form and meaning. NI relates to how morphologically complex words are processed (Bertram et al., 2000; Beyersmann et al., 2016; Creemers et al., 2020; Feldman and Milin, 2018; Leminen et al., 2019; Rastle and Davis, 2008; among others), how words with multiple senses are mentally represented (Brocher et al., 2018; Eddington and Tokowicz, 2015; Klein and Murphy, 2001; Klepousniotou and Baum, 2007; Klepousniotou et al., 2012; Yurchenko et al., 2020; among others), and how novel words integrate the mental lexicon (Bakker et al., 2014; Kaczer et al., 2018; Langacker, 1987; Lavale-Ortiz, 2019; Schmid, 2008, 2017; among others). Variations in NI may indicate that not all neologisms require equal effort to be processed, and could indirectly inform us on how the mental lexicon is organized. On the one hand, the existence of various degrees of neological salience according to morphological types could be due to the fact that these types are stored and accessed differently in the mind. On the other hand, the salience of neologisms with an already existing form but a new meaning may depend on how ambiguous words are represented in the mental lexicon. As such, variations in NI could correlate with the probability of having separate lexical entries for different types of ambiguous words. More generally, speakers' intuitions about lexical novelty could help us understand how novel words are integrated in the lexicon, and in case of NI variation, which lexical properties facilitate word integration. Eventually, models of lexical integration should be able to account for NI phenomena, and for the possible relation between the morphosemantic properties of lexical items and their ability to trigger NI.

2.2. Previous studies

Many experimental studies about the mental lexicon are based on asking participants to distinguish words from non-words (e.g., *cat* vs. *wug*, see Feldman and Milin, 2018; Pexman, 2012; Yap and Balota, 2015; among others). To some extent, such a lexical decision task can be compared to that of assessing whether given words are novel or not. However, the categories of novel word vs. non-novel word differ from those of word vs. non-word, in that novel words as opposed to

non-words are assumed to be meaningful and to possibly integrate the lexicon. Moreover, the intuitive aspect of NI goes beyond the evaluation of single word recognition that is the concern of word vs. non-word experiments. Studies directly addressing the issue of neological salience, i.e., the identification of novel words as opposed to already existing words, appear to be very few.

Gardin et al. (1974) conducted a pioneering study on NI, aiming at a definition of neologism that was congruent with speakers' intuition. In that study, seventeen people were asked to identify what they thought were neologisms in a French newspaper corpus. Many disagreements were observed, which lead to the conclusion that neologisms could not be defined with reference to largely shared intuitions. Authors noted that the speakers' metalinguistic judgments about lexical novelty were highly variable, still without identifying the cause of such a variation. Consequently, unanswered questions remained as to (i) how speakers access NI, and (ii) what causes NI to vary.

As far as NI variations are concerned, two types of factors can be taken into account. First, NI may vary according to extralinguistic factors, in particular sociological factors, for both the individual lexicon and the representation of the shared lexicon depend on age, profession, level of education, living location, etc. Second, NI may vary according to linguistic factors, such as lexical type, morphological structure or syntactic construction. Linguistic factors of NI have been studied by Sablayrolles (2003) and Ben Hariz Ouenniche (2009).

Sablayrolles (2003) presents a study based on an annotation task in which three researchers in French language or literature, including the author himself, were asked to detect neologisms in a given corpus and to classify them according to a pre-existing typology. The goal of the study was to shed light on NI variation factors, and ultimately to reduce annotators' disagreements. Although the study provides some valuable insights, most notably with regard to the influence of neologism type as a factor of NI variation, the methodology can be questioned. First, the use of an existing corpus does not allow for a strict control of the materials, because it can include primarily undetected neologisms. As a consequence, preliminary hypotheses about the influence of specific neologism types on NI cannot be accurately tested. Second, the small number of annotators involved in the study does not allow for quantitative generalization, because results can be dependent on individuals rather than on common behaviors. Third, the data output is limited, especially in terms of variance, which makes it difficult to evaluate the annotators' true disagreements.

Ben Hariz Ouenniche (2009) conducts a similar study to that of Sablayrolles (2003) with four linguists as participants (including Sablayrolles and herself). The study exhibits the same strengths and weaknesses as the previous one, even though the annotation method was modified in order to reduce disagreements between annotators, and the corpus was selected to be more diverse and representative of everyday spoken language. Importantly, the study supports the idea that NI variation is determined by neologism type.

2.3. Linguistic factors of neological intuition

Following Sablayrolles (2003) and Ben Hariz Ouenniche (2009), we aim to determine whether NI depends on specific linguistic factors. Our study nevertheless differs from the previous ones in that it does not focus on neologism types per se, but on two properties that characterize them: *formal novelty* and *lexical regularity*. In this section, we briefly introduce the classification of neologisms used in our study, and analyze it with respect to these two orthogonal properties. We then present hypotheses about the impact of these properties on NI variation.

Many classifications have been proposed in the literature on neologisms (see Cabré, 2006; Cartier, 2018; Gérard et al., 2017; Lavale-Ortiz, 2019; Poix, 2018; Sablayrolles, 2000, among others). We will focus here on six basic types that will be sufficient for our purpose.

Ex nihilo neologisms are created from scratch without any pre-existing lexical base. The word *ptyx* in (1), coined by the poet and unknown in French, instantiates that category.

- (1) Sur les crédences, au salon vide: nul ptyx (Mallarmé, *Sonnet en X*)
'On the sideboards, in the empty drawing room: no ptyx'

Morphological neologisms are new lexemes which result from an operation on the form and generally the meaning of a lexical base. Such is the case of *détatouer* 'remove a tattoo' in (2), which is derived by prefixation from the base verb *tatouer* 'tattoo'.¹

¹ We consider newly derived words as neologisms, regardless of affix productivity. Morphological neologisms can be the output of systematically productive patterns. Previously unattested forms resulting from the application of inflectional rules to existing words are not regarded here as neologisms, because these are word forms as opposed to lexemes.

- (2) Certaines célébrités vont à contre-courant, en se faisant détatouer. (web)
'Some celebrities go against the tide by having their tattoos removed'

Semantic neologisms are existing words that take on new additional meanings (Bastuji, 1974; Gérard and Kabatek, 2012; Renouf, 2013; Sablayrolles, 2010). For example, the adjective *toxique* 'toxic' is used in (3) with a psychological meaning that is metaphorically derived from an original physiological meaning.

- (3) Dans une relation toxique, les tensions et les critiques sont omniprésentes. (web)
'In a toxic relationship, tension and criticism are omnipresent'

Syntactic neologisms are morphologically and semantically specified items that can be used in a new syntactic construction. This class can be illustrated by the verb *cliquer* 'click' in French, that is ordinarily used with an oblique argument (*cliquer sur quelque chose* 'click on something'), but can be transitivized and used with a direct object as in (4).

- (4) Cliquez la photo pour l'agrandir. (web)
'Click the picture to enlarge it'

Phraseological neologisms are new multiword expressions that result from the conventional use of a phrase. They are characterized by non-compositionality. For instance, the expression *être au taquet* (lit. 'be at the cleat') has been recently lexicalized in French with the meaning of 'be going full throttle', as illustrated in (5).

- (5) Toute l'équipe est au taquet et se réjouit de vous retrouver dès ce samedi. (web)
'The whole team is going full throttle and is looking forward to seeing you again this Saturday'

Loanwords are foreign words or expressions that have been borrowed and adapted to a target language (Haspelmath, 2009; Winter-Froemel, 2009). Such is the case of *ghostwriting* in (6), which has been imported from English into French.

- (6) Monsanto est suspecté d'être un adepte du ghostwriting. (web)
'Monsanto is suspected of using ghostwriting services'

As suggested above, these six basic neological types instantiate formal novelty and lexical regularity to various degrees. On the one hand, formal novelty depends on whether the form of a neologism is novel per se or not (Jamet and Terry, 2018). For instance, *ex nihilo* neologisms (e.g., *ptyx*) contrast with semantic neologisms (e.g., *toxique* 'toxic') in that they are completely new lexical forms. On the other hand, lexical regularity depends on whether or not a formation process can be predicted and productively form series of neologisms. Regularity can apply to formal, semantic and/or syntactic properties of lexical items. For instance, the morphological neologism *détatouer* 'remove a tattoo' has predictable form and meaning, because it follows the morphosemantic pattern associated with the prefix *dé-* in French (i.e., the formation of deverbal verbs to denote opposite actions). *Détatouer* fits into the morphological series that includes verbs such as *défaire* 'undo', *déconstruire* 'deconstruct', *débrancher* 'unplug', etc. By contrast, the phraseological neologism *être au taquet* 'be going full throttle' cannot be predicted, and does not instantiate any regular pattern in French.

In the present study, we investigate the influence of formal novelty and lexical regularity on NI, and formulate a directional hypothesis for each property. More precisely, we hypothesize that (i) neologisms with a novel form are more salient (i.e., more easily identified as neologisms) than neologisms with an already existing form, and (ii) irregular neologisms are more salient than regular neologisms. The first hypothesis is rather straightforward; it is based on the assumption that already known forms are less detected than unknown ones, and are consequently less likely to trigger NI. The second hypothesis is motivated by the idea that irregular neologisms require greater cognitive effort than regular ones to be processed, and that this effort results in a stronger impression of novelty. According to Schmid (2008), transparent and unambiguous morphologically complex neologisms (e.g., words prefixed with *over-* in English such as *overbill*) are more easily integrated in the mental lexicon than semantically opaque ones (e.g., blends such as *sheeple* from *sheep* and *people*). Assuming that neologisms that are easier to process are also less salient, it can be hypothesized that novel regular complex words trigger less NI than irregular ones. A similar extrapolation can be drawn from studies showing that ambiguous words are represented differently in the mental lexicon depending on meaning similarity (see Eddington and

Tokowicz, 2015 for an overview). Regular and irregular polysemes may differ in the way they are represented and processed because of the variable predictability and relatedness of their meanings. Consequently, semantic neologisms could be apprehended differently by speakers, depending on the regularity of their semantic construction, which would then influence NI. Such a hypothesis has been explicitly formulated by Carston and Wearing (2014) and Barque et al. (2018), but has never been experimentally tested.

To explore the effect of formal novelty and lexical regularity on NI, we focus on semantic and morphological neologisms. We privilege these two types of neologisms over the others because, first, they allow for a clear distinction between formal novelty and non-formal novelty, and, second, they polarize regularity and can be divided into distinct classes of regular and irregular items.

3. Method

In this section, we describe our experimental methodology, including participants' selection, materials, procedure, and hypotheses.²

3.1. Participants

Sixty-eight French native students from the University of Fribourg (Switzerland) took part in the experiment ($M_{age} = 20.72$, $SD = 2.176$, age range [18;25]). Psychology students ($N = 55$) received academic credits for their participation. The other students ($N = 13$) were volunteers. To collect non-expert judgments, none of the participants were students in linguistics.

3.2. Materials

For the purpose of our experiment, we created 120 stimuli, including 80 experimental items (i.e., sentences with a neologism) and 40 filler items (i.e., sentences without any neologism). In the following subsections, we successively present how target neologisms were selected and how stimuli sentences were formed.

3.2.1. Target neologisms

To ensure that target neologisms were unknown to participants, most of them were especially created for the experiment. All selected items, be they morphological or semantic neologisms, were absent from two reference dictionaries of French (*TLFi*³ and *Le Petit Robert*⁴). Furthermore, morphological neologisms were controlled through corpus investigations. Target morphological neologisms had 0 occurrence in the *frWaC* corpus (Baroni et al., 2009), and no more than 5 occurrences in the *frTenTen17* corpus (Jakubiček et al., 2013).⁵ We did not dismiss words with very few occurrences in extensive corpora, because these are usually nonce words, coined to fill some immediate lexical need and not intended to integrate the lexicon (Dal and Namer, 2018; Štekauer, 2002). We assume that such words remain salient as novel words – even for the speaker or writer that uses them. As for semantic neologisms, *Google* searches were carried out to ensure that the target items were not frequently used in the new meaning they were given in the experiment.

We eventually selected 54 nouns and 36 verbs as target neologisms. Different parts of speech were selected so as to avoid participants focusing on one part of speech only. Adjectives and adverbs were discarded because their polysemy has been less studied than that of nouns and verbs, and because some of the irregular morphological processes tested in the experiment rarely apply to them. All neologisms were designed to be pragmatically plausible, so that participants would not notice them because of unlikely meaning.

Regular and irregular morphological neologisms were created in line with the distinction between grammatical and extragrammatical word-formation processes (Dressler, 2000; Fradin et al., 2009; Mattiello, 2013). A word-formation process is grammatical and regular when its output is predictable in both form and meaning (e.g., the prefix *re-* and the base verb *foudroyer* 'strike' combine to form *refoudroyer* 'strike again', following the deverbal pattern associated with *re-*). A word-formation process is extragrammatical and irregular when its output is unpredictable in form or meaning (e.g., the

² The experiment was preregistered (with materials, procedure, hypotheses, and data analysis plan) on the OSF platform: https://osf.io/k2zwy/?view_only=c466904e779e48e2bda81c3a56b132f7.

³ <http://www.atilf.fr/tlfi>.

⁴ <https://pr.bvdep.com/robert.asp>.

⁵ *frWaC* and *frTenTen17* are extensive French corpora comprising respectively 1.3 and 5.7 billion words collected from the .fr domain in 2010 and 2017. They are available at the following URLs: https://www.clarin.si/noske/run.cgi/corp_info?corpname=frwac&struct_attr_stats=1&subcorpora=1 and <https://www.sketchengine.eu/frtnten-french-corpus/>.

acronym *formupo* corresponding to *formule de politesse* 'greeting formula' is not formally predictable, and could as well be *formulpo*, *formupol*, *formpol*, etc.). Unlike extragrammatical processes, grammatical processes produce series of neologisms and can therefore be evaluated in terms of productivity.

Derivation, compounding and conversion clearly fulfill the requirements for regularity, whereas acronymy, reduplication and blending are irregular. Acronyms have an unpredictable form since the word sections used to form them cannot be predicted (e.g., *Benelux* takes two phonemes from *Belgique* 'Belgium' and *Nederland* 'Netherlands', but three from *Luxembourg*). Reduplication may be more predictable formally than acronymy. According to Plénat (1999), reduplicated parts in French obey some phonetic constraints. However, these constraints are not sufficient to predict the exact form of the output (e.g., the reduplication of *crade* 'dirty' in French produces three attested forms, *cracra*, *cracrade*, and *crade-crade*)⁶. Furthermore, reduplication is mainly connotative, with important unpredictable variations (e.g., *nounours* from *ours* (*en peluche*) 'teddy bear' is hypocoristic whereas *mémère* from *mère* 'mother' is pejorative). As for blends, they have both unpredictable form and meaning (Fradin, 2015). For instance, it cannot be predicted that the blending of *mondialisation* 'globalization' and *banalisation* 'trivialization' will produce *mondiabanalisation* (instead of e.g., *mondibanalisation* or *banamondialisation*) and have the meaning of 'trivialization due to globalization' (instead of, e.g., 'globalization due to trivialization' or 'trivialization of globalization').⁷

The irregular morphological neologisms used in our experiment are formed through blending ($N = 10$), acronymy ($N = 6$), and reduplication ($N = 4$). Acronyms and reduplications are created respectively from existing phrases and from French monosyllabic words, as in *impadem*, constructed from *impression à la demande* 'print on demand', and *beubeurre*, constructed from *beurre* 'butter'. Blends have to be transparent enough to allow for base recognition, as in *inordinateur* 'information machine', constructed from *information* 'information' and *ordinateur* 'computer'. The regular morphological neologisms used in our experiment are suffixed words ($N = 10$), prefixed words ($N = 6$), and compounds ($N = 4$). All of them fit into existing morphological series. For instance, *aide-pianiste* 'piano assistant' follows the word-formation pattern that produces compounds beginning with *aide* 'assistant' (e.g., *aide-cuisinier* 'kitchen assistant'), and *refoudroyer* 'strike again' follows the prefixation pattern associated with *re-* (e.g., *rediscuter* 'discuss again'). Even though derivational productivity was not directly taken into account in the study, we intentionally selected affixes with heterogeneous productivity. For example, the nominalization in *-age*, used to form *assombrissage* 'darkening', is clearly more productive⁸ in contemporary French than the nominalization in *-ure*, used to form *emboutissure* 'bump'.

Regular and irregular semantic neologisms were distinguished according to the definition of regular polysemy introduced by Apresjan (1974). The polysemous relation between two word senses is considered regular if it is instantiated by at least two non-synonymous lexical units. Regular polysemy patterns can be identified and semantically analyzed (Barque, 2008; Copestake and Briscoe, 1995; Nunberg, 1995; Pustejovsky, 1995). For example, the BODY PART → OBJECT PART pattern instantiated in French by nouns such as *pied* 'foot'/'base of an object', *tête* 'head'/'top of an object', *cœur* 'heart'/'central part of an object' is based on a metaphor that relies on a locative or functional analogy (Svorou, 1994). Similarly to regular morphological processes, regular polysemy patterns can be evaluated in terms of productivity, depending on the number of lexical forms that instantiate them. For instance, the FOOD SUBSTANCE → PORTION OF FOOD pattern (e.g., *yaourt* 'yogurt'/'pot of yogurt', *bière* 'bier'/'glass of bier') is more productive in French than the PIECE OF FURNITURE → PLACE pattern (e.g., *bureau* 'desk'/'office', *bar* 'counter'/'bar'). As opposed to regular polysemy, irregular polysemy occurs when only one lexical form instantiates a given polysemous relation (e.g., *licorne* 'unicorn'/'startup company valued at over \$1 billion'). As pointed by Brocher et al. (2018), the different meanings of irregular polysemes bear some similarity to each other but are not derived by means of a semantic rule. The relation between the different meanings of an irregular polyseme is both unpredictable and idiosyncratic.

Since metaphor tends to produce more irregular polysemes than metonymy, most irregular semantic neologisms in our materials are metaphorical. Regular semantic neologisms are the output of 4 nominal metonymy patterns and 6 verbal or nominal metaphor patterns. Polysemy patterns are documented in the literature (Apresjan, 1974; Barque et al. 2018; Goossens, 2009; Lakoff and Johnson, 1980; Srinivasan et al., 2019, among others). For instance, the noun *embauche* 'recruitment', used as a semantic neologism to denote people in charge of recruitment, follows the ACTION → AGENT metonymic pattern instantiated by nouns such as *redaction* 'writing'/'editorial board' and *mission* 'mission'/'people on a mission'. As in the case of regular morphological neologisms, we selected patterns with various degrees of productivity.

⁶ It has to be noted that Plénat (1999) investigates only the reduplication of first names, and that there is no evidence that the phonetic constraints he brings to light also apply to common nouns, which are much less frequently reduplicated. Only reduplicated common nouns were included in our experimental materials.

⁷ Clipping is often mentioned as an extragrammatical morphological process. However, this view can be challenged considering that clipping may be more productive and regular than prototypical extragrammatical processes (Arndt-Lappe, 2018; Bauer et al., 2013). By default, we did not consider here clippings as illustrative of irregular morphological neologisms, and did not use them in the experiment.

⁸ *Productive* is used in this paper as a scalar term. It refers to a relative, as opposed to absolute, property. In other words, a productive process does not necessarily apply systematically to items that meet its input requirements.

Table 1
Length of the stimuli for each condition.

Condition	Mean (char.)	Variance (char.)	Mean (syl.)
Irregular morphological	54.05	6.763	14.75
Regular morphological	54.05	6.909	14.7
Irregular semantic	54.6	6.809	14.75
Regular semantic	54.35	6.846	14.6
Filler	54.775	6.685	14.325

However, the absence of theoretical work on scalar regular polysemy, and the lack of a lexical resource listing the items that instantiate the different polysemy patterns made it impossible to precisely quantify regularity variations, and therefore to analyze them as a variable in our experimental results.

All the neological items used in the experiment are presented in the supplementary materials (see Appendix B).

3.2.2. Experimental sentences

Target neologisms were included in simple sentences of ordinary French such as (7), which includes the regular semantic neologism *moustache* 'mustache' as an instance of the BODY PART → OBJECT PART pattern.

- (7) Il faudrait absolument nettoyer la moustache du balai.
'We should definitely clean the mustache of the broom'

Neologisms were placed in different syntactic positions, provided these were not prominent positions (i.e., beginning or end of sentence). Sociolinguistic and register variations were avoided so as to control their effect on neologism identification. Filler sentences were designed to be syntactically similar to experimental ones.

Semantic neologisms required some additional conditions. Namely, the context of the sentence had to block the original interpretation of the target lexeme. A subsequent difficulty is that stimuli including semantic neologisms might appear absurd to speakers, especially in the case of irregular semantic neologisms. We will come back to this issue in Section 5.

The length of the stimuli, measured as the number of characters and syllables, was standardized across experimental and filler sentences, so that all conditions have the same mean and variance, as shown in Table 1.

One possible concern with the experimental design described above is that the syntactic simplicity of the stimuli sentences could by contrast make the target neologisms salient, thus inducing a ceiling effect on neologism identification. To reduce that ceiling effect, we unbalanced the number of experimental and filler items to two thirds/one third, assuming that participants would tend to maintain a balance between positive and negative answers. Each experimental condition therefore consists of 20 stimuli, which results in a total of 80 neological stimuli, whereas the filler condition consists of 40 stimuli.

3.3. Procedure

Our experimental study design is based on a within-subject factorial plan using formal novelty and lexical regularity as the two main factors. These properties are referred to respectively as *Change* (morphological vs. semantic neologisms) and *Regularity* (irregular vs. regular neologisms). The experiment took the form of a survey completed online on the *Qualtrics* platform. The survey took an average of 20 minutes to complete, which participants did using their personal computer at home.

Participants were asked to perform two tasks. The first one consisted in reading stimuli sentences displayed on the screen one after the other, and deciding for each sentence, as quickly and accurately as possible, whether they included a neologism or not. The exact instructions were: "Indicate whether the following sentence includes or not a new word or an existing word used with a new meaning".⁹ The term *neologism* was intentionally omitted so as to avoid linguistic terminology and misunderstanding of the notion.

⁹ French instructions were: "Indiquez si la phrase suivante contient un mot nouveau ou un mot existant employé dans un sens nouveau".

Answers were given by pressing a labelled key *Oui* 'yes' with the left index finger or a labelled key *Non* 'no' with the right index finger. Each sentence was presented for as long as needed for participants to make their decision and give their answer. To limit biased results due to fatigue effects, stimuli were presented to participants in a random order. In order to familiarize participants with the task, a short training session was proposed, using neologisms created through processes that were absent from the experiment (e.g., loanwords).

The second task consisted in precisely identifying novel words in the stimuli sentences that were detected as including a neologism in the first task. All these sentences were listed on the same page, and participants were asked to highlight neologisms by clicking on them. Only one word per sentence could be identified. Participants could take as long as they wanted to perform that second task.

Response time, i.e., the time elapsed between the display of a stimulus and the response, was measured in the first task. Neologism detection was evaluated through both the first and second tasks, and target neologisms were accordingly tagged as 'detected' or 'not detected'. Importantly, when a participant gave a positive answer in the first task but was unable to correctly identify the target neologism in the second task, the neologism was considered 'not detected'.

3.4. Hypotheses

Our experimental study was designed to test four hypotheses, which are as follows.

H1: Morphological neologisms have a higher detection rate than semantic neologisms.

H2: Irregular neologisms have a higher detection rate than regular neologisms.

H3: When detected, morphological neologisms cause shorter response times than semantic neologisms.

H4: When detected, irregular neologisms cause shorter response times than regular neologisms.

We additionally hypothesized that there might be an interaction effect of the two predictors (formal novelty and lexical regularity) on detection rates and/or response times, depending on how participants would react to irregular semantic neologisms. These neologisms in particular cause salient semantic incongruity, which could be attributed to a general semantic mismatch in stimuli sentences rather than to the new use of any specific word, and/or cause more reflection time. This could result in lower detection rates and/or longer response times than for regular semantic neologisms (whereas irregular morphological neologisms would still cause higher detection rates and shorter response times than regular morphological neologisms).

4. Results

In this section, we present the results of our experiment. We first focus on detection rates for target neologisms, then examine response times in case of neologism detection.¹⁰ To limit inaccurate responses due to participants' lack of attention, we only considered data from participants who completed the entire survey in more than ten minutes and less than an hour (which resulted in the selection of 68 participants out of the 79 who completed the survey). All responses given in a time shorter than 200 ms (0.02% of the data) were excluded as well, since stimuli sentences or even words alone cannot be read carefully in such a short time.

We used the *lme4* package (Bates et al., 2014) in R (R Core Team, 2015) to perform the analyses using mixed-effects models with crossed random effects for subjects and items (Baayen et al., 2008). We calculated *p*-values using log-likelihood ratio tests, and considered them significant if less than 0.01.¹¹

4.1. Detection

Observed detection rates for each condition are presented in Fig. 1 and Table 2.

We used a mixed logistic regression model to analyze the detection data. The *Change* (morphological vs. semantic) and *Regularity* (irregular vs. regular) factors as well as their interaction were entered into the model as predictors. As random effects, we included intercepts for participants and items, and by-participant random slopes for the two predictors and their interaction. By-item slopes were not included, as they were not justified by the design and would generate

¹⁰ The collected data and R scripts used for statistical analysis are provided at https://osf.io/x6whk/?view_only=9da55b31cbab488aaa64fd8dfac1b54b.

¹¹ A threshold of 0.01 was preferred over the traditional 0.05 following Colquhoun (2014). The 0.01 threshold is also mentioned by Baayen (2008) and Meteyard and Davies (2020).

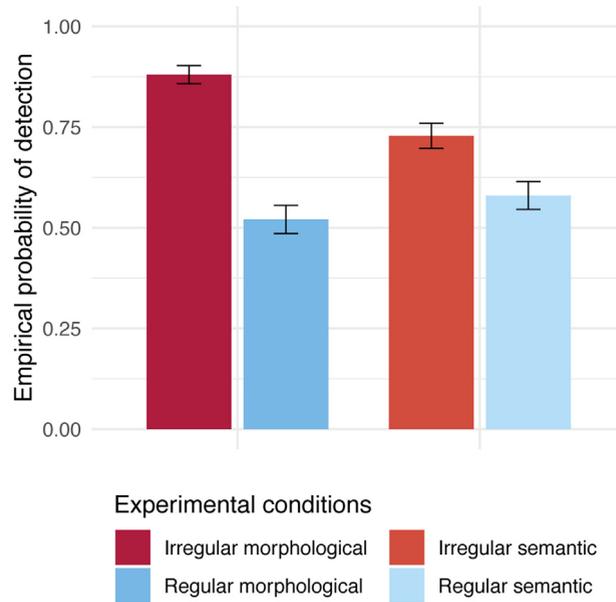


Fig. 1. Empirical probability of detection per experimental condition with 99% Wald confidence intervals. The 1st and 2nd bars indicate rates for morphological neologisms, the 3rd and 4th bars indicates rates for semantic ones. The 1st and 3rd bars indicate rates for irregular neologisms, the 2nd and 4th bars indicate rates for regular ones.

Table 2
Average detection rates per neologism type with 99% confidence intervals.

Neologism type	Mean detection rate	99% CI
Morphological irregular	0.8801471	0.02272230
Morphological regular	0.5205882	0.03495037
Semantic irregular	0.7284768	0.03112586
Semantic regular	0.5801471	0.03452772

Table 3
Summary of the mixed logistic regression model for neologism detection with morphological irregular neologisms as intercept.

Effect	Estimate	SE	z-value
Intercept	2.4465	0.2368	10.333
Change	-1.1142	0.3449	-3.231
Regularity	-2.3224	0.3003	-7.733
Change × Regularity	1.3884	0.4165	3.334

unnecessary model complexity (Bates et al., 2018). The best model to account for the data was selected based on log-likelihood ratio tests (see Table 5 in Appendix A).

Results of the regression model are presented in Table 3. The significance of the interaction was assessed using log-likelihood ratio tests comparing models with and without the effect. The significance of the simple effects was calculated using the same method on models without the interaction. The robust effects that emerge from these analyses are that of lexical regularity ($p = 4.904e-11$), and that of the interaction between formal novelty and lexical regularity ($p = 0.001185$). The main effect of formal novelty is not significant ($p = 0.2312$). The same models were run without residuals greater than 2.5 SD (0.51% of the original data), considering that large residuals could influence the model (Baayen, 2010). The results are similar to those presented in Table 3.

Although morphological neologisms were generally better detected than semantic neologisms, the difference was not significant. The effect of formal novelty on detection rates actually depends on lexical regularity. In the observed results (see Fig. 1), the detection rate decreases by 13.92% between morphological and semantic neologisms in irregular conditions, but slightly increases (by 6.62%) in regular conditions. We further explored these modality shift effects in the model, using the *lsmeans* package (Lenth, 2016), which indicated that the decrease in irregular conditions was significant ($p = 0.0012$), whereas the increase in regular ones was not ($p = 0.3552$). Thus, the effect of formal novelty is only partially consistent with our H1 hypothesis, for it is only significant in the case of irregular conditions.

As for lexical regularity, detection rates clearly decrease between irregular and regular neologisms, with a difference of 35.66% for morphological neologisms and 15.13% for semantic ones. The effect of lexical regularity is significant, both within and across modalities of formal novelty. These results are therefore consistent with our H2 hypothesis.

It appears that the effect of formal novelty varies according to regularity modalities, since unlike irregular neologisms, regular neologisms are not significantly better detected when they are novel forms. This interaction effect is different from the one we predicted, hypothesizing that irregular semantic neologisms might be less detected than expected (see Section 3.4).

4.2. Response times

Only positive answers in the detection task with correct target identification were taken into account in the analysis of response times, which represents 67.62% of all responses. Average response times for each condition are presented in Fig. 2 and Table 4.

To evaluate the significance of each predictor, we computed a full model including interaction between the *Change* and *Regularity* factors. Response times were log-transformed to reduce the skewness of their distribution (Baayen, 2010), and then analyzed using a mixed linear regression model. In addition to the *Change* and *Regularity* factors, we included in the model a predictor corresponding to the length of the stimuli, measured in characters (*LChar*). As random effects, we included intercepts for participants and items. We also included by-participant random slopes for *Change*, *Regularity* and their interaction.

The best model was selected based on log-likelihood ratio tests (see Table 6 in Appendix A). This model includes the main effects, random intercepts and random slope for *Change*, but neither the interaction effect nor the other random slopes. An examination of the residuals did not reveal any violation of the assumption of normal distribution and homoscedasticity. The results of the regression model are presented in Table 5. The analysis shows three main effects,

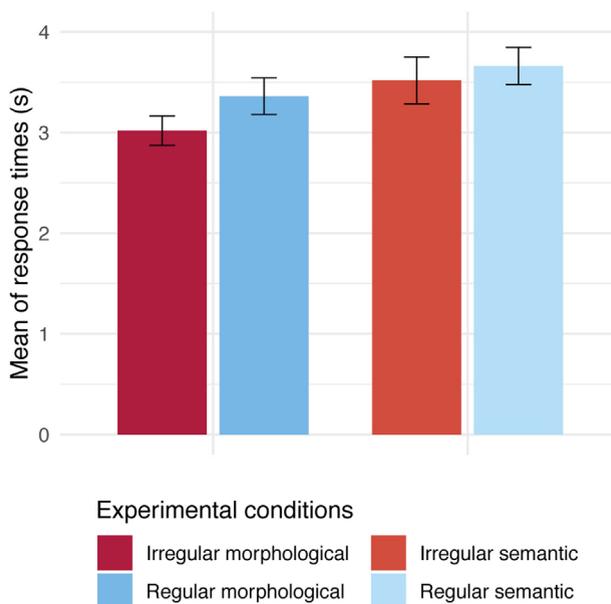


Fig. 2. Average decision latency for stimuli per experimental condition with 99% confidence intervals. The 1st and 2nd bars indicate response times for morphological neologisms, the 3rd and 4th bars indicate response times for semantic ones. The 1st and 3rd bars indicate response times for irregular neologisms, the 2nd and 4th bars indicate response times for regular ones.

Table 4
Average response time in seconds per neologism type with 99% confidence intervals.

Neologism type	Mean response times (s)	99% CI
Morphological irregular	3.017682	0.1452850
Morphological regular	3.360864	0.1815124
Semantic irregular	3.517705	0.2329972
Semantic regular	3.660608	0.1853415

Table 5
Summary of the mixed linear regression model for the log-transformed response times with morphological irregular neologisms as intercept.

Effect	Estimate	SE	t-value
Intercept	0.496448	0.130836	3.794
Change	0.122874	0.030757	3.995
Regularity	0.114317	0.028168	4.058
LChar	0.009782	0.002256	4.337

those of formal novelty ($p = 9.866e-05$), lexical regularity ($p = 7.519e-05$), and sentence length ($p = 2.81e-05$). The model was run again without residuals greater than 2.5 *SD* (1.68% of the original data), and on a set of data trimmed to 2.5 *SD* over and under by-participant and by-condition means (which removed 2.52% of the original data).¹² The estimates and significance levels given by these models are similar to those presented in Table 5.

Note that although the visual exploration of the data seems to indicate a difference (see Fig. 2), the *Change* and *Regularity* effects are of similar size, as indicated by the estimates in Table 5. Overall, morphological neologisms are detected more rapidly than semantic ones, and irregular neologisms are detected more rapidly than regular ones. These results are thus consistent with our H3 and H4 hypotheses. Since no interaction effect is observed, our complementary hypothesis about a possible interaction between formal novelty and lexical regularity is not supported here.

5. Discussion

The goal of our experiment was to determine the influence of formal novelty and lexical regularity on NI. To this end, we examined NI variations among French (non-linguist) speakers who were asked to perform a metalinguistic task of novel word identification. As far as lexical detection is concerned, irregular morphological neologisms were significantly better detected than irregular semantic ones. There was, however, little difference between regular morphological neologisms and regular semantic ones. Our data therefore suggest that formal novelty does not always influence the identification of novel words. By contrast, lexical regularity appears to be a strong predictor of neologism detection, although it is rarely taken into account in studies on neologisms. Detection rates were significantly higher for irregular neologisms than for regular ones, which suggests that the formal aspect of neologisms is not as striking as the unusualness of their lexical construction, be it semantic or morphological. Neologisms appear to be less salient when they can be integrated in dense lexical networks formed by words resulting from the same lexical pattern.

According to our data, the effect of the two predictors on neologism detection depends on how their modalities combine. This interaction effect needs further analysis, and raises questions as to why formal novelty in particular does not have any significant effect on the detection of regular neologisms. A possible explanation is based on the fact that lexical regularity differs substantially depending on whether it applies to morphology or semantics. As defined in Section 2.3, lexical regularity is a higher-order notion that applies to both morphosemantic and semantic properties. A lexeme is regularly formed if it follows a productive, predictable and specifiable pattern at the morphological and/or semantic level. While lexical regularity determines form and meaning in the case of morphological neologisms, it only applies to meaning in the case of semantic neologisms. Such a difference could influence the cognitive processes involved in neologism identification. In the case of morphological neologisms, detection may depend primarily on formal

¹² Although log-transformation of response times has been debated in the literature, it has the advantage of reducing the influence of extreme values in the analysis (Baayen, 2008, 31). Due to some long RTs, our data were very skewed, hence our preference for log-transformation. We still ran our model on raw RTs, trimmed to 2.5 *SD* over and under by-participant and by-condition means. The analysis lead to similar results as those observed with log-transformed data.

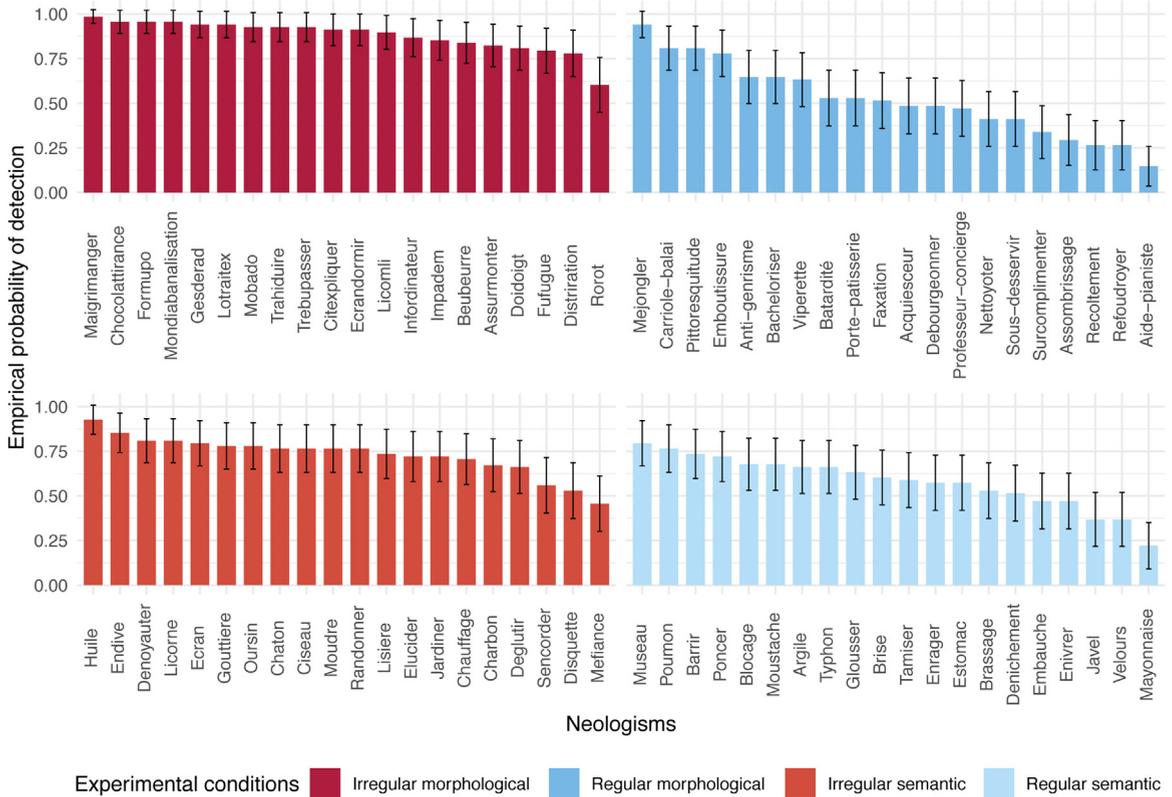


Fig. 3. Empirical probability of detection per neologism with 99% Wald confidence interval. The bars in the left half of the figure indicate rates for irregular neologisms, the bars in the right half indicate rates for regular ones. The bars in the top half of the figure indicate rates for morphological neologisms, the bars in the bottom half indicate rates for semantic ones.

analysis, whereas in the case of semantic neologisms, it is necessarily based on semantic evaluation, which implies both the differentiation with an existing sense and contextual assessment. As regular morphological neologisms are formally predictable and non-context-dependent, they can easily be unnoticed when they are used consistently in discourse. By contrast, regular semantic neologisms require the identification of a novel meaning that is both related to the original one and contextually coherent, which could maintain some neological salience. Lexical regularity could thus have a variable impact on the salience of morphological and semantic neologisms, and cause the interaction effect we observed.

Another interesting point about regular conditions is their heterogeneous detection (see Fig. 3). Detection rates are more dispersed for regular conditions ($SD = 0.181$) than for irregular conditions ($SD = 0.125$). This greater dispersion could be influenced by productivity variations in morphological and semantic patterns. Productivity, and more generally analogy phenomena, can be related to scalar differences in regularity, which could in turn determine differences in neological salience. We observed stronger detection rates for neologisms based on patterns that we considered to be poorly productive than for neologisms based on patterns that we considered to be highly productive. Since productivity, among other factors, plays a role in the retrieval and lexical representation of morphologically complex words (Dal Maso and Giraud, 2019), it may also influence the processing of morphological neologisms, and a similar influence could be observed in the case of semantic neologisms. This would lead us to hypothesize a correlation between productivity, scalar regularity and neological salience. The distinction of different degrees of regularity would certainly have been useful to homogenize the experimental conditions, to control regularity variations, and to reduce the variance of the model. However, such a distinction requires a clear evaluation of regularity degrees based on productivity rates, which is difficult to assess in practice, especially in the case of semantic neologisms.¹³

¹³ To our knowledge, there is not any theoretical account of how regular polysemy productivity should be calculated. A relevant measure should certainly be based on the proportion of members of a semantic class that instantiate a given polysemy pattern. Such a proportion can hardly be evaluated (at least for French), given the lack of lexical resources that propose reliable systematic semantic descriptions.

It can be noted that the dispersion of detection rates is larger for regular morphological neologisms ($SD = 0.210$) than for regular semantic ones ($SD = 0.147$). Formal novelty seems to polarize neological salience when combined with lexical regularity. This can be due to the fact that regular morphological neologisms are easily apprehended when they follow highly productive patterns, given the availability of these patterns and the predictability of their outputs. Highly regular morphological neologisms require little cognitive effort, and are consequently not salient as novel words. By comparison, regular morphological neologisms that follow poorly productive patterns require more effort, as these patterns are less available for morphosemantic analogy. Weakly regular morphological neologisms are consequently more salient. As for regular semantic neologisms, their new meanings have to be inferred both by analogy with an existing polysemy pattern and, as stated above, through contextual evaluation. The corresponding cognitive effort certainly increases neological salience, even for items that follow highly regular polysemy patterns. Highly regular semantic neologisms could thus be more salient than highly regular morphological ones, somehow compensating for the effect of formal non-novelty, and tending to equalize detection probabilities between regular morphological and regular semantic neologisms. Future studies may want to further examine this particular issue.

Response times highlight a different aspect of NI variation. They suggest not only that regular neologisms are more difficult to identify than irregular ones, but also that semantic neologisms are more difficult to identify than morphological ones. The difference in response time with respect to formal novelty could be due to the fact that morphological and semantic neologisms activate different cognitive processes. Morphological neologisms may be analyzed primarily with respect to formal properties, be it through decomposition as some authors suggest (Kaczer et al., 2015; Pollatsek et al., 2011; Smolka and Libben, 2017) or through alternate processes (Baayen et al., 2019; Kinoshita, 2015). The possibility of formal evaluation, as opposed to context-dependent semantic evaluation, could explain why morphological neologisms are detected more rapidly than semantic ones.

Considered together, our observations are consistent with the idea that the novelty of form and meaning is more salient than that of meaning alone. They also clearly indicate that irregular lexical constructions produce more salient neologisms than regular ones. Our general hypotheses about NI are therefore confirmed by the results of the experiment.

6. Conclusion

In this article, we investigated the influence of formal novelty and lexical regularity on neological intuition, through an experimental study of speakers' metalinguistic judgments about neologisms. Our results show that NI depends on neologism type, in particular that it is stronger for morphological neologisms than for semantic ones, and stronger for irregular neologisms than for regular ones. These results especially highlight the importance of the regularity factor, which is, maybe surprisingly, more decisive than formal novelty. It appears that the existence of semantic or morphosemantic lexical patterns into which neologisms fit has an important influence on their salience as novel words. Given that influence, lexical regularity should be taken into consideration more often in experimental studies on neologisms and on word recognition. More generally, the results of our experiment confirm that neologisms are heterogeneous with respect to their capacity to trigger NI. Some neologisms may not be salient at all, as in the case of regular morphological neologisms based on highly productive patterns. This observation could extend to neologism types we did not study here, such as syntactic neologisms (e.g., in the case of transitivized verbal constructions).

To overcome the limitations of our study, we need to develop a method to experimentally evaluate the productivity of lexical patterns, so that lexical regularity can be treated as a scalar factor. Future work will address this particular issue and further investigate the influence of regularity degrees on NI. Separate experiments on morphological and semantic neologisms will be carried out to provide detailed insights into how speakers' judgments about novel words are affected by lexical regularity variations. We also plan to take into account more linguistic predictors, such as the existence of competing synonyms, lexical base frequency, neologism semantic types, and semantic (non-)compositionality.

Conflict of interest

None declared.

Acknowledgement

We are grateful to the two anonymous reviewers for their valuable comments and suggestions on an earlier version of this paper. We also thank Sandra Schwab for her helpful advice in statistical analysis, and Justine Salvadori for proofreading the manuscript.

Appendix A. Log-likelihood ratio tests between models

Table 5

Summary of the log-likelihood ratio tests between different models for the detection of neologisms.

Model	Df	AIC	BIC	logLik	Deviance	Chisq	Chi Df	Pr(> Chisq)
Detection ~ 1 + (1 Participant) + (1 QN)	3	5696.6	5716.4	-2845.3	5690.6			
Detection ~ Regularity + (1 Participant) + (1 QN)	4	5658.2	5684.6	-2825.1	5650.2	40.451	1	2.016e-10
Detection ~ Regularity + (1 Participant) + (1 QN)	4	5658.2	5684.6	-2825.1	5650.2			
Detection ~ Change + Regularity + (1 Participant) + (1 QN)	5	5655.2	5688.2	-2822.6	5645.2	4.9596	1	0.02595
Detection ~ Change + Regularity + (1 Participant) + (1 QN)	5	5655.2	5688.2	-2822.6	5645.2			
Detection ~ Change * Regularity + (1 Participant) + (1 QN)	6	5642.8	5682.4	-2815.4	5630.8	14.381	1	0.0001493
Detection ~ Change * Regularity + (1 Participant) + (1 QN)	6	5642.8	5682.4	-2815.4	5630.8			
Detection ~ Change * Regularity + (1 + Regularity Participant) + (1 QN)	8	5637.1	5689.9	-2810.5	5621.1	9.7534	2	0.007622
Detection ~ Change * Regularity + (1 + Regularity Participant) + (1 QN)	8	5637.1	5689.9	-2810.5	5621.1			
Detection ~ Change * Regularity + (1 + Change + Regularity Participant) + (1 QN)	11	5491.9	5564.5	-2734.9	5469.9	151.17	3	< 2.2e-16
Detection ~ Change * Regularity + (1 + Change + Regularity Participant) + (1 QN)	11	5491.9	5564.5	-2734.9	5469.9			
Detection ~ Change * Regularity + (1 + Change * Regularity Participant) + (1 QN)	15	5483.6	5582.6	-2726.8	5453.6	16.28	4	0.002666

Table 6

Summary of the log-likelihood ratio tests between different models for the log-transformed response times (LogRT).

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(> Chisq)
LogRT ~ 1 + (1 Participant) + (1 QN)	4	4092.2	4117.0	-2042.1	4084.2			
LogRT ~ LChar + (1 Participant) + (1 QN)	5	4082.1	4113.2	-2036.0	4072.1	12.089	1	0.0005071
LogRT ~ LChar + (1 Participant) + (1 QN)	5	4082.1	4113.2	-2036.0	4072.1			
LogRT ~ Regularity + LChar + (1 Participant) + (1 QN)	6	4070.7	4108.0	-2029.3	4058.7	13.41	1	0.0002503
LogRT ~ Regularity + LChar + (1 Participant) + (1 QN)	6	4070.7	4108.0	-2029.3	4058.7			
LogRT ~ Change + Regularity + LChar + (1 Participant) + (1 QN)	7	4056.4	4099.9	-2021.2	4042.4	16.287	1	5.443e-05
LogRT ~ Change + Regularity + LChar + (1 Participant) + (1 QN)	7	4056.4	4099.9	-2021.2	4042.4			
LogRT ~ Change * Regularity + LChar + (1 Participant) + (1 QN)	8	4054.9	4104.6	-2019.4	4038.9	3.5432	1	0.05979
LogRT ~ Change + Regularity + LChar + (1 Participant) + (1 QN)	7	4056.4	4099.9	-2021.2	4042.4			
LogRT ~ Change + Regularity + LChar + (1 + Change Participant) + (1 QN)	9	4043.8	4099.7	-2012.9	4025.8	16.645	2	0.000243
LogRT ~ Change + Regularity + LChar + (1 + Change Participant) + (1 QN)	9	4043.8	4099.7	-2012.9	4025.8			
LogRT ~ Change + Regularity + LChar + (1 + Change + Regularity Participant) + (1 QN)	12	4041.3	4115.9	-2008.7	4017.3	8.4278	3	0.03795
LogRT ~ Change + Regularity + LChar + (1 + Change Participant) + (1 QN)	9	4043.8	4099.7	-2012.9	4025.8			
LogRT ~ Change + Regularity + LChar + (1 + Change * Regularity Participant) + (1 QN)	16	4043.4	4142.8	-2005.7	4011.4	14.391	7	0.04465

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.lingua.2021.103055>.

References

- Apresjan, J., 1974. Regular polysemy. *Linguistics* 42, 5–32.
- Arndt-Lappe, S., 2018. Expanding the lexicon by truncation: Variability, recoverability, and productivity. In: Arndt-Lappe, S., Braun, A., Moulin, C., Winter-Froemel, E. (Eds.), *Expanding the Lexicon: Linguistic Innovation, Morphological Productivity, and Ludicity*. De Gruyter, pp. 141–170. <http://dx.doi.org/10.1515/9783110501933-143>
- Baayen, R.H., 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge University Press, New York.
- Baayen, R.H., 2010. Analyzing reaction times. *Int. J. Psychol. Res.* 3. <http://dx.doi.org/10.21500/20112084.807>
- Baayen, R.H., Chuang, Y.Y., Shafaei-Bajestan, E., Blevins, J., 2019. The discriminative lexicon: a unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity* 2019. <http://dx.doi.org/10.1155/2019/4895891>
- Baayen, R.H., Davidson, D.J., Bates, D., 2008. Mixed-effects modeling with crossed random effects for subjects and items. *J. Memory Lang.* 59, 390–412. <http://dx.doi.org/10.1016/j.jml.2007.12.005>
- Bakker, I., Takashima, A., van Hell, J.G., Janzen, G., McQueen, J.M., 2014. Competition from unseen or unheard novel words: lexical consolidation across modalities. *J. Memory Lang.* 73, 116–130. <http://dx.doi.org/10.1016/j.jml.2014.03.002>
- Baroni, M., Bernardini, S., Ferraresi, A., Zanchetta, E., 2009. The wacky wide web: a collection of very large linguistically processed web-crawled corpora. *Lang. Resour. Eval.* 43, 209–226. <http://dx.doi.org/10.1007/s10579-009-9081-4>
- Barque, L., 2008. *Description et formalisation de la polysémie régulière du français. Thèse en vue de l'obtention du diplôme de docteur de l'Université Paris 7*.
- Barque, L., Haas, P., Huyghe, R., 2018. Polysémie régulière et néologie sémantique. Constitution d'une ressource pour l'étude des sens nouveaux. *Neologica* 12, 91–108. <http://dx.doi.org/10.15122/isbn.978-2-406-08196-8.p.0091>
- Bastuji, J., 1974. Aspects de la néologie sémantique. *Linguages* 36, 6–19. <http://dx.doi.org/10.3406/lgge.1974.2270>
- Bates, D., Kliegl, R., Vasishth, S., Baayen, R.H., 2018. *Parsimonious Mixed Models*. arXiv 1506.
- Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. lme4: Linear Mixed-Effects Models Using Eigen and s4. R Package Version 1.1-6. <http://CRAN.R-project.org/package=lme4>
- Bauer, L., Lieber, R., Plag, I., 2013. *The Oxford Reference Guide to English Morphology*. Oxford University Press, New York. <http://dx.doi.org/10.1093/acprof:oso/9780198747062.001.0001>
- Ben Hariz Ouenniche, S., 2009. Diminuer les fluctuations du sentiment néologique. *Neologica* 3, 37–51. <http://dx.doi.org/10.15122/isbn.978-2-8124-4229-2.p.0041>
- Bertram, R., Schreuder, R., Baayen, R.H., 2000. The balance of storage and computation in morphological processing: the role of word formation type, affixal homonymy, and productivity. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 489–511. <http://dx.doi.org/10.1037/0278-7393.26.2.489>
- Beyersmann, E., Ziegler, J.C., Castles, A., Coltheart, M., 2016. Morpho-orthographic segmentation without semantics. *Psychonom. Bull. Rev.* 23, 533–539. <http://dx.doi.org/10.3758/s13423-015-0927-z>
- Bouzidi, B., 2010. Néologisme et temporalité dans le processus néologique. *Synergies Algérie* 9, 27–36.
- Brocher, A., Koenig, J.P., Mauner, G., Foraker, S., 2018. About sharing and commitment: the retrieval of biased and balanced irregular polysemes. *Lang. Cogn. Neurosci.* 33, 443–466. <http://dx.doi.org/10.1080/23273798.2017.1381748>
- Cabré, M.T., 2006. La clasificación de neologismos: una tarea compleja. *Alfa: Rev. Linguist.* 50, 229–250.
- Cabré, M.T., Estopà, R., 2009. *Les paraules noves: criteris per detectar i mesurar els neologismes*. Eumo Editorial, Barcelona.
- Carston, R., Wearing, C., 2014. Metaphor, hyperbole and simile: a pragmatic approach. *Lang. Cogn.* 3, 283–312. <http://dx.doi.org/10.1515/langcog.2011.010>
- Cartier, E., 2018. *Dynamique lexicale des langues: éléments théoriques, méthodes automatiques, expérimentations en français contemporain. Mémoire d'Habilitation à diriger des Recherches de l'Université Paris 13-Villetaneuse*.
- Colquhoun, D., 2014. An investigation of the false discovery rate and the misinterpretation of p-values. *R. Soc. open sci.* 1, 283–312. <http://dx.doi.org/10.1098/rsos.140216>
- Copestake, A., Briscoe, T., 1995. Semi-productive polysemy and sense extension. *J. Semant.* 12, 15–67. <http://dx.doi.org/10.1093/jos/12.1.15>
- Corbin, D., 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen, Niemeyer.
- Creemers, A., Goodwin Davies, A., Wilder, R.J., Tammings, M., Embick, D., 2020. Opacity, transparency, and morphological priming: a study of prefixed verbs in dutch. *J. Memory Lang.* 110. <http://dx.doi.org/10.1016/j.jml.2019.104055>
- Dal, G., Namer, F., 2018. Playful nonce-formations in french: Creativity and productivity. In: Arndt-Lappe, S., Braun, A., Moulin, C., Winter-Froemel, E. (Eds.), *Expanding the Lexicon: Linguistic Innovation, Morphological Productivity, and Ludicity*. De Gruyter, Berlin, pp. 203–228. <http://dx.doi.org/10.1515/9783110501933-205>
- Dal Maso, S., Giraudo, H., 2019. On the interplay between family and series effects in morphological masked priming. *Morphology* 29, 293–315. <http://dx.doi.org/10.1007/s11525-019-09341-6>
- Devitt, M., 2006. Intuitions in linguistics. *Br. J. Philos. Sci.* 57, 481–513. <http://dx.doi.org/10.1093/bjps/axl017>
- Devitt, M., 2012. Whither experimental semantics? *THEORIA Int. J. Theory Hist. Found. Sci.* 27, 5–36. <http://dx.doi.org/10.1387/theoria.3866>
- Dressler, W.U., 2000. Extragrammatical vs. marginal morphology. In: Doleschal, U., Thornton, A.M. (Eds.), *Extragrammatical and Marginal Morphology*. Lincom Europa, München, pp. 1–10. <http://dx.doi.org/10.1515/9783110501933-205>
- Eddington, C.M., Tokowicz, N., 2015. How meaning similarity influences ambiguous word processing: the current state of the literature. *Psychonom. Bull. Rev.* 22, 13–37. <http://dx.doi.org/10.3758/s13423-014-0665-7>
- Favreau, H., 2018. Commentaires épilinguistiques et sentiment néologique: évaluation comparative de quelques néologismes par deux groupes d'étudiants. 4ème Congrès international de néologie des langues romanes URL: <https://cineo2018.sciencesconf.org/177097>
- Feldman, L.B., Milin, P., 2018. Psycholinguistic studies of word morphology and their implications for models of the mental lexicon and lexical processing. In: Berthiaume, R., Daigle, D., Desrochers, A. (Eds.), *Morphological Processing and Literacy Development: Current Issues and Research*. Routledge, New York, Abingdon, pp. 40–71. <http://dx.doi.org/10.4324/9781315229140>

- Fradin, B., 2015. Blending. In: Müller, P.O., Ohnheiser, B., Olsen, S., Rainer, F. (Eds.), *Handbook of Word-Formation. An international Handbook of the Language of Europe*, Vol. 1. De Gruyter Mouton, Berlin. pp. 386–413, (Chapter 21). <https://halshs.archives-ouvertes.fr/halshs-01428381>.
- Fradin, B., Montermini, F., Plénat, M., 2009. *Morphologie grammaticale et extragrammaticale*. In: Fradin, B., Kerleroux, F., Plénat, M. (Eds.), *Aperçus de morphologie du français*. Presses universitaires de Vincennes, Saint-Denis, pp. 21–45.
- Gardin, B., Lefèvre, G., Marcellesi, C., Mortureux, M.F., 1974. A propos du sentiment néologique. *Langues* 36, 45–52. <http://dx.doi.org/10.3406/lgge.1974.2273>
- Goossens, V., 2009. La polysémie des noms d'affect. *Travaux Neuchâtelois de Linguistique* 50, 147–161. <https://hal.archives-ouvertes.fr/hal-00644518>
- Gérard, C., Bruneau, L., Falk, I., Bernhard, D., Rosio, A.L., 2017. Le logoscope: observatoire des innovations lexicales en français contemporain. In: García Palacios, J., de Sterck, G., Linder, D., Torre del Rey, J., Sánchez Ibanez, M., Maroto García, N. (Eds.), *La neología en las lenguas románicas: recursos, estrategias y nuevas orientaciones*. Peter Lang, Bern. pp. 339–356, <https://hal.archives-ouvertes.fr/hal-01388255>.
- Gérard, C., Kabatek, J., 2012. Introduction: la néologie sémantique en questions. *Cahiers de lexicologie* 100, 11–36. <http://dx.doi.org/10.15122/isbn.978-2-8124-4147-9.p.0011>
- Haspelmath, M., 2009. Lexical borrowing: concepts and issues. In: Martin Haspelmath, M., Tadmor, U. (Eds.), *Loanwords in the World's Languages: A Comparative Handbook*. De Gruyter Mouton, Berlin, pp. 35–54. <http://dx.doi.org/10.1515/9783110218442>
- Jakubiček, M., Kilgarriff, A., Kovář, V., Rychlý, P., Suchomel, V., 2013. The tenten corpus family. In: Hardie, A., Love, R. (Eds.), 7th International Corpus Linguistics Conference CL 2013. UCREL, Lancaster. pp. 125–127, <http://ucrel.lancs.ac.uk/cl2013/>.
- Jamet, D., Terry, A., 2018. Introduction. *Lexis* 12. <http://dx.doi.org/10.4000/lexis.1075>
- Kaczer, L., Bavassi, L., Petroni, A., Fernández, R.S., Laurino, J., Degiorgi, S., Hochman, E., Forcato, C., Pedreira, M.E., 2018. Contrasting dynamics of memory consolidation for novel word forms and meanings revealed by behavioral and neurophysiological markers. *Neuropsychologia* 117, 472–482. <http://dx.doi.org/10.1016/j.neuropsychologia.2018.07.001>
- Kaczer, L., Timmer, K., Bavassi, L., Schiller, N.O., 2015. Distinct morphological processing of recently learned compound words: an ERP study. *Brain Res.* 1629, 309–317. <http://dx.doi.org/10.1016/j.brainres.2015.10.029>
- Kinoshita, S., 2015. Visual word recognition in the Bayesian reader framework. In: Pollatsek, A., Treiman, R. (Eds.), *The Oxford Handbook of Reading*. Oxford University Press, New York, pp. 63–75. <http://dx.doi.org/10.1093/oxfordhb/9780199324576.013.5>
- Klein, D.E., Murphy, G.L., 2001. The representation of polysemous words. *J. Memory Lang.* 45, 259–282. <http://dx.doi.org/10.1006/jmla.2001.2779>
- Klepousniotou, E., Baum, S.R., 2007. Disambiguating the ambiguity advantage effect in word recognition: an advantage for polysemous but not homonymous words. *J. Neurolinguist.* 20, 1–24. <http://dx.doi.org/10.1016/j.jneuroling.2006.02.001>
- Klepousniotou, E., Pike, G.B., Steinhauer, K., Gracco, V., 2012. Not all ambiguous words are created equal: an EEG investigation of homonymy and polysemy. *Brain Lang.* 123, 11–21. <http://dx.doi.org/10.1016/j.bandl.2012.06.007>
- Lakoff, G., Johnson, M., 1980. *Metaphors We Live By*. The University of Chicago Press, Chicago and London.
- Langacker, R.W., 1987. *Foundations of Cognitive Grammar, Vol. I, Theoretical Prerequisites*. Stanford University Press, Stanford.
- Lavale-Ortiz, R.M., 2019. Bases para la fundamentación teórica de la neología y el neologismo: la memoria, la atención y la categorización. *Círculo de Lingüística Aplicada a la Comunicación* 80, 201–226. <http://dx.doi.org/10.5209/CLAC.66608>
- Leminen, A., Smolka, E., Duñabeitia, J.A., Pliatsikas, C., 2019. Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding). *Cortex* 116, 4–44. <http://dx.doi.org/10.1016/j.cortex.2018.08.016>
- Lenth, R.V., 2016. Least-squares means: the R package lsmeans. *J. Stat. Softw.* 69, 1–33. <http://dx.doi.org/10.18637/jss.v069.i01>
- Machery, E., Mallon, R., Nichols, S., Stich, S., 2012. If folk intuitions vary, then what? *Philos. Phenomenol. Res.* 86, 618–635. <http://dx.doi.org/10.1111/j.1933-1592.2011.00555.x>
- Mattiello, E., 2013. Extra-Grammatical Morphology in English: Abbreviations, Blends, Reduplicatives, and Related Phenomena. De Gruyter Mouton, Berlin. <http://dx.doi.org/10.1515/9783110295399>
- Meteyard, L., Davies, A.I.R., 2020. Best practice guidance for linear mixed-effects models in psychological science. *J. Memory Lang.* 112. <http://dx.doi.org/10.1016/j.jml.2020.104092>
- Nunberg, G., 1995. Transfers of meaning. *J. Semant.* 12, 109–132. <http://dx.doi.org/10.1093/jos/12.2.109>
- Pexman, P.M., 2012. Meaning-based influences on visual word recognition. In: Adelman, J.S. (Ed.), *Current Issues in the Psychology of Language, Visual Word Recognition: Meaning and Context, Individuals and Development*. Psychology Press, London and New York, pp. 24–43.
- Plénat, M., 1999. *Prolégomènes à une étude variationniste des hypocoristiques à redoublement en français*. *Cahiers de Grammaire* 24, 183–219.
- Podhorná-Polická, A., Fiévet, A.C., 2018. a. La circulation du néologisme swag: résultats d'une enquête par questionnaires auprès de jeunes de la région parisienne et de nice. In: Jacquet-Pfau, C., Napieralski, A., Sablayrolles, J.F. (Eds.), *Emprunts néologiques et équivalents autochtones: études interlangues*. Wydawnictwo Uniwersytetu Łódzkiego, Łódź, pp. 147–158. <http://dx.doi.org/10.18778/8088-785-5.10>
- Podhorná-Polická, A., Fiévet, A.C., 2018. *La dynamique du français des jeunes: deux périodes à sept ans d'intervalle (1987-1994 et 2010-2017)*. ELAD-SILDA HS1.
- Poix, C., 2018. Neology in children's literature: a typology of occasionalisms. *Lexis* 12. <http://dx.doi.org/10.4000/lexis.2111>
- Pollatsek, A., Bertram, R., Hyönä, J., 2011. Processing novel and lexicalised Finnish compound words. *J. Cogn. Psychol.* 23, 795–810. <http://dx.doi.org/10.1080/204445911.2011.570257>
- Pustejovsky, J., 1995. *The Generative Lexicon*. MIT Press, Cambridge and London.
- R Core Team, 2015. R: A Language and Environment for Statistical Computing. <http://www.R-project.org/>
- Rastle, K., Davis, M.H., 2008. Morphological decomposition based on the analysis of orthography. *Lang. Cogn. Process.* 23, 942–971. <http://dx.doi.org/10.1080/01690960802069730>
- Renouf, A., 2013. *A finer definition of neology in English: the life-cycle of a word*. *Stud. Corpus Linguist.* 57, 177–208.
- Sablayrolles, J.F., 2000. *La néologie en français contemporain: Examen du concept et analyse de productions néologiques récentes*. Honoré Champion, Paris.
- Sablayrolles, J.F., 2003. *Le sentiment néologique*. In: Sablayrolles, J.F. (Ed.), *L'Innovation lexicale*. Champion, Paris, pp. 279–295.

- Sablayrolles, J.F., 2010. Néologisme homonymique, néologisme polysémique et évolution de sens: Pour une restriction de la néologie sémantique. In: Alvès, I.M. (Ed.), *Neologia e neologismos em diferentes perspectivas*. CNPQ, Paulistana. pp. 83–100, <https://halshs.archives-ouvertes.fr/halshs-00736394>.
- Schmid, H.J., 2008. New words in the mind: Concept-formation and entrenchment of neologisms. *Anglia - Zeit. Englisch. Phil.* 126, 1–36. <http://dx.doi.org/10.1515/angl.2008.002>
- Schmid, H.J., 2017. *Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge*. AOA and De Gruyter Mouton, Washington and Berlin.
- Smolka, E., Libben, G., 2017. 'Can you wash off the hogwash?' – Semantic transparency of first and second constituents in the processing of German compounds. *Lang. Cogn. Neurosci.* 32, 514–531. <http://dx.doi.org/10.1080/23273798.2016.1256492>
- Srinivasan, M., Berner, C., Rabagliati, H., 2019. Children use polysemy to structure new word meanings. *J. Exp. Psychol.: Gen.* 148, 926–942. <http://dx.doi.org/10.1037/xge0000454>
- Svorou, S., 1994. *The Grammar of Space*. John Benjamins, Amsterdam <https://www.jbe-platform.com/content/books/9789027276575>
- Štekauer, P., 2002. On the theory of neologisms and nonce-formations. *Aust. J. Linguist.* 22, 97–112. <http://dx.doi.org/10.15122/isbn.978-2-8124-4229-2.p.0083>
- Winter-Froemel, E., 2009. *Les emprunts linguistiques: enjeux théoriques et perspectives nouvelles*. *Neologica* 3, 79–122.
- Xu, Z., 2001. *Le Néologisme et ses implication sociales*. L'Harmattan, Paris. <http://dx.doi.org/10.1093/oxfordhb/9780199324576.013.4>
- Yap, M.J., Balota, D.A., 2015. Visual word recognition. In: Pollatsek, A., Treiman, R. (Eds.), *The Oxford Handbook of Reading*. Oxford University Press, New York, pp. 26–43. <http://dx.doi.org/10.1093/oxfordhb/9780199324576.013.4>
- Yurchenko, A., Lopukhina, A., Dragoy, O., 2020. How meaning similarity influences ambiguous word processing: the current state of the literature. *Front. Psychol.* 11. <http://dx.doi.org/10.3389/fpsyg.2020.02113>