Constructions 16 (2024)
doi 10.24338/cons-659

Individual differences in productivity: intra- and extralinguistic determinants in evaluations of "creative" uses of grammatical patterns[1]

Anouk Van den Stock[1], Anne-Sophie Ghyselen[1], Peter Lauwers[1], Dirk Speelman[2] & Timothy Colleman[1]

[1] Ghent University [2] KU Leuven

ABSTRACT

In light of the partial productivity puzzle (see e.g., Goldberg 2019 for a recent discussion), recent work in Construction Grammar has explored the connection between constructional productivity and linguistic creativity (i.a., Hoffmann 2018, 2019, 2020a; Bergs 2019). While current research into productivity has been mainly concerned with intralinguistic determinants such as type/token frequency and semantic similarity, the present study demonstrates the relevance of including individual, user-related variables as potential extralinguistic determinants of linguistic creativity. Using an acceptability rating experiment focussing on two Dutch argument structure constructions as a case study, we explore individual differences in productivity. The findings indicate considerable inter-individual variation in the extent to which speakers evaluate novel/creative instantiations of the patterns at stake positively or negatively. The results of ordinal regression analyses reveal (i) that participants' ratings are influenced by their social backgrounds, linguistic experiences, and personality traits, and (ii) that intralinguistic and extralinguistic variables are inextricably linked to each other.

1 Introduction

Constructions – in the Construction Grammar (CxG) sense of the term, i.e., conventionalized form-meaning pairings (Goldberg 1995, 2006, i.a.) – "allow us to apply our linguistic knowledge to new situations and experiences" (Goldberg 2019: 2). In the case of clause-level (partially) schematic argument structure constructions (ASCs), for instance, speakers are able to draw on their stored knowledge of a specific schema in order to extend its usage with new verbs – a phenomenon also referred to as syntactic productivity (cf. Barðdal 2008), which is of course not limited to *argument structure* constructions but characterises all schemas with at least one open slot. A vast amount of research within CxG shows that schemas are rarely fully productive, though, i.e., altogether free from lexical constraints on the kinds of items that can be filled into their slots. In other words, constructions typically display partial productivity. Several corpus investigations into the English *way*-construction (Israel 1996; Stefanowitsch & Gries 2005; Perek 2016), for instance, have demonstrated that this pattern occurs with a wide range of verbs, and that it still displays an increasing level of productivity and semantic diversification. Nevertheless, while this pattern regularly occurs with verbs of motion (e.g., *climb*, *squirm*, and *stumble*), Goldberg

[1] This research is part of a larger project on Language Productivity at Work (PI Peter Lauwers; see https://www.languageproductivity.ugent.be/), generously funded by Ghent University's Special Research Fund's Concerted Action Scheme (BOF-GOA2019000501). Special thanks go out to Ludovic De Cuypere (Universiteit Gent, Vrije Universiteit Brussel) for his help with the vector space analysis and to Filip De Fruyt (Universiteit Gent) for his advice on personality research and assessment tools. We are grateful to the two anonymous reviewers for their comments and suggestions, which helped improve the paper.

Corresponding author: Anouk Van den Stock, Ghent University, anouk.vandenstock@ugent.be

(1995, 2019) observes that as the construction is associated with the creation of a path despite difficulty or obstacles, a verb like to walk – denoting an activity that is not particularly challenging– is not readily acceptable as a slot filler (e.g., ?*She walked her way across the room*[2]).

A central issue in the literature on productivity pertains to the question of what actually determines a pattern's degree of productivity. A great body of literature has already described the effects of various language-internal factors on the productivity of constructions, such as type and token frequency, hapax frequency (Baayen 2009; Zeldes 2012), semantic similarity/compatibility (Suttle & Goldberg 2011; Goldberg 2019) and semantic variability/coherence (Goldberg 2006; Barðdal 2008). In sharp contrast, the effects of potential language-external factors have largely been neglected thus far. This is surprising, in a way, since recent work in CxG has described productivity in close relation to the concept of linguistic creativity (Barðdal 2008; Zeschel 2012; Hoffmann 2018, 2019, 2020a, 2022; Bergs 2019; Goldberg 2019), which, it is important to remark, is ultimately "a property of the speaker, and not of the language" (Zawada 2006: 239). As we know from the artistic domain, not all human beings are equally (artistically) creative (Hoffmann 2018). In addition, the pervasive individual differences attested in speaker's grammatical knowledge in work by, most prominently, Dąbrowska and colleagues (e.g., Dąbrowska 2012, 2018, 2019; Dąbrowska & Street 2006; Street & Dąbrowska 2010) demonstrate that different speakers draw on partly different generalisations and store shared constructions at different degrees of schematicity. Consequently, we may also expect important variation in the extent to which individual language users (i) extend constructions creatively in production, and (ii) evaluate creative extensions from other speakers as acceptable. Yet, at this point, as Hoffmann (2018: 263) observes, "the individual differences with respect to verbal creativity are only rarely discussed".

In this paper, we argue that in order to fully grasp the concept of productivity, the effects of individual, user-related variables, as well as their interaction with currently known intralinguistic determinants should be taken into account. After a critical discussion of the existing literature on productivity and its connection to creativity (Section 2.1), we explore which user-related variables can potentially be of interest as extralinguistic determinants of productivity (Section 2.2). Subsequently, these variables are implemented in an online acceptability rating experiment measuring the evaluation of novel/creative instantiations of two Dutch ASCs, namely the *weg*-construction (Verhagen 2002) and the *krijgen*-passive (Colleman 2015). Section 3 motivates the selection of these constructions and describes the design of the experiment. Sections 4 and 5 present and discuss the results of the ordinal regression analyses with respect to the four research questions that will be outlined in Section 3.

2 Intra- and extralinguistic determinants of syntactic productivity

2.1 Productivity and creativity in CxG

So far, (partial) productivity has been mostly approached from a corpus-based perspective (though see e.g., Suttle & Goldberg 2011; Perek 2015: 175-208 for notable exceptions), focusing on the effects of language-internal factors. It has been shown, for instance, that productivity is positively influenced by high type frequency (i.e., the number of unique lexical items that are used in a specific slot of a pattern), as the exposure to many different instantiations of a pattern can make the speaker more confident about its potential to be extended with even more new items (Goldberg 1995, 2006; Bybee & Thompson 1997). In addition, while the frequency with

[2] In the right context, the acceptability of this particular verb-construction combination can improve, though. For example: "[The disabled bride] walked, yes walked, her way down the aisle." Taken from Goldberg (2019: 38).

which a particular lexical item occurs in a specific slot of the construction (i.e., its *token* frequency) is expected to promote its level of entrenchment/familiarity in that construction (Schmid 2015, 2020), high-frequency types tend to be stored autonomously and processed as whole units, making them more resistant to being replaced by new forms created with the regular pattern (Bybee & Thompson 1997).

Apart from frequency-based determinants, semantics play an important role as well. The possibility of using a novel item in a construction, for instance, depends on its degree of semantic similarity to previous usages (Bybee & Eddington 2006; Suttle & Goldberg 2011; Goldberg 2019). Furthermore, the semantic diversity/variability of the lexical items attested in a construction seems to be at least as important as their sheer number, as suggested by Goldberg (2006). Barðdal (2008) unites the frequency-based approach and the semantic dimension and proposes that the productivity of syntactic constructions is a function of the inverse correlation between type frequency and semantic coherence (i.e., the opposite of variability). In this view, the relevance of type frequency for productivity decreases with semantic coherence, in that a pattern attested with a low number of types can still be productive (within a particular semantic domain), as long as these types are highly semantically similar.

Still, it is ultimately the *language user* who needs to find ways to extend constructions with new lexical items. In this respect, Goldberg (2019: 1) notes in her recent monograph on the paradox of partial productivity that "we can be creative in how language is used, but our creativity is constrained in ways that can be hard to articulate." In recent work in CxG, an increasing amount of attention has been paid to the connection between productivity and linguistic creativity (see e.g., a special issue in *Zeitschrift für Anglistik und Amerikanistik* 2018 and a special issue in *Cognitive Semiotics* edited by Hoffmann 2020b, as well as a paper by Bergs 2019). Given the fact that extending stored constructions productively is essentially an act of creativity of the language user, it seems imperative to also consider the question which characteristics of the language user influence their linguistic creativity (see also De Smet 2020, who recommends considering system-external variables, such as aspects of individual creativity and cognition, as potential determinants of individual differences in productivity). In the next section, we explore which individual, user-related variables may potentially be of interest as extralinguistic determinants of productivity.

2.2 Productivity and creativity across individuals

Research in psychology investigating individual levels of general creativity (i.e., not restricted to language)[3] offers a good starting point. It has been shown, for instance, that high levels of general intelligence correlate positively with creativity (Kaufman 2016; Kandler *et al.* 2016). In addition, creativity has been shown to be influenced by personality traits such as Openness and Extraversion (Kandler *et al.* 2016; Jirásek & Sudzina 2020). More specifically, high levels of Openness are associated with characteristics such as intellectual curiosity and a need for new experiences and novel ideas. Extraverted people, moreover, are likely to be relatively more socially active and talkative, tend to engage more easily in risk-taking behaviour and express innovative ideas more easily. With regard to *linguistic* creativity, specifically, Hoffmann (2018, 2019) hypothesizes that more social contact will result in higher exposure to different regional and age-based varieties, which will in turn stimulate the expansion of the constructional network. Finally, exposure to different languages in general, too, may influence a speaker's level

[3] As pointed out by Kandler *et al.* (2016 : 231), psychological research on creativity employs a diversity of definitions, theories, and methods of measurement as the field has not yet converged on a 'explicit and common definition of creativity or what it consists of.' Creativity can be both a product (evaluated by quantity, quality, and usefulness) and a process (an innovative and problem-solving idea), but it may also refer to relatively stable behavioural traits (e.g., personality traits and intrinsic motivation) and cognitive abilities (e.g., intelligence, knowledge, and thinking styles) that are most characteristic for creative persons.

of creativity. In this respect, Kharkurin (2012) shows that multilingualism can act as a facilitator of one's creative potential, as it also impacts cognitive mechanisms that can boost innovative capacities and creative thinking.

Apart from individual levels of creativity, recent work in usage-based grammar highlights the considerable individual differences in speakers' grammatical knowledge and grammatical representations. A broad range of user-related variables have already been explored as potential sources for this variation in work by, most prominently, Dąbrowska and colleagues (see Dąbrowska 2012, 2018, 2019 for overviews). These can be roughly divided into two categories, viz., variables relating to, on the one hand, speakers' linguistic experiences, and, on the other, their cognitive abilities. In a nutshell, a higher amount of exposure to relatively marked linguistic structures – measured through, for instance, the participant's educational level, reading experience, or their performance on vocabulary tasks – will result in better performance on tasks tapping into their knowledge of these specific constructions (e.g., passives, long distance dependencies, and case marking). In a similar vein, it has been shown that the mental representations of speakers (i) working in a specific job field (Verhagen *et al.* 2020) or (ii) with a background in linguistics (Dąbrowska 2010) differ significantly from speakers without this specific background. However, speakers do not just differ in terms of their experience with language, they also differ in cognitive abilities (see e.g., Misyak & Christiansen 2012 on individual differences in statistical learning abilities). In testing speakers' vocabulary size and their knowledge and comprehension of a range of grammatical constructions, Dąbrowska (2018) found that participants' performances were significantly influenced by their nonverbal IQ.

Finally, a long history of variationist sociolinguistic research has abundantly shown that variables relating to the language user's social background are relevant for their language production and perception. These factors include the language user's age, gender, region of origin, level of education or social class (see i.a., Coulmas 1997, 2000; Labov 2001 and Tagliamonte 2011 for overviews), and the speaker's social environment, which is influenced by e.g., (the density of) their social network and their degree of social and linguistic mobility (Coulmas 2003; Milroy 2004; Tagliamonte 2011; Coates 2015). One might hence wonder whether these factors also impact linguistic creativity. Is the association of specific social groups, such as females and adolescents, with more innovative language use reflected in their willingness to extend grammatical constructions with novel lexical items? Similarly, as speakers tend to adjust their language use and norms to "where they live, who they are surrounded by, and who they wish to emulate" (Tagliamonte 2011: 36), does geographic location and linguistic mobility/language contact influence speakers' evaluations of novel/creative instantiations of grammatical patterns? These questions are difficult to answer on the basis of the existing research; they require further empirical scrutiny.

3 A case study on two Dutch Argument Structure Constructions

3.1 Design

In sum, while research on syntactic productivity has abstracted away from social and individual variation, there are good reasons to assume that there might well be important between-speakers variation in the degree to which speakers creatively extend existing schematic constructions. To address this research gap, this study aims to chart the effects of individual, user-related variables on the evaluations of novel/creative instantiations of grammatical constructions, as well as their interplay with currently known intralinguistic variables. While we acknowledge that measuring levels of openness towards productivity from a comprehension

perspective (in a sense 'passive' productivity) is not fully equivalent to productivity in the conventional sense of the concept (i.e., 'active' productivity, from a production perspective), the present approach offers several advantages. By means of a carefully constructed design and controlled test materials, we are able to collect judgements from a large number of speakers, for multiple instantiations of the linguistic phenomena at stake, including very low-frequent patterns. In addition, the survey-based experimental approach allows us to collect fine-grained information about the participating language users, enabling us to explore the effects of each of the potentially relevant individual, user-related variables (Section 2.2) simultaneously. This first exploration of the role of extralinguistic variables on the evaluations of novel/creative uses of grammatical constructions can then serve as a starting point for further (more focused) investigations of these variables in the context of production, from a corpus-based as well as an experimental approach.

More specifically, we set up an online acceptability rating experiment in which participants were asked to evaluate both conventional and unconventional/novel/creative instantiations of two Dutch ASCs, namely the *weg*-construction in (1) (Verhagen 2002, 2003a, 2003b, 2005, 2007; Van Egmond 2006; Pos 2010) and the *krijgen*-passive in (2) (Van Leeuwen 2006; Colleman 2015; Colleman & Rens 2016). As is customary in CxG work on the partial productivity of ASCs, we focus on the *verb* slot. Hence, conventional and unconventional instantiations of the schematic constructions are, for the purposes of the present study, instantiations in which the verb slot is filled by a verb which is or which is not conventionally associated with the construction, respectively.[4] Data were drawn from preliminary corpus investigations into both constructions viz., Oosterlinck (2019) for the *weg*-pattern and Delaby & Colleman (2023) for the *krijgen*-passive. Both studies offer an overview of the attested verbs and their token frequencies in the construction at stake in the Dutch SONAR-corpus (Oostdijk *et al.* 2013). For a comparative corpus study between the alternating *weg*-construction and the transition-to-location (TLC) construction, Oosterlinck (2019) queried the complete SONAR-corpus (500 million words) and attested 1730 instances of the *weg*-pattern (3.46 occurrences per million words). With the aim of investigating preferences in word-order variation in Netherlandic Dutch, Delaby & Colleman (2023) focused on the Netherlandic Dutch newspaper component (59 381 224 words) and found 1010 instances of the *krijgen*-passive in the Netherlandic Dutch newspaper component (17 occurrences per million words). These normalized frequencies indicate that we are dealing with relatively low-frequently occurring constructions.

(1) Hij baande/zocht/toeterde/elleboogde zich een weg door de menigte.
'He made/searched/honked/elbowed his way through the crowd.'

(2) Els kreeg een kaartje aangeboden/opgestuurd/geleverd/toevertrouwd.
'Els was offered/sent/delivered/entrusted a card.' (lit. 'E. got the card offered etc.')

The *weg*-construction [S_i V $REFL_i$ *een weg* PP], which is the Dutch equivalent of the English *way*-construction that was already mentioned above, denotes an event in which the subject referent moves along a path, despite certain obstacles, with the verb typically denoting the means by which the path is created or the manner in which it is travelled. The most typical lexical instantiation, by far (almost 70% of the attestations in our sample), is with the verb *banen* ('to

[4] In theory, it would of course be possible to focus on the lexical filling of other slots, too. For instance, while the verb is pretty run-of-the-mill, there is a degree of unconventionality in the instance of the *krijgen*-passive in (i) in that the subject referent is a building here, while, in the usual case, the *krijgen*-passive denotes a transfer to an *animate* recipient. Such instances have not been included in the experiment.

(i) *Het gebouw kreeg een nieuwe functie toegekend.*
'A new function was attributed to the building.' (lit. The building got a new function attributed).'

pave, to clear or smooth a path'). As described by Verhagen (2003), the exact meaning of *banen* is hard to articulate without referring to its use in the *weg*-construction. In that sense, he continues, the meaning of the verb seems to coincide with the meaning of the construction in which it is mostly used, making it the construction's "default"-verb. Other typical verbs include *zoeken* ('to search'; 7.80%), *vechten* ('to fight'; 3.99%), and *vreten* ('to eat excessively and/or greedily'; 1.91%). With 73 pure hapaxes and an additional 35 dis- or tris legomena out of 134 attested verbs in our sample (81%), the *weg*-construction is very productive.

The Dutch *krijgen*-passive [S *krijgen* NP $V_{past.part}$ (*van/door* NP)] – alternatively the 'receptive construction' or the 'semi-passive' – is a passive counterpart of the active ditransitive construction which enables the speaker to present a transfer from the point of view of the recipient (cf. the active counterpart of [2] with *opsturen* would be *Iemand stuurde Els een kaartje op* 'Someone sent Els a card'). The construction is largely restricted to verbs encoding 'possessional' (e.g., *aanbieden* 'to present'; *betalen* 'to pay) or 'communicative' (e.g., *opleggen* 'to impose'; *toewijzen* 'to assign') transfer. Moreover, even among the class of verbs of transfer, the construction imposes lexical and morphological constraints, as several simplex and highly frequent ditransitive verbs are not regularly combined with *krijgen*, including the prototypical 'transfer of possession'-verb *geven* ('to give') (see Colleman 2015 for details). Consequently, when directly compared to each other, the *krijgen*-passive is less productive in Dutch than the *weg*-pattern. This is also reflected in corpus data: among the 59 verb types attested, there are 12 pure hapaxes and 10 dis- or tris legomena (37%).

The *weg-* and *krijgen*-constructions do not only differ in terms of their degree of productivity but, possibly, also in their level of "extravagance" (see Ungerer & Hartmann 2020). Given its more "salient" and "unconventional" properties, we suspect that the *weg*-pattern is in this case the more "extravagant" construction of the two (see Section 5 for more details). Additionally, the *weg*-pattern has a comparatively more productive counterpart in English while the *krijgen*-passive does not.[5] According to Verhagen (2002; 2003b), for instance, the incidental-action sense of the English *way*-construction (as in *He whistled his way to the front door*; see i.a., Israel 1996; Perek 2018) is not possible for Dutch (?*Hij floot zich een weg naar de voordeur*; Verhagen 2003b: 337; but see Colleman 2020 on sporadic counterexamples in recent language use). This allows us to examine the potential effect of exposure to a construction in a (dominant) neighbouring language. Taking these two Dutch ASCs as a case-study, this study constitutes a first attempt at measuring inter-individual (between-speakers) variation in evaluations of productivity/creativity in grammatical constructions. The following research questions are at the core of our investigation:

RQ1. Is there inter-individual (between-participants) variation in language users' evaluations of novel/creative uses of the selected grammatical constructions?

RQ2. If so, can we relate this inter-individual variation to user-related variables, such as
- i. social background (age, gender, and region),
- ii. linguistic experience (amount and type of exposure),
- iii. personality traits (e.g., extraversion and openness), and
- iv. cognitive abilities (general intelligence)?

Furthermore, we aim to maximally integrate insights and methods from the corpus-based approach most often taken in productivity research. More specifically, we want to examine the effect of several language-internal variables, which will be described in more detail in Section

[5] While the Dutch *krijgen*-passive bears a degree of structural resemblance to the English *get*-passive (see e.g., Givón & Yang 1994; Collins 1996), it has different, much more restricted semantics, being largely restricted to 'transfer' events. Present-day German, by contrast, does display a construction that closely resembles the Dutch *krijgen*-passive, both structurally and functionally, namely the *bekommen*-passive (Leirbukt 1997; Bader 2012).

3.2.2, and moreover, investigate how language-external and language-internal variables interact:

RQ3. Can we find evidence for the role of (i) token frequency, (ii) lemma frequency, and (iii) semantic compatibility as intralinguistic determinants of acceptability ratings of conventional and unconventional uses?

RQ4. What is the interplay between currently known intralinguistic determinants of productivity and currently unexplored extralinguistic determinants of linguistic creativity?

3.2 Materials

3.2.1 Judgement task

In our experiment, participants were asked to evaluate stimulus items on 7-point Likert scales. More specifically, the instructions read *In welke mate vind je dit een goede Nederlandse zin?* ('To what degree do you find this a good Dutch sentence?'). While it is generally acknowledged that the exact formulation of such an instruction, ultimately, has little influence on the results (Cowart 1997; Schütze & Sprouse 2013; Schütze 2019), we explicitly avoided potentially inadequate or confounding content words in our phrasing, such as 'grammatical' (as linguistically trained and linguistically naïve subjects might have a different notion of this concept; see Schütze 2019 for a discussion), 'acceptable' (which may be taken to refer to the content of the stimulus), or 'correct' (which may invoke a contrast between standard language and dialect). The lowest score was represented by the number 1 on the scale and was accompanied by the label *Kan zeker niet* ('Definitely not possible'). The highest score, in turn, was represented by the number 7 on the scale and was accompanied by the label *Kan perfect* ('Perfectly possible)'. The intermediate points of the scale were numbered from 2 to 6 but did not have a verbal label. In order to familiarize the participants with the scale, two example sentences were presented with extensive instructions and the first five stimuli to be rated were filler items. To ensure authentic responses, it was emphasized that there were no right or wrong answers and that the researchers were solely interested in participants' personal linguistic intuitions.

3.2.2 Stimulus items

For the selection of conventional and unconventional instantiations of the constructions at stake, the concept of 'productivity'/'creativity' was operationalised on the basis of corpus data. As mentioned in Section 3.1, we started from already existing corpus investigations of the two selected Dutch ASCs, which provided us with lists of attested verbs for each construction and their token frequencies, i.e., how often a specific verb occurs as a slot-filler in the construction. Following the assumption that items high in token frequency are more entrenched (Schmid 2015, 2020), we used the top 5 most frequently occurring verbs in each construction to represent its most conventional instantiations. On the other side of the frequency spectrum, the group of hapax legomena (73 for the *weg*-pattern and 22 for the *krijgen*-passive) were considered to represent the most "creative" uses of each construction, their one-off character being taken to suggest the status of freshly computed rather than previously stored instantiations of a pattern, as is customary in the literature on (morphological) productivity (Baayen & Lieber 1991; De Smet 2020). In addition, we distinguished a third group of verbs with an intermediate token frequency.

A number of studies, however, have observed that there seems to be a discrepancy or gap between the frequency of an item and its perceived acceptability, the so-called 'frequency/acceptability mismatch' (Kempen & Harbusch 2005; Bader & Häussler 2010; Divjak 2017; Flach 2020). More specifically, these studies observe that corpus frequencies tend

to be a poor predictor of judgements, especially when it comes to rare or unattested items at the lower end of the frequency spectrum. Therefore, for the selection of hapaxes to be included in the test materials, two additional parameters were taken into account (see also Divjak 2017 who takes into account similar variables). First, the verb's general lemma frequency was considered, as this has been shown to be a crucial predictor in word recognition tasks and in memory performance (Brysbaert *et al.* 2011, 2018). As for how this parameter may affect the ratings of conventional and unconventional uses of constructions, two competing hypotheses are plausible. First, as high frequency words tend to be better entrenched in the minds of language users, a verb's high overall lemma frequency might contribute to the sense of familiarity of the construction in which it occurs, even if that verb has a low contextual occurrence in that construction (i.e., low token frequency). Consequently, higher general frequencies of the verb might influence the ratings positively (= hypothesis 1). On the other hand, for an overall *high* frequency verb, a very low frequency of occurrence in a given construction can be seen as more significant than in the case of a verb with an overall *low* frequency (which is the basic idea behind Stefanowitsch & Gries' 2003 method of simple collexeme analysis, cf. esp. their idea of significantly repelled verbs, i.e., verbs which occur *less* often in a given construction than one would expect on the basis of their overall frequency). In this respect, then, higher overall frequencies of verb that are hapaxes in the target construction might *negatively* influence the ratings (= hypothesis 2). Frequency information was drawn from Zipf-transformed frequencies found in the SUBTLEX-NL database (Keuleers *et al.* 2010). The hapaxes were first divided on the basis of their semantic compatibility to the construction (see below), after which the lemma frequencies of potential verbs in the eligible semantic cluster were compared in relation to each other to distinguish between verbs higher and lower in lemma frequency. Where possible, we tried to restrict the selection to a Zipf-value cut-off point of above 3.5 for high frequent verbs and below 3.0 for low frequent verbs.

Second, given that the success rate of a newly introduced item in a construction is highly dependent on its semantic similarity to previous usage (Section 2.1), or, in other words, its 'semantic compatibility' to the dominant semantic fillers classes attested in the construction, a distributional semantics approach was employed to model the semantic similarity of a hapax to all other attested verbs in the construction (following work by i.a., Perek 2016, 2018). More specifically, we distinguished between verbs high vs. low in semantic compatibility by selecting them from different types of clusters, which originated from a cluster analysis using the PAM-method (partitioning around medoids). To do this, we used the Dutch COW-corpus (Schäfer 2015; Schäfer and Bildhauer 2012) to construct vector-space models according to the Bag-of-Words paradigm (Jurafsky & Martin 2024). For each construction, a distance matrix based on the cosine dissimilarity scores between the attested verbs in that construction and the 58K most frequent lemmas in the Dutch COW-corpus served as the basis for a cluster analysis. The optimal amount of clusters for each construction was determined by comparing the average silhouette widths for different solutions, which resulted in 70 clusters for the *weg*-pattern and 35 clusters for the *krijgen*-passive. Verbs in the same cluster are semantically close; verbs far away from each other in the clusters are semantically different. For the choice of hapaxes high in semantic compatibility, we focused on verbs from dense clusters consisting of at least five verbs, and ideally, including at least one verb from the top 10 most frequent verbs in the construction. In addition, we took into account the combined token frequencies of the individual verbs in the cluster. As an illustration, consider two verbs high in semantic compatibility selected for the *weg*-pattern in Table 1, *breken* ('to break') and *stoten* ('to thrust'). These two verbs appear in the same cluster, which consists of seven verbs, including one top-10 most frequent verb (*schieten* 'to shoot'). Adding up the token frequency of each individual verb in the construction results in a combined token frequency of 34, which is the third highest number of all clusters. For the sake of comparability, a pair consisting of a verb high in lemma frequency (in this case *breken*) and a verb low in lemma frequency (in this case *stoten*) was selected from the same semantic cluster. In

contrast, hapaxes low in semantic compatibility were clustered in isolation, that is, they formed a cluster of their own (e.g., *gokken* 'to gamble' and *stuntelen* 'to stumble'), which we take to imply that they are less semantically similar to the other attested verbs in the construction than hapaxes that were part of a more dense cluster.

Consequently, the experiment included six categories of experimental target items per construction, as illustrated in Table 1. Two verbs were selected per category, resulting, in theory, in 12 experimental items per construction. However, due to the smaller amount of attested hapaxes for the *krijgen*-passive, only one item could be selected for each subcategory of hapaxes for this construction. To compensate for this, two additional experimental items were constructed containing two verbs that were *not* attested in the corpus sample at all, namely *weigeren* ('to deny'*)* and *afpakken* ('to take away'). As the *krijgen*-passive denotes a transfer and these are verbs of refusal and taking away, respectively, these verbs have been described in previous studies as semantically 'incompatible' with the underlying meaning of the construction, but not *altogether* unattested, i.e., occurring highly sporadically in large corpora (see Clement & Glaser 2014; Colleman 2015 for details). When converting the selected hapaxes and unattested verbs to stimulus items, we ensured that the experimental sentence was modelled on a real-language example from the corpus-based dataset, though sentences were often simplified and standardised. For instance, based on the basic syntactic structure of each pattern (i.e., [S V *zich een weg* PP] for the *weg*-construction and [S *krijgen* NP V] for the *krijgen*-passive), we restricted the total length of all experimental sentences to maximally 8 to 12 words. In addition, we aimed to include only lexically high-frequent content words in the other open slots of the construction. A complete list of the experimental items per construction can be found in Appendix 1A.

TOKEN FREQUENCY OF THE VERB	**LEMMA FREQUENCY**	**SEMANTIC COMPATIBILITY**	**VERBS *WAY*-CX**	**VERBS *KRIJGEN*-CX**
HIGH			*banen* ('to clear a path') *vechten* ('to fight')	*aanbieden* ('to offer') *opleggen* ('to impose')
INTERMEDIATE			*schieten* ('to shoot') *hakken* ('to cut')	*voorleggen* ('to present') *opsturen* ('to send')
LOW	High	High	*breken* ('to break') *huilen* ('to cry')	*bieden* ('to offer')[6]
	High	Low	*gokken* ('to gamble') *bombarderen* ('to bomb')	*belonen* ('to reward')
	Low	High	*stoten* ('to thrust') *grommen* ('to growl')	*doorbetalen* ('to continually pay')
	Low	Low	*stuntelen* ('to stumble') *liften* ('to hitchhike')	*inpeperen* ('to preach')
UNATTESTED				*weigeren* ('to refuse') *afpakken* ('to take away')

Table 1. Overview of the types of experimental stimuli.

Given the fact that participants were asked to rate ten or twelve highly similar sentences (i.e., with a comparable syntactic structure and reoccurring lexical items such as the combinations *zich een weg* in the *weg*-pattern and the verb *krijgen* in the passive-cx), we had to ensure that there were sufficient filler items to distract them from the experimental items. Therefore, the ratio between experimental items and fillers was 1:3, resulting in 72 filler items. Two groups of ten fillers were constructed to ensure that the participants used the full extent of the Likert scale for the evaluation of the stimuli: (i) perfectly acceptable items (i.e., both syntactically and semantically very plausible), (ii) grammatically unacceptable/incorrect/implausible items (e.g., *Er wordt door de gasten om 7 uur gearriveerd* 'There will be arrived by the guests at 7 o'clock'). Additionally, we included ten fillers instantiating various grammatical structures unrelated to the *krijgen*- and *weg*-constructions which were of interest to the researches and for which we suspected variation in judgements would occur (based on varying acceptability reported in

[6] *Bieden* and *aanbieden*, both of which can be glossed as 'to offer' are morphological variants, the latter being a separable complex verb with the particle *aan* as its first element (lit. 'on-offer'). The *krijgen*-passive has a certain preference for complex over simplex verbs (cf. Colleman 2015), and *aanbieden* is among the very top verbs in terms of attested token frequency (and is included as such in the experiment) whereas *bieden*, though of course semantically highly compatible with the construction, too, is much more infrequently attested.

previous research, in most cases; e.g., *Thomas zei daarstraks hij kwam morgen zeker* 'Thomas said earlier he would come tomorrow definitely'). An overview of these three groups of fillers can be found in Appendix 1C. The remaining 42 filler sentences served as 'pure' distractors and were designed to either syntactically, lexically, or semantically resemble the experimental items without instantiating the very same constructions.

Ultimately, the acceptability rating experiment consisted of 94 Dutch sentences (i.e., 22 experimental items and 72 fillers). Stimulus items were presented in semi-randomized order, guaranteeing that each experimental item was followed by at least two fillers. Each sentence was presented in isolation and participants were not allowed to return to previously evaluated items.

3.2.3 User-related variables

To collect detailed information about their social backgrounds and linguistic experiences, participants filled out an extensive socio-biographic questionnaire (see Appendix 2). Apart from their age, gender, and region of growing up, participants were asked to indicate with speakers from which provinces in Belgium and/or The Netherlands they frequently (i.e., at least on a monthly basis) interacted, which gives us an indication of participants' communicative radius (cf. social and linguistic mobility as often-used sociolinguistic variables). To further gauge their linguistic experience, participants were asked to specify the number of years spent in full-time education (cf. Dąbrowska 2018), including all stages from primary education to higher education. In general, we can expect that participants who spent more years in full-time education have generally read more frequently and more diversely (see Dąbrowska 2018 for a correlation between education and print exposure; $r = 0.47$***), and thus, are likely to have been more exposed to the patterns at stake (amount of linguistic experience) – especially since we are dealing with relatively low-frequently occurring ASCs (see Section 3.1). With respect to the type of linguistic experience, participants indicated whether or not they received a higher education that focused on language/linguistics. Finally, participants evaluated their own proficiency in a number of neighbouring languages of Dutch on a 10-point Likert scale, namely English, French, German, Italian, and Spanish. Participants then moved on to the acceptability rating experiment.

In the third part of the survey, participants filled out a Dutch adaptation (Denissen *et al.* 2019) of a standardized personality test called the Big Five Inventory 2 (BFI-2; Soto & John 2017). The BFI-2 questionnaire is a revised version of the Big Five Inventory (BFI) developed by Soto & John (2009) and is designed to measure personality on the basis of the 'Big Five' personality domains, namely Extraversion, Agreeableness, Conscientiousness, Open-Mindedness, and Negative Emotionality, and their respective personality 'facets' (Goldberg 1990; John *et al.* 2008; McCrae & Costa 2008):

1. EXTRAVERSION: Sociability, Assertiveness, Energy Level
2. AGREEABLENESS: Compassion, Respectfulness, Trust
3. CONSCIENTIOUSNESS: Organization, Productiveness, Responsibility
4. OPEN-MINDEDNESS: Intellectual Curiosity, Aesthetic Sensitivity, Creative Imagination
5. NEGATIVE EMOTIONALITY: Anxiety, Depression, Emotional Volatility

The fourth and final part of the survey consisted of a 40-item multiple-choice Dutch vocabulary test (Vander Beken *et al.* 2018). As receptive vocabulary has been shown to significantly correlate with general intelligence ($r = 0.43$, $p < 0.001$; Dąbrowska 2018), the results of the vocabulary test can be used as a proxy for general intelligence.

3.3 Participants

To recruit as many Dutch native speakers as possible, the survey was distributed by means of mailing lists, social media and various online channels that are themed around the Dutch language (e.g., https://neerlandistiek.nl/; https://taalunie.org/).[7] In total, 717 native speakers of Dutch fully completed the survey. Seventeen responses were excluded from the analysis for one of two reasons. First, five participants indicated to have been diagnosed with dyslexia. These participants were excluded as dyslexia might have influenced their processing of the stimuli. Second, twelve participants indicated to belong to a group the sample size of which turned out to be too small for inclusion in the statistical models (e.g., participants who did not identify as either 'male' or 'female'; participants who did not grow up in either Belgium or The Netherlands). Ultimately, a dataset of 700 responses was obtained for statistical analyses. Table 2 shows the distribution of the sample with respect to the main socio-biographic variables. For information on the distribution of the remaining user-related variables, we refer to Appendix 3.

Age (continuous)	Range: 20 – 91 years old $\mu = 57.76$, $\sigma = 18.47$
Gender (binary)	Female: 356 (50.86%) Male: 344 (49.14%)
Region of origin (binary)[8]	Belgium: 305 (43.57%) The Netherlands: 395 (56.43%)
Background in linguistics (binary)	Yes: 346 (49.43%) No: 354 (50.57%)

Table 2. Distribution counts of the main socio-biographic variables (n = 700).

These superficial group counts hide a number of underlying sample imbalances. Sample analyses show, for instance, that in the dataset older participants have a higher chance of being male and originating from The Netherlands, whereas younger participants are more likely to be females coming from Belgium. These sample imbalances do not lead to empty cells in the cross tabulations of the categorical variables. To account for these imbalances in the analyses, the variables AGE[9], GENDER and REGION are always included together as predictors in the statistical models.

3.4 Statistical analyses

As Likert-type responses constitute an ordinal level of measurement, ordinal regression analyses were carried out. More specifically, we ran Cumulative Link Mixed Models using the *clmm*-function from the *ordinal*-package (Christenssen 2019a; see also Christenssen 2015, 2018, 2019b) in the R software environment (www.r-project.org). Participants and items were included as random intercepts. No random slopes were included in the model, as this led to convergence issues. We started from a maximal model with (i) all user-related variable of interest (see Sections

[7] The only restriction was that participants had to be of age. As the full experiment would take around 30 to 45 minutes to complete, thirty vouchers of €20 valid in the Dutch webshop *bol.com* were randomly given away amongst the participants as an incentive to fill out the survey.

[8] Participants also specified the province(s) in Belgium and/or The Netherlands they grew up in, but to reduce the amount of cross-tabulations between the included categorical variables, we limited the analyses to the binary distinction between the two countries.

[9] This formatting will be used when referring to the effects of a specific variable in the statistical models. For the readability of the text, these labels have been translated from Dutch to English, while the output of the statistical models include Dutch labels. The translation equivalents for each user-related predictor can be found in Appendix 3.

3.1 and 3.2.3), (ii) the relevant linguistic variables (i.e., type frequency; lemma frequency and semantic compatibility for hapaxes), and (iii) all two-way interactions between the linguistic variables and user-related variables. We then used a stepwise backward elimination approach via *anova* to determine the optimal structure of the model. The ordinal modelling was conducted for each construction (i.e., the *weg*-construction and the *krijgen*-passive) separately, and moreover, at several levels of analysis. As the hapaxes were selected on the basis of two additional parameters, the creative instantiations had to be kept separate in order to be able to explore the effects of LEMMA FREQUENCY and SEMANTIC COMPATIBILITY. However, in order to investigate the effects of the third intralinguistic variable TOKEN FREQUENCY, we also ran models for all experimental items together, i.e., both conventional as well as unconventional instantiations. In addition, the same analyses were conducted for several other clusters of (non-target) stimulus items. This allowed us to check whether the observed patterns were restricted to evaluating creative and/or conventional instantiations of both constructions, or whether they might be more general effects related to the task of acceptability rating as such. The same modelling approach was therefore applied to the following groups of stimuli:

a) More conventional instantiations of the construction that contain a verb with high or intermediate token frequency.
b) Individual experimental items, which also allowed for a more fine-grained analysis of the effects of the individual/user-related variables. Note that at this level of analysis, the random effects of participants and items no longer apply, and, as a consequence, we ran regular Cumulative Link Models by means of the *clm*-function.
c) Filler items (see Section 3.2.2): the group of (i) perfectly acceptable items, (ii) grammatically unacceptable items, and (iii) grammatical structures of interest to the researchers.

4 Results

4.1 Inter-individual variation

Visualisations of the distribution of the ratings per experimental item reveal considerable inter-individual variation in the evaluations of the unconventional/novel/creative instantiations of the constructions at stake (RQ1). Figure 1 shows the distribution of the ratings of four experimental items – two for each construction. The stacked barplot at the top represents the ratings of the experimental item from the *weg*-pattern that contains its most frequently occurring verb, namely *banen* ('to pave, to clear or smooth a path'). Unsurprisingly, most participants agree that this is a perfectly acceptable sentence in Dutch. In contrast, there is little agreement on the acceptability of a sentence such as *De speler gokte zich een weg door de Verenigde Staten* ('The player gambled their way through the United States'), which is an instantiation of the construction with a hapax (i.e., *gokken* 'to gamble').[10] This is reflected in the second stacked barplot: whereas around a quarter of the participants judges this utterance to be perfectly acceptable in Dutch Individual (27.57%), another fourth of the participants categorizes this sentence as completely impossible in Dutch (20.29%). Around half of the participants fall somewhere in between these two extremes. A similar picture emerges for the *krijgen*-passive, as can be observed in the final two barplots in Figure 1. The distribution of ratings for the remaining experimental items can be found in Appendix 1B.

[10] Note that this instance potentially resembles the 'incidental activity'-type that is claimed, by Verhagen (2005), to be impossible in Dutch, rather than the 'means' or 'manner' type (i.e., it can be understood as denoting a situation in which the player moved through the United States *while* gambling). A more standard '*by* gambling'-interpretation is not entirely impossible either, though, without further context).

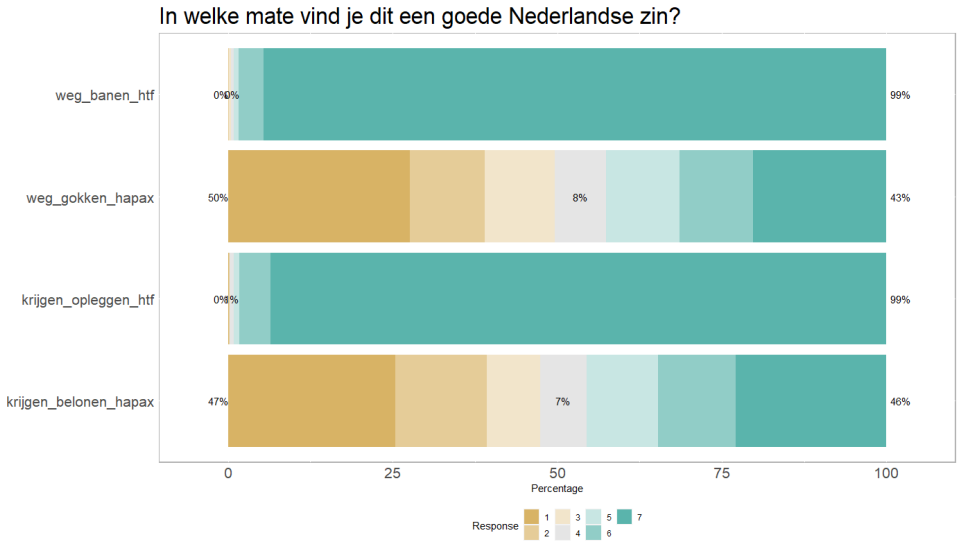

Figure 1. Distribution of ratings of four experimental items.

Tables 3 and 4 present the output of the final (i.e., after stepwise backward elimination) ordinal regression models for the ratings of the creative instantiations of the *weg*-pattern and the *krijgen*-passive, respectively. The full output of the statistical analyses and the visualisations of the significant effects can be accessed https://osf.io/mye3a/.

	estimate	std. error	p-value
Age	0.001246	0.005092	0.806680
Gender-FEMALE	-0.245634	0.164402	0.135147
Region of origin-NL	-0.324256	0.150394	0.031080*
Education_years	0.025903	0.019056	0.174049
Background_linguistics-NO	-0.0296173	0.137197	0.030870*
Communicative radius	0.021893	0.049767	0.660006
English proficiency	0.160343	0.048489	0.000944***
German proficiency	-0.017356	0.034841	0.618380
Extraversion	-0.089474	0.118750	0.451166
Conscientiousness	-0.257423	0.121035	0.033433*
Agreeableness	-0.037304	0.146840	0.799462
Open Mindedness	0.058861	0.126792	0.642480
Negative Emotionality	-0.060025	0.103879	0.563373
Semantic compatibility-LOW	0.436813	0.787734	0.579224
Lemma frequency-LOW	1.139022	0.830116	0.170025
Age:semantic compatibility-LOW	-0.014889	0.003089	1.43e-06***
German proficiency: semantic compatibility- LOW	-0.062394	0.023294	0.007395**
Gender-FEMALE: semantic compatibility- LOW	-0.224532	0.112218	0.045407*
Age: lemma frequency- LOW	-0.008527	0.003157	0.006917**
Education_years: lemma frequency- LOW	-0.032468	0.013970	0.020124*
Communicative radius: lemma frequency- LOW	0.077209	0.037280	0.038351*

English proficiency: lemma frequency-LOW	-0.067415	0.033457	0.043909*

Table 3. Effects for the ratings of the creative instantiations of the *weg*-cx.

	estimate	std. error	p-value
Age	0.01317	0.004828	0.006372**
Gender-FEMALE	-0.07205	0.1611	0.654634
Region of origin-NL	0.6758	0.1660	0.0000466***
Education_years	0.03504	0.01936	0.070345
Background_linguistics-NO	0.02998	0.1134	0.791565
Communicative radius	0.001604	0.03794	0.966280
English proficiency	-0.005269	0.03536	0.881550
German proficiency	0.004052	0.02519	0.872214
Extraversion	-0.00009289	0.09684	0.999235
Conscientiousness	-0.1624	0.09246	0.079098
Agreeableness	-0.1504	0.1126	0.181448
Open Mindedness	0.04817	0.09804	0.623238
Negative Emotionality	-0.09171	0.1091	0.400683
Semantic compatibility-LOW	-0.3200	1.002	0.749462
Semantic compatibility - UNATTESTED	-0.2836	1.101	0.796777
Lemma frequency-LOW	2.054	.09997	0.039887*
Age: semantic compatibility -HIGH vs. UNATTESTED -LOW vs. UNATTESTED	 -0.01941 -0.0250335	 0.005471 0.0051023	 0.000388*** 9.28e-07***
Education_years:semantic compatibility -HIGH vs. UNATTESTED -LOW vs. UNATTESTED	-0.05435 -0.0464082	0.02285 0.0213537	0.014377* 0.02976*
Negative Emotionality -LOW vs. UNATTESTED	 0.3397068	 0.1146834	 0.00306**
Gender-FEMALE: semantic compatibility -HIGH vs. UNATTESTED	-0.3885	0.1870	0.037769*
Region of origin-NL: semantic compatibility -HIGH vs. LOW -HIGH vs. UNATTESTED -LOW vs. UNATTESTED	 -0.7733 -1.512 -0.7521399	 0.1769 0.1905 0.1872518	 0.000123*** 2.09e-15*** 5.90e-05***
Communicative radius: lemma frequency-LOW	0.1261	0.0583	0.021456*
Conscientiousness: lemma frequency-LOW	0.3040	0.1260	0.015813*
Region of origin-NL: lemma frequency-LOW	-0.3396	0.1685	0.043851*
Background_linguistics: lemma frequency-LOW	-0.4652	0.1544	0.002584**

Table 4. Effects for the ratings of the creative instantiations of the *krijgen*-cx.

The tables clearly show that various effects are at play, of both intralinguistic variables (see e.g., an effect of LEMMA FREQUENCY for the *krijgen*-passive) and extralinguistic variables (e.g., an effect of REGION OF ORIGIN for both constructions). Importantly, however, several of the effects are only visible in the interactions[11]. For instance, the influence of AGE for the *weg*-cx is only significant in its interaction with the intralinguistic variables (i.e., LEMMA FREQUENCY and SEMANTIC COMPATIBILITY). Similarly, in the case of the *weg*-pattern, the two intralinguistic variables only seem to influence the ratings in the interaction with a few of the extralinguistic variables. In general, we can therefore conclude that a broad range of intra- and extralinguistic factors should be taken into account when considering the evaluations of the creative instantiations of the patterns at stake. The significant effects will be described in more detail in Sections 4.2-4.3.

Prior to that, it should be observed that the marginal and conditional R^2-coefficients[12] for the models call for some nuance. With conditional R^2-values of 0.522 for the *weg*-cx and 0.635 for the *krijgen*-cx, we can infer that a fair amount of variation in the data is accounted for by the models. However, the marginal R^2-coefficients (0.048 for the *weg*-cx; 0.467 for the *krijgen*-cx) also indicate that a considerable proportion of that variation cannot be attributed to the linguistic and/or user-related variables entered in the model as fixed effects, but is caused by other differences between participants and items. The discrepancy is especially remarkable in the case of the *weg*-pattern, which suggests that even though more significant fixed effects are visible when compared to the *krijgen*-passive, the actual contributions of these predictors are much smaller. Detailed ad hoc analyses of the variance structure of the different models described in this paper (i.e., for the creative instantiations, the conventional instantiations, and all critical items together) reveal that a large amount of the variation that is not captured by the fixed effects seem to reside in individual differences between the participants, which is also reflected in the intercepts of the random effects in the models (Table 5).

	PARTICIPANTS (*SURVEYID*) **VARIANCE (STD.DEV.)**	**ITEMS (CRITICALITEM) VARIANCE (STD.DEV.)**
***WEG*-CX**	2.2957 (1.5152)	0.9622 (0.9809)
***KRIJGEN*-CX**	0.9414 (0.9702)	0.5791 (0.7610)

Table 5. Intercepts for the random effects in the models for the creative instantiations.

4.2 Extralinguistic determinants

4.2.1 Social background

Concerning the social background of the speakers, the analyses indicate that the age, gender and region of origin of the speaker all significantly influence the rating behaviour. With respect to age, the effects are different for the two constructions, though. For the *krijgen*-passive, the ordinal regression model reveals a significant positively correlated main effect for the numerical

[11] In addition, it is worth pointing out that the visualisations of the significant fixed effects and the significant interactions (see https://osf.io/mye3a/) show that the majority of the effects emerge mostly at the end points of the 7-point Likert scale. For instance, with respect to the effect of region of origin on the evaluations of the creative items of the *weg*-pattern, participants from the Netherlands have a higher chance of providing a rating of 1, whereas participants from Belgium have a higher chance of providing a rating of 7. The difference in ratings is not as pronounced for the intermediate points of the scale (2 to 6). In general, however, we can conclude that participants from the Netherlands are more critical in their rating behaviour.

[12] The R^2 is the model's coefficient of determination and corresponds to the proportion of variance in the dependent variable that is explained by the independent variable(s). In case of mixed-effects models, the marginal R^2 considers only the variance of the fixed effects (without the random effects), while the conditional R^2 takes both the fixed and random effects into account (i.e., the total model).

predictor AGE ($p = 0.006$). This means that older participants give comparatively higher ratings for creative instantiations of the *krijgen*-passive, whereas younger participants tend to be more critical in their evaluations. Yet, for the *weg*-construction, the effect of AGE runs in the opposite direction, though it is only visible in the interactions with the two linguistic variables (i.e., LEMMA FREQUENCY and SEMANTIC COMPATIBILITY). More specifically, the visualisations show that for items that contain a hapax low in lemma frequency ($p = 0.007$) and for items that contain a hapax low in semantic compatibility ($p < 0.001$), older participants provide significantly lower ratings and thus tend to be *less* tolerant towards creative uses of the *weg*-pattern than younger participants. Interestingly, however, the analyses of the more conventional instantiations of both constructions (i.e., containing a highly or intermediately frequent verb), reveal yet another picture. Here, for the *krijgen*-passive, a significant negative correlation emerges ($p = 0.003$), signalling that older participants give lower ratings to the more conventional instantiations of the construction. The interaction with TOKEN FREQUENCY shows that, for the items of the *weg*-construction that contain a verb of intermediate token frequency, older participants are significantly more likely to provide higher ratings ($p = 0.014$). The effects with respect to AGE are summarized in Table 6 and will be further discussed in Section 5.

	CONVENTIONAL INSTANTIATIONS	**CREATIVE INSTANTIATIONS**
***WEG*-CX**	Interaction with TOKEN FREQUENCY: Positive correlation for verbs with intermediate token frequency (vs. high token frequency)	Interactions with SEMANTIC COMPATIBILITY and LEMMA FREQUENCY: Negative correlation for verbs low in compatibility and verbs low in lemma frequency
***KRIJGEN*-CX**	Main effect: Negative correlation	Main effect: Positive correlation

Table 6. The effects of age on the ratings of the experimental items. Included effects are significant.

Participants' gender and region of origin are found to influence their ratings, too. As for GENDER, we can observe identical effects for both constructions at the level of the interactions with the linguistic variable SEMANTIC COMPATIBILITY, that is, female participants tend to evaluate the creative instantiations of both constructions that contain a hapax low in semantic compatibility significantly lower than male participants ($p = 0.045$ for the *weg*-cx – see Figure 2; $p = 0.038$ for unattested verbs vs. hapaxes high in semantic compatibility and $p = 0.068$ for hapaxes low vs. high in semantic compatibility for the *krijgen*-cx – see Figure 3). A similar effect can be observed in the interaction with items that contain a hapax low in lemma frequency, though it is not significant ($p = 0.096$ for the *weg*-cx; $p = 0.092$ for the *krijgen*-cx). It is not entirely clear, however, whether this effect pertains solely to evaluating creative instantiations of grammatical constructions, as female participants rate instantiations of the *weg*-pattern that contain a highly or intermediately frequent verb significantly lower than male participants as well ($p = 0.026$).

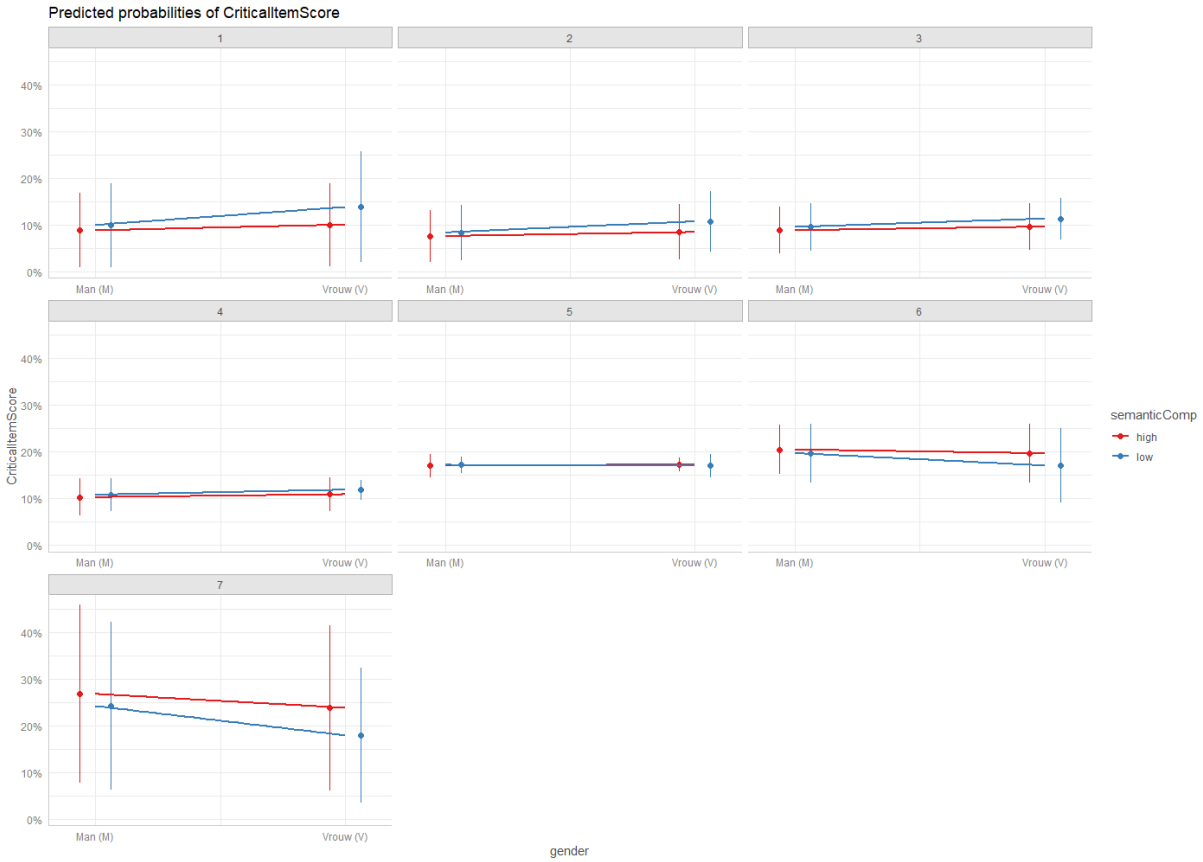

Figure 2. Visualisation of the interaction between gender and semantic compatibility for the *weg*-construction.

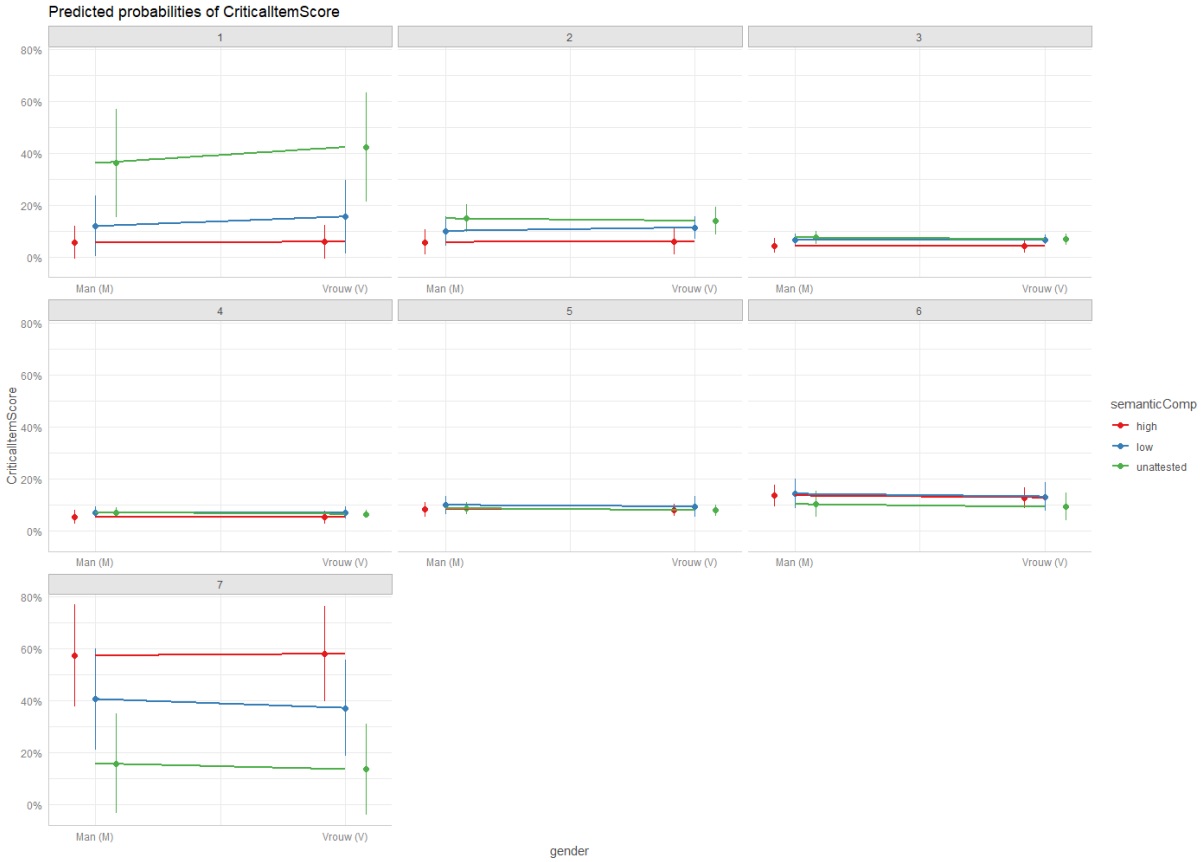

Figure 3. Visualisation of the interaction between gender and semantic compatibility for the *krijgen*-construction.

With respect to the participant's REGION OF ORIGIN, the analyses for both constructions indicate that participants from the Netherlands provide significantly lower ratings for the creative instantiations than participants from Belgium ($p = 0.031$ for the *weg*-cx; $p < 0.01$ for the *krijgen*-cx). For this variable, the effect is not significant in case of the more conventional instantiations, indicating that participants from the Netherlands specifically evaluate instances of linguistic creativity more critically.

More evidence for the effects of both gender and region of origin can be found at the level of the filler items, where we can observe that both female participants and participants from the Netherlands gave significantly lower ratings to items from the category with grammatical structures of interest to the researchers ($p < 0.001$ for both variables) and the grammatically unacceptable fillers ($p = 0.007$ for GENDER; $p < 0.001$ for REGION). The effect is not significant for the category of perfectly acceptable fillers, though, which leads us to believe that both females and speakers of Netherlandic Dutch tend to be more critical in their ratings when presented with stimuli that somehow deviate from the grammatical expectations a bit.

4.2.2 Linguistic experience

Following the literature on individual differences in grammatical knowledge and grammatical representations (Section 2.2), we distinguish between the potential effects of the participants' amount of linguistic experience (measured through the number of years participants spent in full-time education) and their type of linguistic experience (whether or not participants have a background in linguistics) on their evaluations of creative uses of the constructions at stake. In general, we hypothesize that more linguistic experience promotes more tolerant evaluations of stored grammatical patterns and, by extension, creative uses of those, which can be attributed to a greater level of familiarity with those structures (see Verhagen *et al.* 2018 for evidence of an effect of usage frequency on representational strength/entrenchment levels measured through familiarity ratings).

With respect to the type of linguistic experience, we find a very consistent pattern that participants with a training in languages/linguistics rate the experimental items significantly higher than participants without a specific background in linguistics. The analyses reveal main effects of EDUCATION_LINGUISTICS for both the creative instantiations ($p = 0.031$) and the conventional instantiations ($p = 0.030$) of the *weg*-pattern and for the conventional instantiations of the *krijgen*-passive ($p < 0.001$). The same holds for the creative instantiations of the *krijgen*-construction, but only in the case of hapaxes low in lemma frequency, as shown in an interaction with LEMMA FREQUENCY ($p = 0.003$). We can detect the same effect for the category of perfectly acceptable fillers ($p < 0.001$), but not for the grammatically unacceptable fillers ($p = 0.156213$) or the grammatical structures of interest to the researchers ($p = 0.19099$). Given these observations, it is important to bear in mind that this may be a more general pattern and that speakers with specific training in languages/linguistics may be more tolerant in their evaluations of any grammatical pattern. These findings corroborate Dąbrowska's (2010) finding that linguists' intuitions about and judgements of grammatical structures diverge from those of "naïve" speakers (see also Schütze & Sprouse 2013; Schütze 2019).

The data also demonstrate that participants' amount of linguistic experience affects their judgements, though we wish to acknowledge the narrow and relatively one-sided manner in which it was operationalized. While we can expect that participants who spent more years in full-time education have generally read more frequently and more diversely, and thus, are likely to have been more exposed to the patterns at stake (Section 3.2.3), other factors might also be at play (e.g., assumed higher levels of general intelligence, creativity, and openness; cf. Section 2.2 for a discussion of the complex correlations between these variables). That being said, the analyses display a significant positive correlation with EDUCATION_YEARS for the creative instantiations of the *weg*-pattern that contain a hapax high in lemma frequency. In addition, as

shown in Figure 4, the interaction with SEMANTIC COMPATIBILITY reveals an intriguing pattern for the creative instantiations of the *krijgen*-passive. More specifically, the visualisations show the same positive trend for the items that contain a hapax high or low in semantic compatibility. However, a significant negative correlation arises for the group of unattested verbs. As pointed out in Section 3.3.2, the two verbs that make up this category – namely *weigeren* ('to deny') and *afpakken* ('to take away') – have been previously described as being semantically 'incompatible' with the underlying meaning of the construction. In this sense, more experience with a specific construction also seems to entail a greater awareness of a pattern's constraints, which negatively affects language users' judgements of instantiations that stretch the limits of productivity. Similarly, we can recognize the same pattern at the level of the fillers, that is, a significant positive correlation in case of perfectly acceptable sentences ($p = 0.009$) and a significant negative correlation in case of grammatically unacceptable structures ($p = 0.043$).

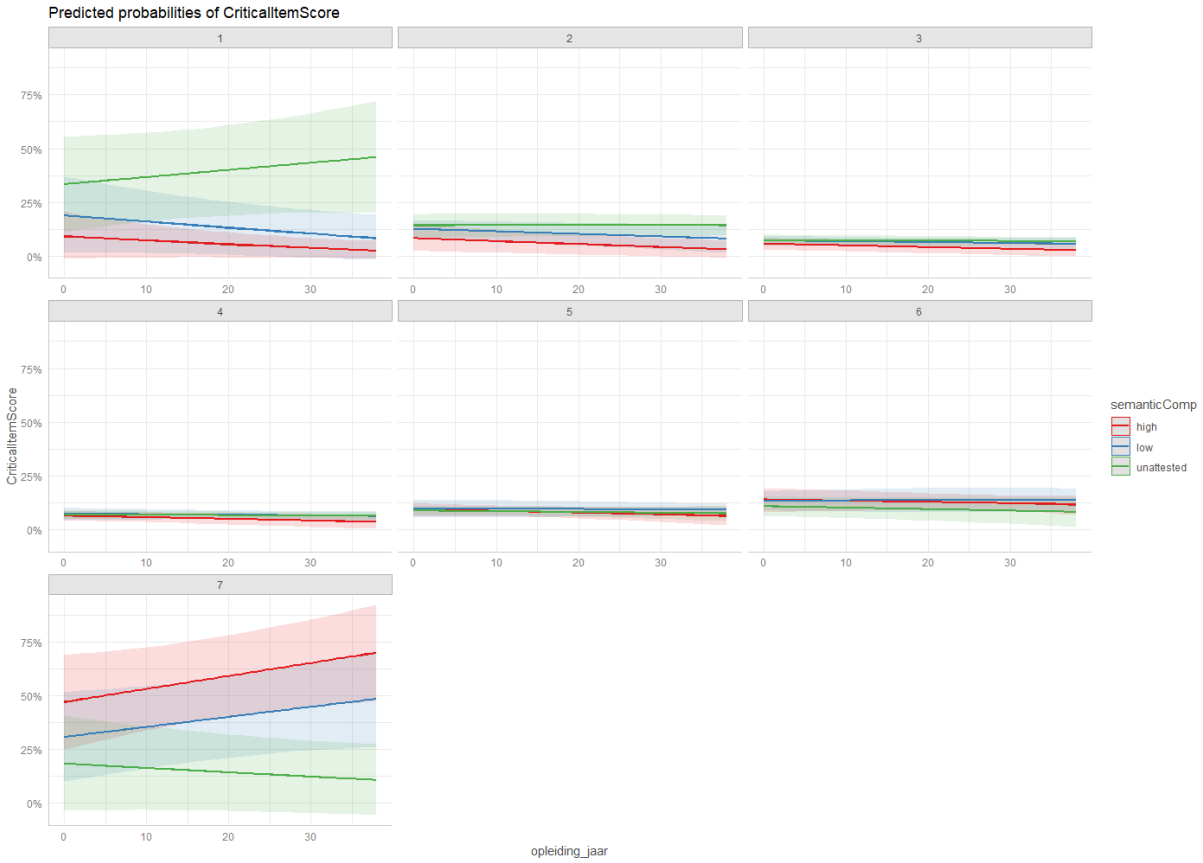

Figure 4. Visualisation of the interaction between education_years and semantic compatibility for the *krijgen*-construction.

Finally, the analyses reveal a robust positive correlation between participants' level of proficiency in English (ENGLISH PROFICIENCY) and their evaluations of conventional as well as creative instantiations of the *weg*-pattern. Bearing in mind that the *weg*-pattern has a more productive counterpart in English (see Section 3.1), this effect indicates that the exposure to a construction in a (dominant) neighbouring language can affect how speakers judge that construction in their native language. More specifically, we observe this pattern as a main effect for the creative instantiations of the construction ($p < 0.001$) and in the model considering all types of experimental items (i.e., hapaxes, as well as highly frequent and intermediately frequent verbs; $p = 0.003$). Additionally, with respect to the creative instantiations, an interaction with LEMMA FREQUENCY shows that the effect is more pronounced for hapaxes high in lemma

frequency. As for the conventional instantiations of the *weg*-pattern, the model reveals that the effect is only at play for verbs with an intermediate token frequency (and not for high frequent verbs). The analyses at the level of the individual verbs uncover that this latter interaction is most likely caused by the effect of one specific verb, viz. the most frequently occurring verb *banen* ('to pave, to clear or smooth a path'), for which a negative trend emerges that borders significance ($p = 0.052$). This is interesting, as the English *way*-construction of course has another "default"-verb, viz. *to make* (Verhagen 2003). Possibly, some of the participants with the highest (reported) proficiency in English (and who are potentially more exposed to the English *way*-cx than to the Dutch *weg*-cx) may have expected the Dutch cognate of 'to make', i.e. *maken*, more so than the Dutch top verb *banen*. For now, the fact that the positive correlation between the ratings and English proficiency can only be attested for the *weg*-construction, and not for the *krijgen*-passive[13], strengthens the hypothesis that this effect is promoted by exposure to its English counterpart.

4.2.3 Personality traits

Though these might have been expected following research in psychology on individual variation in general levels of creativity (see Section 2.2), the data do not reveal main effects of Extraversion or Openness to Experience on participants' evaluations of creative uses of the constructions at stake.

We do, however, observe a significant negative correlation between CONSCIENTIOUSNESS and the ratings of the creative instantiations of the *weg*-construction ($p = 0.033$). In other words, participants who score higher on the Conscientiousness domain evaluated the creative uses of the *weg*-pattern more critically than participants who score lower on this personality trait. The same negative correlation is visible for the creative uses of the *krijgen*-passive, though, for this construction, it is not significant ($p = 0.079$). The pattern does not occur at the level of the more conventional instantiations of either construction. We do find the negative effect for the grammatically unacceptable fillers, though ($p = 0.013$). The conclusion can be drawn that highly conscientious participants – characterized by higher levels of responsibility, dutifulness and reliability – are more critical of test sentences that, for one reason or another, are not immediately recognized as conventional.

Finally, ordinal analyses were also carried out at the level of the BFI facets (see Section 3.2.3 for more information). These indicated a consistent significant positive correlation for INTELLECTUAL CURIOSITY – a facet linked to the personality domain Openness to Experience; measured through questions such as *I am someone who… is curious about many different things / avoids intellectual, philosophical discussion* – for all instantiations of the *weg*-pattern ($p = 0.032$ for the creative uses; $p = 0.004$ for the conventional instantiations; $p = 0.009$ for all types of verbs together). However, the effect is not as pronounced for the *krijgen*-passive ($p = 0.079$ for the conventional instantiations).

4.2.4 Cognitive abilities

As almost 80 percent of participants scored at least 38, 39, or 40 out of 40 on the Dutch vocabulary test (Vander Beken *et al.* 2018), there is too little variation in the sample to include vocabulary knowledge as a factor in the analyses. We suspect this may be related to the web-based setting in which the experiment was conducted. First, when in doubt, some participants

[13] For the *krijgen*-passive, which structurally and functionally resembles the German *bekommen*-passive (cf. footnote 5 in Section 3.1), the analyses do not display any positive effect of exposure to the construction in a neighbouring language. However, we can observe a moderately significant negative correlation ($p = 0.048$) for German proficiency on the evaluations of the conventional instantiations of the *krijgen*-passive. The analyses at the level of the individual verbs indicates that this effect is most likely related to the verb *opsturen* ($p = 0.067$). In this case, it seems that a better familiarity with German and the *bekommen*-passive provokes a more inhibitory effect.

may have quickly looked up the correct answers on the internet. Second, it is reasonable to assume a sample bias, that is, that the survey inherently attracted more participants who are interested in and/or have an aptitude for (the Dutch) language. Future research will have to explore alternative measures.

4.3 Intralinguistic determinants and their interplay with extralinguistic determinants

The previous sections already demonstrated that user-related variables interact closely with intralinguistic determinants of productivity. In this section, we examine the effects of currently known intralinguistic determinants on speakers' evaluations of grammatical constructions more thoroughly.

4.3.1 Token frequency

The data for the *krijgen*-passive provide evidence for the role of TOKEN FREQUENCY on participants' judgements of different instantiations of this grammatical pattern. A significant main effect shows that the creative uses of the construction, i.e., those containing hapaxes, are rated significantly lower than more conventional uses ($p < 0.001$ compared to verbs with intermediate token frequency; $p = 0.007$ compared to high frequent verbs). This is the effect we expected, as items high in token frequency are taken to be more entrenched, which promotes their level of conventionalization and familiarity (cf. Sections 2.1 and 3.2.2). Hapaxes, on the other hand, can be assumed to be less entrenched. The difference in ratings between verbs high in token frequency (HTF) and verbs with an intermediate token frequency (ITF) is not significant ($p = 0.626$). Furthermore, several significant interactions reveal that the effect is more pronounced for (i) younger participants ($p < 0.001$ for the difference between hapaxes and both ITF and HTF), (ii) participants from the Netherlands ($p = 0.002$ for ITF; $p = 0.003$ for HTF), and (iii) participants with a background in linguistics ($p < 0.001$ for ITF; $p < 0.001$ for HTF). Additional significant interactions show that the difference in ratings between hapaxes and HTF-verbs (but not ITF-verbs) is more pronounced for participants with higher scores on Conscientiousness ($p = 0.016$) and with lower levels of proficiency in German ($p = 0.023$).

The effect of token frequency is also visible for the *weg*-pattern, though only for specific participant groups. While creative uses of this construction are consistently rated lower than instantiations containing a HTF-verb across all ages, degrees of educational experience, and levels of Conscientiousness, the behaviour with respect to items containing an ITF-verb differs. More specifically, we observe that younger participants and those with less time spent in full-time education rate these items similarly to hapaxes, meaning they give them lower ratings, whereas older participants and participants who received longer full-time education rate them more similarly to items containing HTF-verbs (i.e., they give higher ratings). These effects are visualised in Figures 5 and 5, respectively. Assuming that older age and longer education imply that speakers have, in general, been more exposed to the selected constructions, this effect suggests a key role for linguistic experience – namely, that the acceptability of items that occur with intermediate token frequency depends on the amount of general experience the language user has with the patterns at stake.

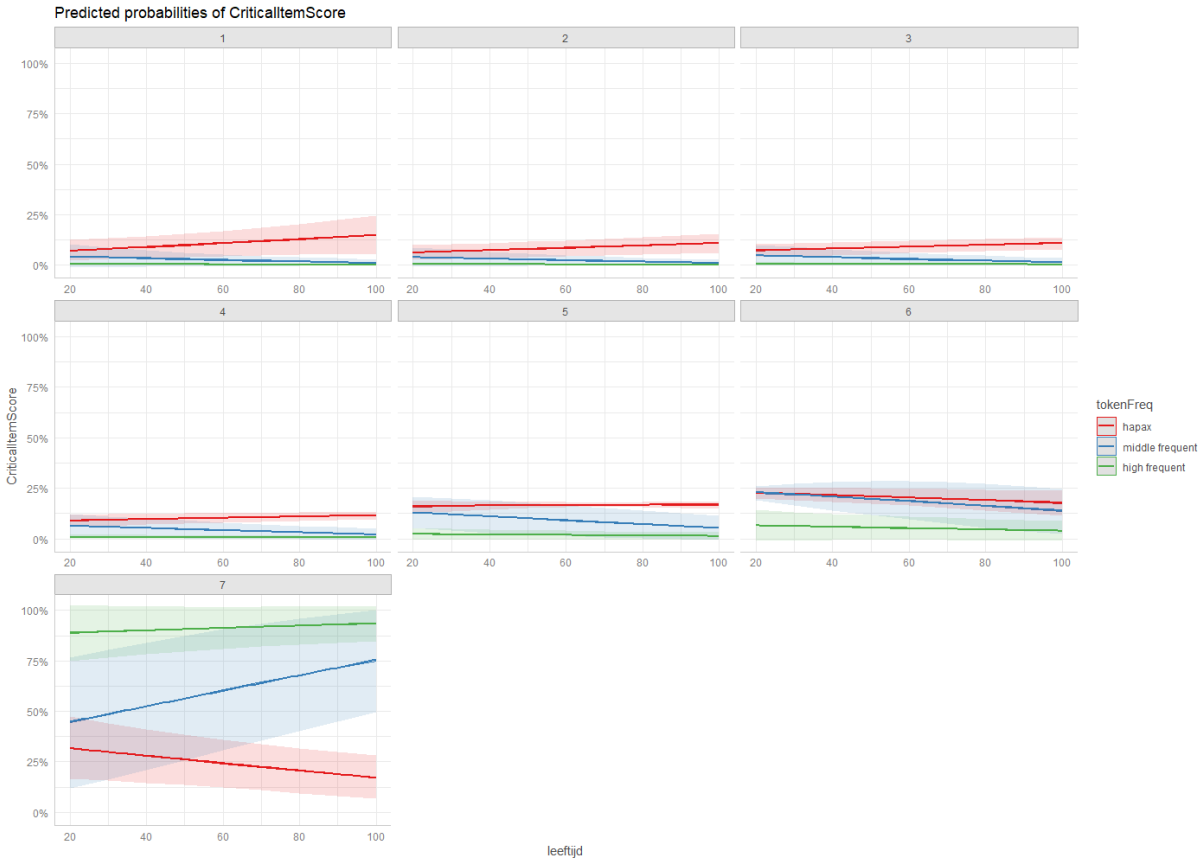

Figure 5. Visualisation of the interaction between age and token frequency for the *weg*-construction.

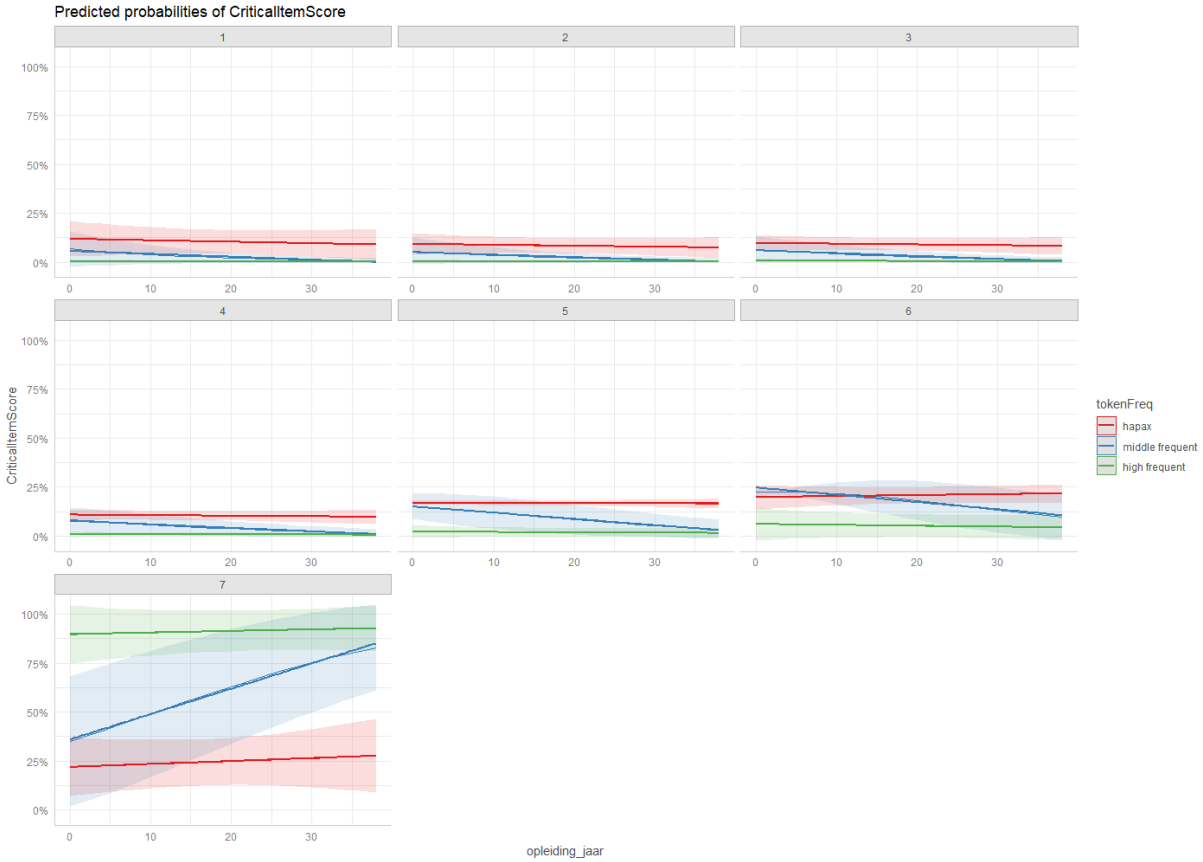

Figure 6. Visualisation of the interaction between education_years and token frequency for the *weg*-construction.

For Conscientiousness, a similar pattern emerges, though in the opposite direction: participants with high levels of Conscientiousness evaluate items with ITF-verbs more similarly to hapaxes, whereas participants with lower levels of Conscientiousness give those items higher ratings, similar to items with HTF-verbs. This can be linked to our observations in Section 4.3.2, viz. that participants who score higher on this personality trait are more critical of items that can be seen as less conventional in one way or another.

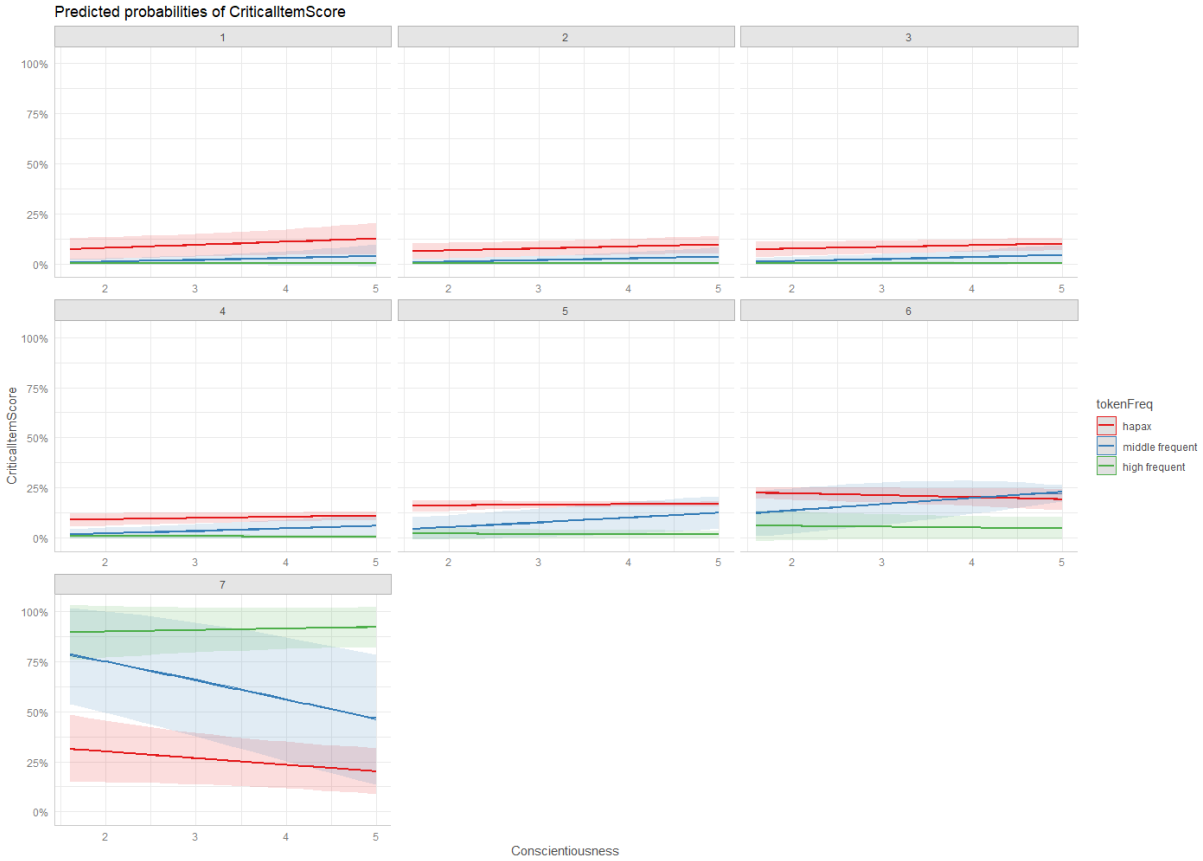

Figure 7. Visualisation of the interaction between Conscientiousness and token frequency for the *weg*-construction.

4.3.2 Lemma frequency

With respect to the *weg*-pattern, the analyses do not uncover a main effect for LEMMA FREQUENCY as a linguistic determinant. However, the visualisations of the interactions reveal that items containing a hapax low in lemma frequency are rated significantly lower than items containing a hapax high in lemma frequency by specific groups of individuals, namely older participants ($p = 0.007$ for AGE x LEMMA FREQUENCY) and participants who spent more time in full-time education ($p = 0.020$ for EDUCATION_YEARS x LEMMA FREQUENCY). In that sense, we can apply a similar reasoning as with the effect of token frequency and postulate that higher levels of linguistic experience (as indicated by the participants' age and educational level) imply higher levels of awareness and familiarity with the overall frequency of the verbs in question, which contributes to their acceptability in the construction (cf. our first hypothesis with respect to lemma frequency in Section 3.2.2).

The opposite pattern can be observed for the *krijgen*-passive, that is, items containing a hapax low in lemma frequency are rated significantly *higher* than items containing a hapax high in lemma frequency in a significant main effect ($p = 0.040$). Furthermore, the interactions indicate that this effect is more pronounced for (i) participants from Belgium ($p = 0.044$ for REGION x LEMMA FREQUENCY), (ii) participants with a background in linguistics ($p = 0.003$ for BACKGROUND_LINGUISTICS x LEMMA FREQUENCY), (iii) participants who score higher on

Conscientiousness ($p = 0.016$), and (iv) participants who score higher on linguistic mobility ($p = 0.021$). In this case, participants seem to be more sensitive to the discrepancy between how often they would *expect* a verb to occur in the construction based on its lemma frequency and its actual frequency of occurrence in the construction, which was the second hypothesis with respect to lemma frequency (Section 3.2.2). However, it should be noted that these analyses include the two unattested verbs (i.e., *weigeren* and *afpakken*) among the group of "hapaxes" high in lemma frequency. As will be shown in Section 4.3.3, these items receive significantly lower ratings than the other hapaxes, which likely considerably weighs on the effect of LEMMA FREQUENCY. Tentatively, this difference in the direction of the effect between the two constructions might be related to their position in the constructional network, or, more specifically, their relationships to other, semantically related ASCs (see also Section 5). In Dutch, there are several different ways to express motion, the *weg*-pattern is only one of which, and a quite special and infrequent one at that. Compare, for instance, *Messi dribbelt zich een weg door de verdediging* 'Messi dribbles his way through/between the defence' with the intransitive motion construction *Messi dribbelt door de verdediging* 'Messi dribbles through/between the defence', a construction that is more "regular" (e.g., does not have a fake 'path' object), has a different, more general semantics (lacking the suggestion of movement in spite of obstacles or difficulty), occurs with all kinds of motion verbs including, unlike the *weg*-pattern, basic verbs such as *gaan* 'go' or *lopen* 'run, walk', etc. The *weg*-construction, as is also observed by Verhagen (2003a: 52) , is a bit of a grammatical 'island'. The *krijgen*-passive, by contrast, is the direct counterpart of the active ditransitive construction and the regular ditransitive passive, three constructions which only differ from each other in terms of the linking of argument roles to grammatical functions (cf. Section 3). As a consequence, participants may be more sensitive to verbs which frequently occur in these other two 'transfer' constructions but not, or only infrequently so, in the *krijgen*-passive. Note, for instance, how *aanbieden* ('to offer') is the top verb in the *krijgen*-passive, whereas its morphological variant *bieden*, which has a similar lemma frequency in the SUBTLEX-database and occurs in the active and "regular" passive ditransitive constructions at least as frequently as *aanbieden*,[14] only is a hapax in the *krijgen*-passive. Speakers might be more sensitive to such "discrepancies" than to cases where a verb does not appear as often in the relatively isolated *weg*-construction as might be expected in view of the verb's lemma frequency. We leave it to future research to further investigate this possibility.

4.3.3 Semantic compatibility

Several interactions for the *krijgen*-passive validate the assumption that the two unattested verbs – i.e., *weigeren* ('to deny') and *afpakken* ('to take away') – are rather 'incompatible' with the construction's semantics. The ratings for these verbs are significantly lower than for the verbs chosen from the corpus investigation in the following groups: (i) older participants ($p < 0.01$ vs. high semantic compatibility items [HSC] and vs. low semantic compatibility items [LSC]), (ii) participants who spent more time in full-time education ($p = 0.017$ HSC; $p = 0.029$ LSC] , (iii) participants from the Netherlands ($p < 0.001$ HSC and LSC), and (iv) females ($p = 0.038$ HSC). More interesting is that the data also indicate that the success of a newly introduced token is promoted by its degree of semantic similarity to previous attestations (cf. Section 2.1). For the *krijgen*-passive, this was observed in one specific participant group, namely speakers of Netherlandic Dutch ($p < 0.001$). For the *weg*-construction, items that contain a hapax low in semantic compatibility are rated significantly lower than items that contain a hapax high in

[14] In Colleman's (2009) data, *bieden* and *aanbieden* are both among the ten strongest collexemes of the active ditransitive construction, with 52 and 39 instances in a one-million-word corpus, respectively.

semantic compatibility, but only for older participants ($p < 0.001$), (ii) female participants ($p = 0.045$), and (iii) participants with a higher proficiency in German ($p = 0.007$).

5 Discussion

Taking an online acceptability rating experiment set up around two Dutch ASCs as a case study, this paper substantiates the idea of individual differences in grammatical representations. The considerable between-speaker variation in the evaluations of different types of instantiations of the two Dutch ASCs corroborates the assumption that linguistic knowledge is usage-based and that our linguistic representations are shaped by our own, individual experience with language (cf. Langacker 1995; Barlow & Kemmer 2000; Bybee 2010). Our data reveal that these differences impact judgements of more unconventional/novel/creative uses of grammatical constructions, meaning that there are clear individual differences in the evaluation of linguistic creativity as well.

In general, we observed that a higher amount of general exposure to linguistic structures in education (e.g., also via an education in linguistics) – results in more tolerant evaluations of the creative instantiations of the investigated constructions. This is expected, since exposure implies a greater level of familiarity with those structures their different types. A particularly intriguing effect related to linguistic experience, is the effect of English proficiency on the evaluations of instantiations of the *weg*-pattern. As already touched upon in Section 4.2.2, our data provide strong evidence for a degree of cross-linguistic transfer where, at least from a comprehension perspective, exposure to a construction in English can affect the processing of that construction in Dutch. The dataset compiled by Oosterlinck (2019) also provides evidence for this transfer-effect in production. While early work on the *weg*-construction initially claimed that incidental-action uses of the *way*-construction (as attested in English) are not possible in Dutch (Verhagen 2002, 2003b; cf. Section 3.1), the dataset contains several counterexamples (cf. Colleman 2020 for more counterexamples in recent language use), including the examples in (3). While further research is required to investigate how and when the incidental-action uses of the *weg*-construction were introduced in Dutch, it seems reasonable to assume a role of cross-linguistic transfer between two variants of the same construction given the effects observed in our data.

(3) a. De bus is overvol, de geur is niet te harden. Na vijf uur slingeren op een aarden baan over een bergpas van 4.500 meter, komen we aan in Santa María, een klein dorpje waar we een combi nemen naar Santa Teresa. De afgeleefde combi **kreunt zich een weg** naar boven.
'The bus was overcrowded, the smell was unbearable. After five hours of swinging back and forth on a mountain trail of 4.500 kilometres, we arrive in Santa María, a small village from where we take a station car to Santa Teresa. The worn-down station car **groans its way** upwards.'

b. De onverwachte comeback van Alex Van Damme bleek van groot nut. Zonder zich in de kijker te spelen, deed hij waar hij goed in is: de ploeg samenhouden. Als één van de laatste **zuchtte hij zich een weg** uit de kleedkamer.
'The unexpected comeback of Alex Van Damme proved to be of large value. Without playing to stand out, he did what he does best: keep the team together. As one of the last ones, he **sighed his way** to the dressing rooms.'

Furthermore, participants' ratings are influenced by factors relating to their social background (Section 4.2.1). More specifically, we can detect an influence of age, gender, and region of origin. However, more research is required to fully grasp how age affects speakers' linguistic creativity, as we encounter reversed effects in the two constructions (see Table 6). Theoretically, either direction of the effect of age would make sense. On the one hand, from a usage-based

perspective, the fact that older participants are, over time, likely to build up more experience with specific grammatical patterns may facilitate the process of analogy that is involved in extending constructions with new lexical items (Bybee 2010; 2013). However, we also have to take into account potential age-related cognitive decline in language processing (Marini & Andreetta 2016), which may at the same time impede processes of analogy and linguistic creativity. On the other hand, from a variationist sociolinguistic perspective, the well-studied phenomenon of age grading entails that the use of non-prestigious linguistic features (i.e., features that are not part of the standard language and that are associated with a high degree of social awareness and/or are seen as stigmatized forms) peaks during adolescence (the so-called 'adolescent peak'), whereas in middle to older age, speakers tend to become more conservative as they adapt to the norms and values of mainstream society (Chesire 2005; Tagliamonte 2011; Eckert 2017). In this specific study, we suspect specific characteristics of the constructions at stake lie at the basis of the attested patterns. More specifically, while we cannot yet adduce empirical evidence to support this, both constructions seem to differ in terms of their degree of "extravagance": while the *krijgen*-passive is a rather formal "perspectivizing" construction used to express transfer, which differs from the active ditransitive construction only in terms of the linking of argument roles to grammatical functions, the *weg*-construction enables the speaker to express a motion event in a more unconventional, imaginative, and vivid manner compared to more "vanilla" ways of encoding such events, such as the intransitive-motion construction (compare *Hij slalomde door het strafschopgebied* 'He slalomed through the penalty area' to *Hij slalomde zich een weg door het strafschopgebied* 'He slalomed his way through the penalty area', where, arguably, the latter has "richer" semantics, for instance by implying that there were important obstacles along the path, thus lending greater prominence to the subject's skill in avoiding/overcoming these). These are all characteristics that are mentioned in the literature on so-called extravagant constructions (see Ungerer & Hartmann 2020 for an overview). Tying in this hypothesis with the literature on age grading, this may explain why older participants give higher ratings to the more conventional instantiations of the *weg*-pattern and tend to be more open to productive extensions of the more formal *krijgen*-passive. In contrast, younger participants seem to be more drawn to the unconventional/imaginative/vivid nature of creative instantiations of the *weg*-pattern.

Another aspect of the participants' social background that influences their ratings pertains to their geographic location. More specifically, the demonstrated effect of region of origin shows that participants from the Netherlands tend to give lower ratings to the creative uses of the investigated constructions (and, by extension, to the more questionable filler items) than participants from Belgium[15]. This might on first sight be a surprising effect, as one might assume from a language-historical perspective that speakers of *Belgian* Dutch would be more severe in their evaluations. After all, whereas the history of Netherlandic Dutch is described as a process of *normal* and *spontaneous* standardization that quite naturally evolved out of a dominant dialect, Flanders (i.e., the Dutch-speaking part of Belgium) is said to have *imported* the standard language from the North. Following these historical developments, Netherlandic Dutch is said to be a fully standardized variety, whereas the standardization process of Belgian Dutch is often characterized as *delayed* or *unfinished* (see i.a., Geeraerts *et al.* 1999 Chapter 2; Jaspers 2001; Grondelaers & Van Hout 2011; Van Hoof & Jaspers 2012). The "more advanced standardization" of Netherlandic Dutch (Grondelaers *et al.* 2008: 186) is argued to be reflected in *more stable* and *less complex* distributions of both lexical (see Geeraerts *et al.* 1999 for lexical

[15] A rather evident explanation for this pattern resides in the fact that the stimulus materials were constructed by speakers of Belgian Dutch. However, then we would expect to see this pattern for the conventional instantiations of the constructions as well, which is not the case. Moreover, the survey was pretested among eleven participants, among whom four speakers of Netherlandic Dutch who were specifically instructed to pay attention to words and phrases that were less commonly known or used in Netherlandic Dutch.

choices in the domain of clothing and football terms) and syntactic phenomena (see Grondelaers *et al.* 2020 for an overview). As a consequence, speakers of Belgian Dutch are hypothesized to display generally less linguistic security than speakers from Netherlandic Dutch, so that they might be expected to be on the whole more "conscientious" and/or "critical" in their assessment, resulting in generally lower ratings (cf. the demonstrated effect of Conscientiousness discussed below). A potential explanation for why the reverse effect is observed in our study can be found in Impe (2010). In an auditory lexical decision task with both existing and nonsense words, recorded in ten different varieties of Belgian and Netherlandic Dutch, Impe (2010) found that speakers of Netherlandic Dutch systematically identified the nonsense stimuli faster than speakers of Belgian Dutch, irrespective of the variety in which these were recorded. This implies that speakers of Netherlandic Dutch more quickly break off the cognitive process of searching for a match in their mental lexicon before rejecting a word, which Impe (2010) attributes to their greater level of linguistic security – i.e., speakers of Netherlandic Dutch seem to equate a non-match to the non-existence of a word, whereas speakers of Belgian Dutch do not so easily take a non-match as evidence for the non-existence of a word. With respect to our findings, this may also explain why speakers of Netherlandic Dutch are more severe in their rating of the unconventional/creative instantiations of both constructions, which generally occur with very low frequency and are likely not part of participants' individual constructicon.

With respect to personality traits, while our data do not provide any evidence for an effect of Extraversion or Openness – which might have been expected given their influence on levels of general creativity as attested in psychological research (Section 2.2) – we can observe that participants who score higher on Conscientiousness generally give lower ratings to less conventional phrases. We hypothesize this may (partly) be a task-specific effect, assuming that speakers with higher levels of this personality trait are more likely to be earnestly engaged in the task at hand and, as a consequence, evaluate items only after careful consideration. Finally, the analyses revealed a significant positive correlation between participants' levels of Intellectual Curiosity and their ratings of both conventional and unconventional instantiations of the *weg*-pattern.

6 Conclusion

To conclude, we return to the four research questions outlined in Section 3.1. First, the distribution of ratings of the experimental items demonstrates substantial inter-individual/between-speaker variation in the evaluations of novel/creative instantiations of the selected grammatical constructions (RQ1).

Second, as already hinted at by De Smet (2020), this study demonstrates that not only language-internal predictors, but also *language-external* factors play a role in explaining individual variation in (evaluations of) productivity. Mixed ordinal regression analyses demonstrate that the ratings are irrefutably shaped by a range of characteristics of the language user, such as their social background, linguistic experience, and their personality traits (RQ2). Following several calls to better integrate the social aspects of language in cognitive linguistic approaches in general (Croft 2009; Schmid 2016; Dąbrowska 2016) and in productivity research in specific (De Smet 2020), the results of the current investigation corroborate the importance of considering individual variation, especially with respect to productivity/creativity and related linguistic phenomena such as extravagance, innovation, and language change, where the language user plays a central role. After all, as Dąbrowska (2016: 486) observes, "language is not only an instrument *for* social interaction; it is also a system that emerges *through* interaction, and we cannot hope to understand its structure without considering both cognitive and social factors *and their interactions*." In both corpus-based and experimental approaches, this begs the question of whose language we are actually studying. Is the chosen corpus sufficiently

representative of the targeted population? While an increasing amount of attention has been paid to individuality in corpus data (Barlow 2013; Dąbrowska 2014; Petré 2016; a special issue in *Cognitive Linguistics* edited by Petré & Anthonissen 2020), recent experimental studies in usage-based grammar question to what extent amalgamated corpus data are informative of mental representations of individual speakers (Barlow 2013; Verhagen 2019; Barking *et al.* 2022). In experimental studies as well, the characteristics of the chosen sample may influence the patterns that are encountered.

Third, the analyses provide evidence that is in line with expectations from the literature for the role of the three included intralinguistic determinants, viz. token frequency, lemma frequency, and semantic compatibility (RQ3).

Fourth, intralinguistic and extralinguistic determinants of productivity prove to be inextricably linked to each other (RQ4). The two-way interactions reveal that, on the one hand, the effects of certain user-related variables only play at the level of specific classes of linguistic items (e.g., the effect of gender only applies for items that contain a hapax low in semantic compatibility; see Section 4.2.1), and, concomitantly, that the effects of the included linguistic variables are more pronounced (or, in some cases even, are only at play) at the level of certain groups of individuals (Section 4.3). One important consequence of the latter observation is that a more nuanced characterization may be required of previously described system-internal factors, that accounts for their heterogenous effects across different types of speakers.

In sum, the present exploration of the role of individual, user-related variables on the evaluations of novel/creative uses of grammatical constructions demonstrates that individual language users differ in the extent to which they evaluate productive extensions from other speakers as acceptable and that they are hence also likely to differ in the degree to which they extend constructions creatively themselves (Section 1). Further research is therefore required to investigate the effects of extralinguistic variables, and their interplay with intralinguistic variables, on the productivity of constructions – in the context of both comprehension and production, from a corpus-based as well as an experimental approach.

References

Baayen, R. Harald. 2009. Corpus linguistics in morphology: morphological productivity. In Anke Lüdeling & Merja Kyto (ed.), *Corpus Linguistics. An international handbook*, 900–919. Berlin/Boston: De Gruyter Mouton.

Baayen, R. Harald & Rochelle Lieber. 1991. Productivity and word-formation in English: A corpus-based study. *Linguistics* 29(5). 801-483.

Bader, Markus. 2012. The German *bekommen* passive: A case study on frequency and grammaticality. *Linguistische Berichte* 2012(231). 249-298.

Bader, Markus & Jana Häussler. 2010. Toward a model of grammaticality judgments. *Journal of Linguistics* 46(2). 273-330.

Barðdal, Jóhanna. 2008. *Productivity. Evidence from case and argument structure in Icelandic*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Barking, Marie, Ad Backus & Maria Mos. 2022. Individual corpus data predict variation in judgments: testing the usage-based nature of mental representations in a language transfer setting. *Cognitive Linguistics* 33(3). 481-519.

Barlow, Michael. 2013. Individual differences and usage-based grammar. *International Journal of Corpus Linguistics* 18(4). 443–478.

Barlow, Michael & Suzanne Kemmer. 2000. *Usage-based models of language.* Stanford, CA: CSLI Publications.

Bergs, Alexander. 2019. What, if anything, is linguistic creativity?. *Gestalt Theory* 41(2). 173-183.

Bybee, Joan & Sandra Thompson. 1997. Three frequency effects in syntax. *Annual Meeting of the Berkeley Linguistics Society*. 378-388.

Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, Joan. 2013. Usage-based theory and exemplar representations of constructions. In Thomas Hoffmann & Graeme Trousdale (ed.), *The Oxford handbook of construction grammar*, 49-69. New York: Oxford University Press.

Bybee, Joan & David Eddington. 2006. A usage-based approach to Spanish verbs of 'becoming'. *Language* 82(2). 323-355.

Cheshire, Jenny. 2005. Age- and Generation-Specific Use of Language. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (ed.), *Sociolinguistics Volume 2: An International Handbook of the Science of Language and Society*, 1552-1563. Berlin/New York: De Gruyter Mouton.

Christensen, Rune Haubo B. 2015. Analysis of ordinal data with cumulative link models—estimation with the R-package ordinal. R-package version, 28, 406.

Christensen, Rune Haubo B. 2018. Cumulative link models for ordinal regression with the R package ordinal. Submitted in *J. Stat. Software* 35.

Christensen Rune Haubo B. 2019a. ordinal: Regression models for ordinal data. R-package version 2022.11-16. https://cran.r-project.org/web/packages/ordinal/ordinal.pdf [Accessed March 28th 2023].

Christensen, Rune Haubo B. 2019b. A Tutorial on fitting Cumulative Link Mixed Models with clmm2 from the ordinal Package. *Tutorial for the R Package ordinal https://cran. r-project. org/web/packages/ordinal/Accessed*, *1*.

Clement, Marja & Elvira Glaser. 2014. Hoe krijg je dit gepubliceerd ? 'Krijgen'-constructies in het Nederlands. In Freek Van de Velde, Hans Smessaert, Frank Van Eynde & Sara Verbrugge (ed.), *Patroon en argument: Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst,* 195-214. Leuven : Leuven University Press.

Coates, Jennifer. 2015. Women, Men and Language. A Sociolinguistic Account of Gender Differences in Language [3rd edition]. London : Routledge.

Colleman, Timothy. 2009. The semantic range of the Dutch double object construction. A collostructional perspective. *Constructions and Frames* 1(2). 190-221.
Colleman, Timothy. 2015. Constructionalization and post-constructionalization: The constructional semantics of the Dutch *krijgen*-passive in a diachronic perspective. In Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (ed.), *Diachronic Construction Grammar*, 213-255. Amsterdam/Philadelphia: John Benjamins.
Colleman, Timothy. 2020. Twee vroege constructiegrammaticale studies in Nederlandse Taalkunde: een terugblik op Schermer-Vermeer (2001) en Verhagen (2003a). *Nederlandse Taalkunde* 25(2-3). 333-343.
Colleman, Timothy & Dario Rens. 2016. Het *krijgen*-passief en de werkwoordelijke eindgroep: Een diachrone voorstudie op basis van Delpher. *Studies van de Belgische Kring voor Linguïstiek/Travaux du Cercle Belge de Linguistique/Papers of the Linguistic Society of Belgium* 10. 1-15.
Collins, Peter C. 1996. *Get*-passives in English. *World Englishes* 15(1). 43-56.
Coulmas, Florian. 1997. *The Handbook of Sociolinguistics*. Oxford: Blackwell.
Coulmas, Florian. 2003. Sociolinguistics. In Mark Aronoff & Janie Rees-Miller (ed.), *The Handbook of Linguistics*, 563-581. Oxford: Blackwell.
Cowart, Wayne. 1997. Experimental syntax: Applying objective methods to sentence judgements. Calfifornia: SAGE Publications.
Croft, William. 2009. Toward a social cognitive linguistics. In Stéphane Pourcel & Vyvyan Evans (ed.), *New directions in cognitive linguistics*, 395-420. Amsterdam: John Benjamins.
Dąbrowska, Ewa. 2010. Naïve v. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review* 27(1). 1-23.
Dąbrowska, Ewa. 2012. Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2(3). 219-253.
Dąbrowska, Ewa. 2014. Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics* 25(4). 617-653.
Dąbrowska, Ewa. 2018. Experience, aptitude and individual differences in native language ultimate attainment. *Cognition* 178. 222-235.
Dąbrowska, Ewa. 2019. Individual differences in grammatical knowledge. In: Ewa Dąbrowska & Dagmar Divjak (ed.), *Cognitive Linguistics – key topics*, 231-250. Berlin/Boston: De Gruyter Mouton.
Dąbrowska, Ewa & James Street. 2006. Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences* 28(6). 604-615.
Delaby, Gauthier & Timothy Colleman. 2023. Volgordevariatie in groepen met receptief *krijgen* + voltooid deelwoord in de Nederlands Nederlandse krantentaal. Een synchroon en diachroon onderzoek. *Nederlandse Taalkunde* 28(1). 1-35.
De Smet, Hendrik. 2020. What predicts productivity? Theory meets individuals. *Cognitive Linguistics* 31(2). 251-278.
Denissen, Jaap J. A., Rinie Geenen, Christopher J. Soto, Oliver P. John & Marcel A.G. van Aken. 2019. The Big Five Inventory–2: Replication of psychometric properties in a Dutch adaptation and first evidence for the discriminant predictive validity of the facet scales. *Journal of Personality Assessment* 102(3). 309–324.
Divjak, Dagmar. 2017. The role of lexical frequency in the acceptability of syntactic variation. Evidence from that-clauses in Polish. *Cognitive Science* 41(2). 354- 382.
Eckert, Penelope. 2017. Age as a sociolinguistic variable. In Florian Coulmas (ed.), *The handbook of Sociolinguistics*, 151-167. Oxford: Blackwell.
Flach, Suzanna. 2020. Schemas and the frequency/acceptability mismatch: Corpus distribution predicts sentence judgments. *Cognitive Linguistics* 31(4). 609-645.

Geeraerts, Dirk, Stefan Grondelaers & Dirk Speelman. 1999. Convergentie en divergentie in de Nederlandse woordenschat: een onderzoek naar kleding- en voetbaltermen. Amsterdam: P.J. Meertens Instituut.

Givón, Talmy & Lynne Yang. 1993. The ruse of the English *get*-passive. In Barbara Fox & Paul J. Hopper (ed.), *Voice. Form and Function*, 119-149. Amsterdam/Philadelphia: John Benjamins.

Goldberg, Adele. 1995. Constructions: a construction grammar approach to argument structure. Chicago, IL: University of Chicago Press.

Goldberg, Adele. 2006. Constructions at work: The nature of generalization in language. Oxford: Oxford University Press.

Goldberg, Adele. 2019. Explain me this. Creativity, Competition, and the Partial Productivity of Constructions. Princeton: Princeton University Press.

Goldberg, Lewis R. 1990. An alternative "description of personality": the Big-Five Factor structure. *Journal of Personality and Social Psychology* 59(6). 1216-1229.

Grondelaers, Stefan, Hilde Van Aken, Dirk Speelman & Dirk Geeraerts. 2001. Inhoudswoorden en preposities als standaardiseringsindicatoren. De diachrone en synchrone status van het Belgische Nederlands. *Nederlandse Taalkunde* 6. 179-202.

Grondelaers, Stefan, Dirk Speelman & Dirk Geeraerts. 2008. National variation in the use of *er* "there". Regional and diachronic constraints on cognitive explanations. In Gitte Kristansen & René Dirven (ed.), *Cognitive Sociolinguistics. Language variation, cultural models, social systems*, 153-204. Berlin/New York: De Gruyter.

Grondelaers, Stefan, Robbert De Troij, Dirk Speelman & Antal van den Bosch. 2020. Vissen naar variatie. Digitaal op zoek naar onbekende Noord/Zuid-verschillen in de grammatica van het Nederlands. *Nederlandse Taalkunde* 25(1). 73-99.

Grondelaers, Stefan & Roeland van Hout. 2011. The standard language situation in the Low Countries: top-down and bottom-up variations on a diaglossic theme. *Journal of Germanic Linguistics* 23(3). 199-243.

Hoffmann, Thomas. 2018. Creativity and construction grammar: Cognitive and psychological issues. *Zeitschrift für Anglistik und Amerikanistik* 66(3). 259-276.

Hoffmann, Thomas. 2019. Language and creativity: a Construction Grammar approach to linguistic creativity. *Linguistics Vanguard* 5(1). 20190019.

Hoffmann, Thomas. 2020a. Construction grammar and creativity: Evolution, psychology, and cognitive science. *Cognitive Semiotics* 13(1). 20202018.

Hoffmann, Thomas (ed.). 2020b. Construction Grammar and Creativity [Special Issue]. *Cognitive Semiotics* 13(1).

Hoffmann, Thomas. 2022. Constructionist approaches to creativity. *Yearbook of the German Cognitive Linguistics Association* 10(1). 259-284.

Impe, Leen. 2010. *Mutual intelligibility of national and regional varieties of Dutch in the Low Countries* [Unpublished doctoral dissertation]. Katholieke Universiteit Leuven.

Israel, Michael. 1996. The *way* constructions grow. In Adele Goldberg (ed.), *Conceptual Structure, Discourse and Language*, 217-230. Stanford: CSLI.

Jaspers, Jürgen. 2001. Het Vlaamse stigma. Over tussentaal en normativiteit. *Taal & Tongval* 53(2). 129-153.

Jirásek, Michal & Frantisek Sudzina. 2020. Big five personality traits and creativity. *Quality Innovation Prosperity* 24(3). 90-105.

John, Oliver P., Laura P. Nauman & Christoper J. Soto. 2008. Paradigm shift to the integrative Big Five trait taxonomy. History, measurement, and conceptual ideas. In Oliver P. John, Richard W. Robins, & Lawrence A. Pervin (ed.), *Handbook of personality: Theory and research*, 114-158. New York: Guilford Press.

Jurafsky, Daniel & James H. Martin. 2024. Vector semantics and embeddings. In Daniel Jurafsky & James H. Martin (ed.), Speech and Language Processing. An introduction to

Natural Language Processing, Computational Linguistics, and Speech Recognition, 105-136. New Jersey: Pearson Education International.

Kandler, Christian, Rainer Riemann, Alois Angleitner, Frank M. Spinath, Peter Borkenau & Lars Penke. 2016. The nature of creativity: The roles of genetic factors, personality traits, cognitive abilities, and environmental sources. *Journal of Personality and Social Psychology* 111(2). 230–249.

Kaufman, James C. 2016. *Creativity 101*. 2nd ed. New York: Springer Publishing Company.

Kempen, Gerard & Karin Harbusch. 2005. The relationship between grammaticality ratings and corpus frequencies: A case study into word order variability in the midfield of German clauses. In Stephan Kepser & Marga Reis (ed.), *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 329-349. Berlin/New York : De Gruyter Mouton.

Keuleers, Emmanuel, Marc Brysbaert & Boris New. 2010. SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods* 42(3). 643-650.

Kharkhurin, Aanatoliy V. 2012. *Multilingualism and Creativity*. Bristol, Blue Ridge Summit: Multilingual Matters.

Labov, William. 2001. Principles of linguistic change vol.2: Social factors. Oxford: Blackwell.

Langacker, Ronald W. 1995. *Foundations of cognitive grammar : Volume I : Theoretical prerequisites.* 4th print. Stanford, CA: Stanford university press.

Leirbukt, Oddlef. 1997. Untersuchungen zum bekommen-Passiv im heutigen Deutsch. Tübingen: Niemeyer.

Linguistic Creativity [Special Issue]. 2018. *Zeitschrift für Anglistik und Amerikanistik* 66(3).

Marini, Andrea & Sara Andreetta. Age-related effects on language production. In Heather Harris Wright (ed.), *Cognition, language and aging*, 55-97. Amsterdam: John Benjamins.

McCrae, Robert R. & Paul T. Costa. The Five-Factor theory of personality. In Oliver P. John, Richard W. Robins, & Lawrence A. Pervin (ed.), *Handbook of personality: Theory and research*, 159-181. New York: Guilford Press.

Milroy, Lesley. 2002. Introduction: Mobility, Contact, and Language Change--Working with Contemporary Speech Communities. *Journal of sociolinguistics* 6(1). 3-15.

Misyak, Jennifer B. & Morten H. Christiansen. 2012. Statistical learning and language: an individual differences study. *Language Learning* 62(1). 302–331.

Oostdijk, Nelleke, Martin Reynaert, Véronique Hoste & Ineke Schuurman. 2013. The Construction of a 500 Million Word Reference Corpus of Contemporary Written Dutch. In Peter Spyns & Jan Odijk (ed.), Essential Speech and Language Technology for Dutch: Results by the STEVIN-programme, 219-250. Heidelberg/New York/Dordrecht/London: Springer.

Oosterlinck, Ellis. 2018. Constructie in taal en tekst. De alternantie tussen de zich een weg constructie en de zich-verplaatsingsconstructie. [Unpublished master's thesis]. Gent: Universiteit Gent.

Perek, Florent. 2015. Argument structure in usage-based construction grammar: experimental and corpus-based perspectives. Amsterdam: John Benjamins Publishing Company.

Perek, Florent. 2016. Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics* 54(1). 149-188.

Perek, Florent. 2018. Recent change in the productivity and schematicity of the way-construction: A distributional semantic analysis. *Corpus Linguistics and Linguistic Theory* 14(1). 65-97.

Petré, Peter. 2016. Unidirectionality as a cycle of convention and innovation. Micro-changes in the grammaticalization of [BE *going to* INF]. *Belgian Journal of Linguistics* 30. 115-146.

Petré, Peter & Lynn Anthonissen (ed.). 2020. Constructionist approaches to individuality in language [Special issue]. *Cognitive Linguistics* 31(2).

Pos, Michaela. 2010. Under Construction. *Cognitive and Computational Aspects of Extended Lexical Units* [Doctoral dissertation]. Amsterdam: LOT/Netherlands Graduate School of Linguistics.

Schäfer, Roland. 2015. Processing and querying large web corpora with the COW14 architecture. In *Proceedings of the 3rd Workshop on Challenges in the Management of Large Corpora*. pp. 28-34.

Schäfer, Roland & Felix Bildhauser. 2012. Building large corpora from the web using a new efficient tool chain. In *Proceedings of the eighth international conference on language resources and evaluation*. 486-493.

Schmid, Hans-Jörg. 2015. A blueprint of the entrenchment-and-conventionalization model. *Yearbook of the German Cognitive Linguistics Association* 3(1). 3-26.

Schmid, Hans-Jörg. 2016. Why Cognitive Linguistics must embrace the social and pragmatic dimensions of language and how it could do so more seriously. *Cognitive Linguistics* 27(4). 543-557.

Schmid, Hans-Jörg. 2020. The dynamics of the linguistic system: Usage, conventionalization, and entrenchment. Oxford : Oxford University Press.

Schütze, Carson T. 2019. The empirical base of linguistics: Grammaticality judgments and linguistic methodology. Berlin: Language Science Press.

Schütze, Carson T. & Jon Sprouse. 2013. Judgment data. In Robert J. Podesva & Devyani Sharma (ed.), *Research methods in linguistics*, 27–50. Cambridge, England: Cambridge University Press.

Soto, Christopher. J. Oliver P. & John. 2009. Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality* 43. 84-90.

Soto, Christopher. J. Oliver P. & John. 2017. The next big five inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology* 113(1). 117–143.

Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209-243.

Stefanowitsch, Anatol & Stefan Th. Gries. 2005. Covarying collexemes. *Corpus Linguistic and Linguistic Theory* 1(1). 1-43.

Street, James A. & Ewa Dąbrowska. 2010. More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers?. *Lingua* 120(8). 2080-2094.

Suttle, Laura & Adele Goldberg. 2011. The partial productivity of constructions as induction. *Linguistics* 49(6). 1237-1269.

Tagliamonte, Sali A. 2011. Variationist sociolinguistics: Change, observation, interpretation. Wiley-Blackwell.

Ungerer, Tobias & Stefan Hartmann. 2020. Delineating extravagance: Assessing speakers' perceptions of imaginative constructional patterns. *Belgian Journal of Linguistics* 34(1). 345-356.

Vander Beken, Heleen, Evy Woumans & Marc Brysbaert. 2018. Studying texts in a second language: No disadvantage in long-term recognition memory. *Bilingualism: Language and Cognition* 21(4). 826-838.

Van Egmond, Marie-Elaine. 2006. Two 'way'-constructions in Dutch: motion along a path and transition to a location [Master's thesis]. Canterbury: University of Canterbury.

Van Hoof, Sarah & Jürgen Jaspers. 2012. Hyperstandaardisering. *Tijdschrift voor Nederlandse taal- en letterkunde* 128(2). 97-125.

Van Leeuwen, Maarten. 2006. Een baan aangeboden krijgen? Dat krijg ik nooit gedaan ! Een synchroon en diachroon onderzoek naar de gebruiksmogelijkheden van krijgen +

participium in het kader van de constructiegrammatica [Master's thesis]. Leiden: Universiteit Leiden.
Verhagen, Arie. 2002. From parts to wholes and back again. *Cognitive Linguistics* 13(4). 403-439.
Verhagen, Arie. 2003a. The Dutch *way*. In Arie Verhagen & Jeroen van de Weijer (ed.), *Usage-Based Approaches to Dutch*, 27-58. Amsterdam: LOT.
Verhagen, Arie. 2003b. Hoe het Nederlands zich een eigen weg baant. Vergelijkende en historische observaties vanuit een constructie-perspectief. *Nederlandse taalkunde* 8. 328-346.
Verhagen, Arie. 2005. Constructiegrammatica en 'usage based' taalkunde. *Nederlandse Taalkunde* 10. 197-222.
Verhagen, Véronique. 2019. *Illuminating variation. Individual differences in entrenchment of multi-word units*. [Doctoral dissertation]. Amsterdam: LOT/Netherlands Graduate School of Linguistics.
Verhagen, Véronique, Maria Mos, Ad Backus & Joost Schilperoord. 2018. Predictive language processing revealing usage-based variation. *Language and Cognition* 10(2). 329-373.
Verhagen, Véronique, Maria Mos, Joost Schilperoord & Ad Backus. 2020. Variation is information: Analyses of variation across items, participants, time, and methods in metalinguistic judgment data. *Linguistics* 58(1). 37-81.
Zawada, Britta. 2006. Linguistic creativity from a cognitive perspective. *Southern African Linguistics and Applied Language Studies* 24(2). 235-254.
Zeschel, Arne. 2012. Incipient productivity: A construction-based approach to linguistic creativity. Berlin/Boston: De Gruyter Mouton.
Zeldes, Amir. 2012. Productivity in Argument Selection: From Morphology to Syntax. Berlin/Boston: De Gruyter Mouton.