Storage or decomposition in Dutch morphological processing: Arguments for the use of the MMN paradigm*

Hernán Labbé Grünberg

Fred Weerman

Judith Rispen

Jan Don

ACLC, University of Amsterdam

Research into the processing of morphologically complex words has yielded an extensive body of work over the last decades. However, still no consensus has been reached about the mechanisms involved. In Dutch, empirical evidence seems to contradict the predictions of most models. We will argue this is due to theoretical and methodological limitations of the models and the experiments that tested their predictions. After examining what is known and what remains to be investigated, we will propose the use of the mismatch negativity paradigm as a tool to tackle the remaining gaps in the study of Dutch morphological processing.

1 Introduction

An important research topic in psycholinguistics is the processing of morphologically complex words. Morphemes are defined as the smallest unit of meaning, and they include both free standing morphemes (like Dutch *hond* 'dog' or *werk* 'work'), as well as bound morphemes (the plural *-en* in *honden* 'dogs', the infinitive *-en* in *werken* 'to work', or the past tense *-te* in *werkte* 'worked'). Morphologically complex words are in principle decomposable into their constituent elements. In the example above, *honden* can be interpreted as [[domestic canine] plus [more than one]], yielding the meaning speakers associate with the string *honden*. This view, known as the decompositional hypothesis, assumes lexical representation would only exists for individual morphemes (like *hond*, *werk-*, plural *-en*, past tense *-te*, infinitive *-en*, etc.), and

* This article is based on the first chapter of my PhD thesis, and is meant to act as a theoretical background to justify the use of the MMN paradigm as a tool to study the processing of Dutch simple and morphologically complex words.

Linguistics in Amsterdam 11,1 (2018): 22–45

that the meaning of morphologically complex words would be created on-line, during the course of word recognition. There is, however, no agreement as to whether complex words are analyzed in this way (via on-line decomposition into their constituent morphemes) or whether they are stored in lexical memory as whole units, just as monomorphemic words would.

This debate touches upon an important question in linguistics: The psychological reality of morphemes. Assuming that all words, regardless of morphological complexity, are stored as whole units in lexical memory, implies taking the morphological level away from the processes and mechanisms responsible for word comprehension. It would mean, in other words, that the morpheme does not exist in its own right, but would rather be a descriptive construct created by linguists to refer to certain phonological-semantic correlations (e.g. the correlation between Dutch verbs ending in sounds /də/, or /tə/ and the reference to actions that occurred in the past). On the other side of the debate are linguists who see morphemes as psycholinguistically real objects of study. This view assumes morphology has its own level of analysis during word comprehension processes, since morphemes are concrete instantiations at the behavioral, cognitive and physiological levels.

This debate has motivated a large body of research over the past decades. Experiments have been carried out in different languages and with different methods to test the hypotheses of storage and decomposition. However, after decades of research, there is little agreement between psycholinguists at both sides of the debate. Specifically, evidence has been found to support both hypotheses, and it does not seem that the methodologies and experiments used in the past will easily solve the controversy (for recent and promising approaches to these questions, see e.g. Fruchter & Marantz 2015; Lignos & Gorman 2012; Zhang, van Heuven & Conklin 2011). Unsurprisingly, most theories have settled somewhere in the middle, acknowledging the existence of both lexical storage and decompositional mechanisms, but with vague criteria to explain why or when storage and/or decomposition would act during the course of word comprehension. In this paper I will argue this is not because there is consistent evidence of some sort of complementary or parallel process by which both mechanisms come into play. Instead, while the evidence for storage seems robust enough, the evidence for decomposition shows highly variable results, which make it impossible to either rule out or incorporate the mechanism of decomposition into models of word comprehension. The purpose of this paper is, first, to pin down some of the theoretical and methodological barriers that have prevented researchers to arrive at more conclusive evidence for the use of decompositional mechanisms, and, second, to propose a neuropsychological approach that can solve some of the pending issues in the field.

The paper starts by critically examining three representative models of word processing. We then cover commonly used experimental paradigms, together with the results that have been obtained from such experiments. Theoretical and methodological limitations will be examined in detail, in order to make sense of the apparently contradicting results in the literature. Finally, we will present the prospects brought about by the use of a relatively new paradigm to test this question: The mismatch negativity (MMN) response and, specifically, its syntactic version, the syntactic mismatch negativity (sMMN). We end with some concluding remarks about the possibility for future research under this paradigm.

2 Models of morphological processing

Language use involves the perception of sensory stimuli (acoustic variations over time, or visual inputs), of interpretation of hierarchically structured information (acoustic form, grammar) and of meaning interpretation (semantics, pragmatics). Its scientific understanding is therefore regarded as a window into the workings of human cognition. Indeed, the two main camps in models of word comprehension are based on two different views of how cognition is organized. This debate can be traced back to the philosophical camps of rationalism and empiricism. While the former sees cognition as the result of symbol manipulation through a constrained set of computations, the latter assumes cognition starts as a *tabula rasa*, or blank slate, and works its way up to higher cognition from processing of specifics in the environment and their co-occurrences. In what follows, two models will be presented that to a certain extent embody these two sides.

2.1 Single-mechanism connectionist model

Based on the notion that cognitive processes must be instantiated not only at the level of but exclusively *by* neurons or neuron-like units (Rumelhart & McClelland 1986) single-mechanism models were proposed where all words, regardless of morphological complexity, are represented as inter-connected networks of transistor, neuron-like units. These units become active (go from off to on) if enough incoming information agrees with the features these units instantiate. Figure 1 shows how a past tense form in English is represented. The pattern of activation that corresponds to the surface form of the past tense is achieved by the interconnectedness of the units of the network.

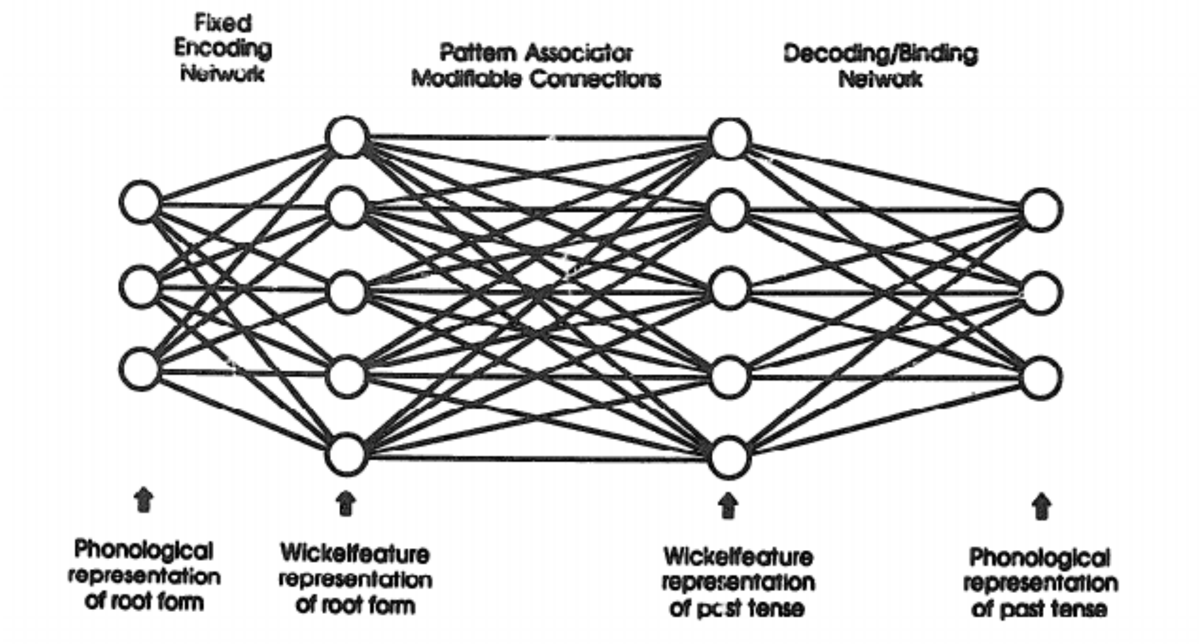

Figure 1: Rumelhart and McClelland's model (1986: 222): Each node represents a property of the input item, and can be either on or off

Co-occurrence of sound or letter on the one hand, and of meaning on the other, creates connections between lexical representations that are related by sound or meaning, thus boosting activation levels for related lexical entries. In order to tune the model to the context-dependent representation of letters in language, the model includes intermediate levels of representations, where so-called wickel-features represent not only each letter, but the immediate surrounding of each letter. The word *kat*, for example, is represented by 3 wickel-features: ${}_{\#}K_{A}$, ${}_{K}A_{T}$, and ${}_{A}T_{\#}$. In theory, this model only presupposes the level of acoustic form and of content (semantic) as determining the connections between words, thereby eliminating the level of morphological information. Morphologically related words like *hond* and its plural *honden* would have different representations, with connections between the two based *only* on their phonological and semantic similarities.

Single-mechanism models based purely on lexical storage have not been able to offer a satisfactory account of the phenomenon of morphological processing. In empirical terms, this type of single-mechanism model attempts to replicate behavioral data of production and comprehension via simulations run under a connectionist architecture (see, for example, Christiansen, Conway & Curtin 2000). Success in mimicking the behavior of humans is interpreted as evidence that language processing proceeds as in these neural networks, although these models are often criticized for oversimplifying the input given to the models (Pinker & Ullman 2002). There are, however, more fundamental reasons why a model that does away with morphology cannot successfully account for how morphologically complex words are processed. Linguists tend to argue whether the different morphological versions of a word are constrained paradigmatically (choices within a given inflectional paradigm) or syntagmatically (choices given by surrounding words in the sentence), but not whether the morphological level is there in the first place (Clahsen 2006). This is a fact that has been accepted amongst the great majority of researchers that study language.

Furthermore, the argument that morphology can be reduced to acoustic form and semantics can be dismissed with purely linguistic arguments: The phenomenon of allomorphy (where different acoustic forms have the same grammatical function) shows that suffixes can be independent of their acoustic form. For example, what is known as the English past tense *-ed* suffix is, in reality, three different acoustic forms: /d/, /t/ and /ɪd/. Each of these is used for the same purpose (to mark past on a verb) but their use depends on the last sound of the stem to which they are added ('walk' receives /t/ because of the voicelessness of its final consonant, whereas 'add' receives /ɪd/ because of the final /d/ in the stem). There are also cases where the same form in a suffix can instantiate different meanings. For example, in Dutch the simple present can mean a present action, as in *Ik werk nu* 'I work now' or a fact about the world, as in *Ik werk aan de universiteit* 'I work at the university'. These examples show that the grammatical function of a suffix and its surface realization are two distinct phenomena (Marantz 2013). But as Yang (2017) has argued, models that concentrate exclusively on the implementation level (in this case, the neuron-like units and their success at mimicking human behavior) are fundamentally incomplete, as this level should be the instantiation of algorithms, which in turn should be based on theories at the computational (abstract) level.

2.2 Dual-mechanism models

The position in early modern linguistics was that only morphemes were stored, and then put together via combinatorial mechanisms based on symbolic rules (Chomsky 1965; Chomsky & Halle 1968). However, the empirical evidence for the existence of a lexicon with surface-form lexical representations led later models that favored decomposition to include the mechanism of lexical storage in their models. This gave rise to what became known as dual-mechanism models. One of the main exponents of this type of models is Word & Rules (Pinker & Prince 1994; Pinker 1998, 1999). This model argues that regular and irregular morphology are fundamentally different, and it is this difference which triggers, on the one hand, the use of surface-form lexical storage for irregularly inflected and monomorphemic words, and, on the other hand, of rule-based decomposition for regularly inflected words. This model was complemented by, and is partly based on, the procedural/declarative model by Ullman (2004, 2007), which distinguishes a fast, automatic, unconscious mechanism that enacts routines and habits (procedural), and a conscious, slower, volitional mechanism that handles fact-based knowledge (declarative). The procedural/declarative distinction was made equivalent to the grammar/lexicon distinction in Words & Rules, thus suggesting that the reason for the existence of two mechanism (i.e. the dual-mechanism view) is the architecture of cognition and the natural separation of processes into these two domains.

Dual-mechanism models have had a good acceptance in the study of word processing. Moreover, the procedural/declarative distinction from Ullman's model fits nicely into the dual-mechanism models' separation between grammar and lexicon. An important feature of dual-mechanism models was to combine the more linguistically-oriented views of language (rule-based, hierarchically organized, etc.) with the findings coming from neural network simulations that suggest the existence of an interconnected lexical memory.

One unsolved issue in dual-mechanism models is the relation *between* the mechanisms of lexical storage and decomposition. Do they operate in parallel, or are they mutually exclusive? In its original formulation, Pinker (1998) claims procedural processing is truly independent of declarative lexical memory, to the point where procedural mechanisms can proceed without accessing the contents of declarative memory. However, later formulations claim that all morphemes, both free standing lexical units (e.g. *dog, walk, happy*) as well as bound morphemes (e.g. the *-s* in *dogs*, the *-ed* in *walked* or the *-ly* in *happily*), are all stored in lexical memory (Pinker & Ullman 2002). If this is the case, then regardless of whether a word is represented in its surface form in lexical memory, or if it is created online via rule-based decomposition, the language mechanism will *always* need to access lexical memory. In other words, there is no decomposition without lexical access. This renders the Words & Rules model empirically unfalsifiable, since evidence of access to the mental lexicon cannot be taken as evidence that the decompositional mechanism is not used.

Some premises of dual-mechanism models are based on sound assumptions on the working of cognition in general, but ill-informed with respects to knowledge from the study of language. For example, The Words & Rules model claims the storage versus decomposition balance in language is related to the known distinction between regular and irregular inflections. The regular/irregular distinction can be rephrased as a matter of structural transparency. i.e. how easily can the structure of a word be inferred from its surface form. Conversely, words would be structurally opaque when the language mechanism cannot infer the internal structure of the word from its surface form. The latter scenario would result in the storage of the opaque word as a whole chunk of sound-meaning. For example, a verb like *kussen* 'to kiss' in Dutch forms its past tense by adding the past tense suffix *-te*. The high structural transparency of the inflected form *kuste* would result in decomposition into the constituents *kus-* and the past tense suffix *-te*. The strong verb *krijgen* 'to get', however, forms the past tense through ablaut (*kreeg)*, thus leaving no surface evidence of added structural elements (like suffixes).[1] The past tense of a verb like *krijgen*, then, would be predicted by dual-mechanism models to be lexically stored in its surface form.

The reliance on surface forms, however, is problematic. While it seems reasonable to assume more frequently encountered forms are processed faster and with less use of cognitive resources, the role assigned to frequency in these psycholinguistic paradigms stands in contrast with more linguistically-informed models. While Words & Rules assumes that word representations in the brain must "look like" words as used in the visual or written domain, important linguistic theories see words as the end results of the merging and moving of abstract features into lexical units (e.g., Chomsky 1995; Harley & Noyer 1999), and therefore the surface form of words should not determine how they are processed, or at least not so much that the language parser cannot compute that *kreeg* is the past of *krijgen*. Moreover, Albright (2002) showed that irregulars are just as regular as the so-called regular verbs. The only difference is that the rules of regulars apply to a wider range of items, while the rules of irregulars apply to smaller sets of words. Both sets of rules, however, are equally 'regular' and systematic. For example in Dutch, *krigen* is not the only verb that undergoes such ablaut. Other verbs like *kijken*, *prijzen* and *lijken* also undergo the same, systematic and highly regular phonological modification (*keek*, *prees, leek,* respectively). It seems unreasonable to assume our language processing mechanism are not sensitive to these regularities.

2.3 The Parallel Race Model

Although there are many more models that move between full storage, full decomposition and somewhere in the middle, the scope of this paper is to specifically review a model based on evidence from Dutch. In this regard, the Parallel Race Model (PRM) by Baayen, Dijkstra & Schreuder (1997) stands out as one of the best known models based on Dutch evidence. This model belongs to the family of dual-mechanism models. Although it never explicitly adopted the premise of the Word & Rules model about the key role of regularity and structural transparency, it did empirically test the hypothesis by comparing frequency effects across a wide range of Dutch inflectional paradigms, like

[1] A more striking example of irregularity in Dutch is the verb *zijn* (to be), where the whole form of the verb changes for every subject in which the verb is conjugated. Our purpose is to argue that irregular verbs inflected through ablaut can still exhibit regularities that could make them candidates for systems of decomposition, but we make no claim regarding special cases like the verb *zijn*, which we assume must be stored separately for every conujugation.

singulars and plurals (ibid; Schreuder 2003), past tense (Bertram, Schreuder & Baayen 2000), perfect participle and plural past tense (Baayen et al. 2002), agentive, second and third person, comparatives, and regular and irregular past participles (Bertram, Baayen & Schreuder 2000). The original formulation of the model assumes that both mechanisms are in principle equally likely to process a word, and that only the 'fastest' one will ultimately yield the target representation.

Despite its name, the term *parallel* in the PRM is misleading. The mathematical instantiation of the PRM assumes lexical storage of surface forms is the first mechanism to act (the default one). This implies that decomposition is a parallel mechanism (i.e. acting at the same time), but rather a complementary one that is used *if* lexical retrieval fails or takes too much time to retrieve the relevant lexical representation. Given that lexical access is a faster mechanism and also the default one, it becomes difficult to imagine a scenario where the decomposition mechanism could retrieve the representation faster than direct lexical access. In fact, the authors of this model do not explicitly state the conditions that would have to be met for decomposition to prevail over lexical access. Notice that having lexical storage as the default mechanism is also contradictory with the notion by Ullman (2004, 2007) that procedural processes (decomposition) are executed faster and more automatically than declarative processes (lexical access).

The PRM has the same limitations described above for dual-mechanism models: It lacks an explicit relation between mechanisms, does not offer experimental predictions that could support the existence of decompositional mechanisms, and is based on premises that might not hold as strongly as was once thought (e.g. the regular-irregular distinction). In fact, the original architecture of the PRM, which included a decompositional mechanism, was not supported by the experiments carried out in Dutch morphological processing (see next section for details). In fact, it was found that a wide array of inflections (plurals, past tense, participles, agentives, comparatives, and other suffixed words) were all stored in the lexicon as whole forms. The authors of the model assumed that because the suffixes they tested had more than one grammatical function (e.g. *-en* does not only mark plural nouns, but also plural verbs and infinitives), the system was more efficient storing them as independent lexical units. This would suggest that the mechanism of decomposition into constituent morphemes is actually not part of the linguistic repertoire of Dutch speakers. It comes as no surprise, then, that some of the authors of this model have recently come to propose models where there is actually no built-in mechanism for decomposition (Baayen et al. 2011).

The results from experiments that tested the PRM oppose not only the assumption of most dual-mechanism models that there must be some sort of decompositional mechanism, but also the experimental data from other languages. Evidence for decomposition, although harder to elicit reliably, has been found in many languages (e.g. Diependaele, Sandra & Grainger 2005; Feldman et al. 2010; Diependaele et al. 2011; Zhang, Van Heuven & Conklin 2011; Morris & Stockall 2012). A lack of decomposition in Dutch would mean that either Dutch is a special language, or that something about the way in which the data was obtained prevented Dutch researchers from replicating the findings of other languages. A critical examination of the premises behind dual-mechanism models, and the PRM in particular, suggests that these models are not offering complete enough hypotheses to test competing theories about word and morphological processing. In what follows we will examine the empirical evidence that led some authors to exclude decomposition from Dutch. We will argue that, together with the theoretical gaps presented in this section, methodological deficiencies are partly responsible for the lack of empirical evidence for decomposition in Dutch. In the last section we will propose an electrophysiological paradigm that could answer the pending question of whether there is decomposition in Dutch, or whether Dutch speakers rely solely on stored lexical items.

3 Empirical Evidence

Evidence from comprehension studies that favor decomposition of regularly inflected words has mainly been obtained through masked priming experiments (Forster et al. 2003).

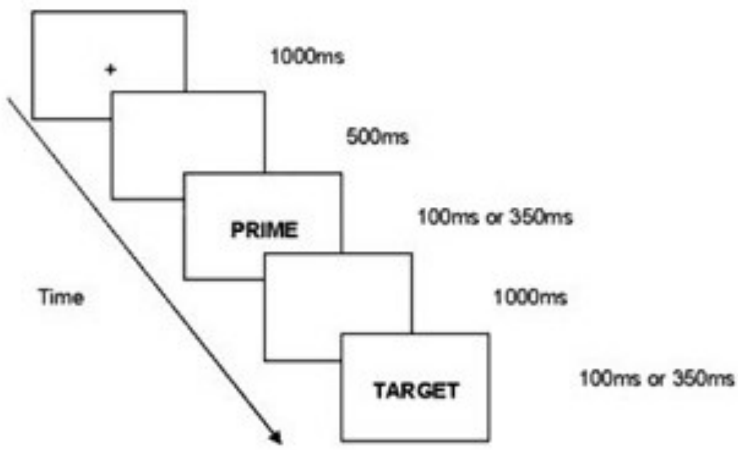

Figure 2: The masked priming paradigm: The prime is presented for such short duration that the participant is unaware of having seen it. Nevertheless the prime is often found to affect processing of the target word, measured in response time and accuracy

In these experiments (see Figure 2) the recognition of a word is preceded by the subliminal presentation of a morphologically related word (e.g. *kuste* priming *kussen*). If the two words are the manifestation of a single representation, as decomposition-based models would predict, then processing *kuste* should facilitate the recognition of *kussen* because both would be processed by accessing the same stem *kus*-. So far, most masked priming experiments have found evidence for decomposition in inflected and, to a lesser extent, in derived words. Interestingly, recent studies using masked priming have challenged the hypothesis of Words & Rules that irregular words should be lexically stored. Gor & Cook (2010) showed priming for regular and irregular verbs in Russian. Feldman et al. (2010) showed morphological facilitation was present for regulars and irregulars in native English speakers. Additionally, it was found that non-native speakers with morphologically rich languages were able to decompose complex English verbs regardless of regularity. Decomposition has even been found based purely on orthographic input (Rastle, Davis & New 2004), suggesting this might be a very early, automatic and over-reaching mechanism in word processing (for a review of evidence obtained from priming studies, see Amenta & Crepaldi 2012).

In Dutch, most of the evidence comes from behavioral experiments on morphological processing where the main dependent variable is response time. In such designs, frequency counts (e.g. how frequent a word or a part of a word is found in language corpora) are a commonly used independent variable (Davis, van Casteren & Marslen-Wilson 2003). The premise of the frequency effect as evidence of lexical storage is straightforward: If it is assumed that complex words are (de)composed online during processing via activation of its constituent morphemes, then frequency counts based on the constituent morphemes (the stem and/or the suffix) should predict the ease of access to the word instead of the surface frequency of the whole word. On the other hand, if a word is stored in its surface form, it means that the lexical representation of a word matches its acoustic form when heard. In such cases, the frequency of that acoustic form should predict how fast the word is responded to. For example, if the past tense of *kussen* (*kuste*) were decomposed, the surface frequency of *kuste* should not determine the reaction time to the word. Instead, only the frequency of the stem *kus*- should predict how fast people can recognize the word. Conversely, if *kuste* is stored as a lexical entry, then the frequency of *kuste*, as a whole unit, should predict how fast the word is recognized. By controlling for one type of frequency (e.g. base) and manipulating another type (e.g. surface) researchers can test whether a word is being decomposed or retrieved as a whole from lexical memory.

There are, nonetheless, a few problems with the notion of frequency effects. Frequency effects are really correlations (between a certain frequency count and response times) and, as such, the phenomenon can also be explained as the outcome of two interacting variables, instead of as the process of a single one (lexical retrieval). Lignos and Gorman (2012) introduced the concept of suffix conditional probability, which is expressed in the following way:

$$p\ (suffix \mid base) = p\ (base,suffix)\ /\ p\ (base)$$

This mean that a surface frequency effect can be conceptualized as an interaction between the frequency of the stem and the conditional probability of the suffix given that particular stem. In other words, surface frequency effects could be explained as the outcome of a decompositional process. In fact, early in the study of morphological effects on word processing, Taft (1979) showed that surface frequency effects might only be tapping the end-result of decompositional processes. Indeed, if a word like *kuste* is decomposed into *kus-* and *-te*, the end result of such decomposition is still a surface-form representation of *kuste*. The key difference with a model that assumes lexical retrieval is that the surface form is created online, instead of being retrieved as a whole chunk from lexical memory. Until there is evidence that shows that the cause of the correlation found between high frequency words and faster recognition times is the retrieval of surface forms, the alternative hypothesis implying decomposition cannot be discarded.

The bulk of evidence upon which the PRM was based on came from lexical decision tasks that manipulated the variable of surface frequency. Participants saw a target word and had to decide whether the word belonged to their language or not. One of the earliest studies on Dutch morphological processing was done by Baayen, Dijkstra & Schreuder (1997), replicated by Schreuder (2003). The 1997 study found marginal evidence for decomposition of the plural *-s* suffix, as well as for the past tense *-te* and *-de* suffixes (via lack of surface frequency effects). Surprisingly, another plural allomorph (the *-en* suffix) was affected by surface frequency, suggesting plural *-en* Dutch nouns but not the plural *-s* ones, would be stored as whole words rather than decomposed. In the replication study of 2003 however, the authors found no evidence of decomposition for any suffix, including the plural *-s* and the past tense suffixed words. The authors hypothesized that the frequency difference was not big enough between low and high frequency past tense words in the 1997 study. Indeed, choosing more extremes values in the 2003 experiments yielded significant surface frequency effects for past tenses. Regarding the results of the plural *-en* words, the authors concluded that the ambiguity of the *-en* suffix (it marks both infinitive and plurality in verbs), and the change of the phonological form of the nominal stem when adding the plural *-en* suffix, caused plural *-en* nouns to be retrieved more efficiently via direct lexical access. However, these hypotheses were not empirically tested, but rather offered as possible explanations of the data *after* the results were examined.

The experiments in Dutch morphological processing were useful in showing that words with different degrees of morphological complexity were processed differently by native speakers. Such findings, based on significant response-time differences between high and low frequency items, paved the way to further investigate what it is that make some Dutch words show evidence of lexical storage. One possibility is methodological: The experiments by Schreuder (2003) made use of predictor variables that were likely correlated, thus introducing multicollinearity in their study. Furthermore, the use of dichotomous frequency values (high or low) is only useful in cases where data collection is limited, difficult or too expensive, which is not the case when looking for words in modern linguistic corpora. These practices increase the likelihood of type II errors (false positives). Factorized designs (that divide variables into discrete categories, like high and low frequency) are also limited to answering whether there is an effect of a variable or not, but cannot say much about the interactions between variables that lead to the observed effects (Baayen 2004). This is why the studies that test the PRM are constrained to yes or no research questions (such as “is there evidence of storage or decomposition?”). But this type of research questions are ill-suited for models that already assume the existence of both processes (like dual-mechanism models). Instead, experiments testing these models should be asking how different variables affect the way each mechanism behaves, rather than asking whether the mechanism is there or not. It is no wonder, then, that the results of these experiments only allow the researchers to speculate on possible explanations for their results, instead of testing such post-hoc hypotheses experimentally.

There is also a need to better control for variables that might play a role in how words are responded to. In experiments 3 and 4 of Schreuder (2003), high frequency past plural verbs were responded to faster than low frequency past plural verbs, suggesting these are lexically stored (otherwise they should show no surface frequency effects). However, in these experiments, infinitive verbs were compared with plural verbs that were also inflected for past tense, and that were, moreover, irregular. Setting aside the fact that an infinitive was being compared with a verb inflected in both past and plural, it is surprising that the authors used irregular verbs, since almost all dual-mechanism models would predict different effects for regular and irregular verbs. There is a contradiction between assuming structural transparency plays a role, and collapsing regulars and irregulars into one group. It is possible that the different responses that were obtained were due to the words being linguistically highly dissimilar in terms of word category and regularity of the inflection.

The choice of task can also have important consequences for how a phenomenon is investigated. Contrary to most psycholinguistic research exploring morphological processing, the evidence gathered for the PRM was consistently obtained under unprimed lexical decision tasks. In this way, Bertram, Baayen & Schreuder (2000) showed that inflected adjectives in the comparative *-er* suffix (*groot* - *groter* 'big - bigger'), past tense verbs and derived nouns with the agentive *-er* suffix (*gebruiken* - *gebruiker* 'use - user') were sensitive to family size (a type of surface frequency count), with higher family size predicting faster responses, but not to family frequency (a count sensitive to the constituents of complex words). It remains to be seen if these results would hold in an experimental paradigm that taps into earlier stages of word processing, like masked priming experiments have done.

There are also more fundamental problems with the "semantic ambiguity" hypothesis of the PRM. Namely, it contradicts its own premise that lexical retrieval of surface forms is the first mechanism to act, and that rule-based decomposition acts only when memory fails to provide an adequate representation. Assuming that ambiguous suffixes are stored as surface-forms *because* decomposition takes too long, implies that decomposition occurs first, and that only when there are specific obstacles, like a polysemantic suffix, lexical retrieval should steps in. Indeed, the explanations given to account for the results in the Dutch experiments claim that lexical retrieval, at least in the case of ambiguous suffixes, acts because parsing failed to provide an adequate representation. These models, then, seem to lack appropriate predictions for how their data should behave under different morphological circumstances. By focusing exclusivly on whether there are differences or not, the question of what triggers the observed differences has been neglected, leaving the field with incomplete models and unclear empirical predictions to test.

4 On the need for a neuropsychological approach: The syntactic mismatch negativity response

We have reviewed some of the reasons why experiments in Dutch morphological processing have not offered convincing evidence for the claim that all Dutch complex words have lexically stored surface forms. There are, however, more fundamental reasons why these types of experiments are unlikely to solve the pending issues in the debate of storage versus decomposition. In the beginning of the previous section we outlined some of the evidence that has favored decomposition of complex words in other languages. Most of this evidence is based on experimental paradigms that are much more sensitive to early cognitive processes than the lexical decision tasks used in the Dutch experiments reported here. Lexical decision tasks require cognitively demanding decisions to be made regarding the words before a response can be made: A participant must process the words at all relevant levels (phonological, possibly grammatical, semantic, etc) before being able to make a conscious judgment on the lexicality of the word. Given that the studies testing the PRM did not use tasks that tap into earlier processes (like masked priming), it should come as no surprise that no evidence for decomposition was found.

An example of how a neuropsychological approach can offer different results is found when comparing behavioral and neurophysiological experiments that study the same phenomenon. For example, Bertram, Baayen & Schreuder (2000) found no difference between inflected and derived words when comparing the effect of family size, suggesting these two types of words are stored as whole-forms in lexical memory. However, other studies that are able to tap into early processes (e.g. Bozic & Marslen-Wilson 2010; Leminen et al. 2013) do find evidence for decomposition not only in inflected, but also on derived words. This suggests that decompositional processes could be part of early mechanisms that cannot be probed by behavioral responses, such as those elicited by lexical decision tasks. If this was the case, what the authors classify as evidence of surface-form lexical storage would simply correspond to the inevitable activation of a lexical representation at the last stages of the word comprehension process, even if the process consisted in decomposition of the constituent elements (Taft 2004).

Indeed, many researchers have pointed out the fact that behavioral responses can only show us how the processes are affected after said processes have finished (Taft 2004; Van der Lely 2005; Morgan-Short et al. 2012). Behavioral experiments do not allow us to know what is happening while words are being processed, but can only assess effects on word processing at the response stage, which is after processing has taken place (Clahsen & Neubauer 2010; Diependaele et al. 2011). Since recent evidence indicates that morphological comprehension occurs at the earliest stages of word processing, i.e. only a few hundred milliseconds after the stimulus is recognized (Pulvermüller & Shtyrov 2006; Hasting, Kotz & Friederici 2007; Morris & Stockall 2012), it is likely that more time-sensitive measurements will offer a clearer picture of how lexical access and rule-based decomposition interact in word processing.

A type of measurement that is particularly suited to capture such time-sensitive phenomena is the event-related potential (ERP). This is a neurophysiological response obtained by recording electrical activity from the cortex through electrodes placed on top of a participant's scalp. Temporal resolution is perhaps the biggest advantage of using ERP responses. It gives the possibility of comparing phenomena very early in the time course of the cognitive processes they instantiate (usually within the first second after stimulus presentation). ERP responses have been used to complement the behavioral responses obtained from linguistic tasks. For example, ERP responses can be used in masked priming experiments that use lexical decision tasks. Because ERP responses are time-locked to the presentation of the stimuli (usually the target word), they can show responses to the processing of the word that are much earlier than the motor responses of lexical decision tasks.

Importantly for psycholinguistics, different ERP components have been shown to be sensitive to different levels of linguistic analysis. The N400, for example, has been associated with integration of unexpected (semantic) information to an ongoing parsing process. The P600, on the other hand, has been linked with revision processes that result from ungrammatical or hard to parse sentences. Despite its sensitivity to syntactic information, the P600 component is a response that is assumed to occur after the sentence has been parsed, and thus it does not tell much about what goes on during the actual parsing of words.

The general picture of the neurophysiology of language is that form and phonology produce the earliest responses, while higher level processes like morphosyntax, semantic and pragmatics, would occur later on (see e.g. Friederici 2002; Friederici & Kotz 2003). However, this view has been challenged by recent evidence that shows very early responses to morphosyntactic and even lexical processing. For example, the early left anterior negativity (ELAN) peaks around 100 to 250 ms after encountering a word that does not fit with the phrase-structure being processed. A slightly later component (the left Anterior Negativity, or LAN) peaks around 300 ms after encountering an ungrammatical word (for a summary of recent findings, see Pulvermüller & Shtyrov 2006).

Some ERP studies on morphological processing have challenged many of the theories based on behavioral evidence. For example, the M350 study by Stockall & Marantz (2006) challenges the hypothesis of Word & Rules that structural transparency is a requisite for decomposition. Similar results were obtained by Morris & Stockall (2012), who showed significant priming for both regular and irregular past tense in an early (approximately 200 ms) response, although a difference emerged at later behavioral measures. In general, ERP studies on morphological processing suggest that morphological effects occur much earlier than previously thought, and that they might be more widespread across types of words than was assumed.

There are, however, important drawbacks to the use of ERPs in pyscholinguistic research. First, although this is also the case in behavioral experiments, responses in more classic psycholinguistic experiments are dependent on particular tasks (e.g. a button press in a lexical decision task), and so the brain response gets contaminated by task-specific strategies that do not represent the mechanisms under study (Pulvermüller & Shtyrov 2006). Another disadvantage of ERPs is known as the physical variance problem (Pulvermüller 1999). In order to get an ERP response, an average must be made of several individual responses. However, each response is unique, since in linguistic experiments, like masked priming, all target words are different. If the point at which a word is recognized is different for every word (e.g. because of length differences), the processes that later lead to syntactic and semantic comprehension will also occur at different times. This physical variance problem raises doubt about how representative an ERP response is of a particular linguistic process, given that the average ERP response includes physically distinct stimuli that elicit different ERP responses. There is also a great deal of individual variation that is overlooked on grand average ERPs. For example, Tanner (2015) showed that responses that are considered to represent different stages, like the N400 and the P600, could really be artifacts of the average of one big biphasic response that varies across individuals. Likewise, Helenius et al. (2007) showed that the LAN and the N400 component share the same left-anterior temporal lobe generator, which together with their close temporal proximity might suggest a common underlying response.

The mismatch negativity (MMN) paradigm (Näätänen, Gaillard & Mäntysalo 1978) is able to solve some of these drawbacks in ERP research. It was first used as a diagnosis tool to check for healthy acoustic change detection in the auditory system. In the context of a passive odd-ball design, a repeating sound, referred to as the standard stimulus, is acoustically presented in a consecutive manner. The presentation of the standard sound is interrupted by the occasional appearance of a sound which deviates from the standard stimuli on a particular acoustic feature (amplitude, frequency, pitch, etc.). The MMN component is obtained by subtracting the averaged response to the standard stimuli from the averaged response to the deviant ones. By using only two acoustical stimuli that differ on a concrete acoustic feature, the problem of physical variance is solved. The MMN allows to average the response to the same word (the deviant stimulus), thus offering a response that is truly representative of the processing for that particular word. The MMN component can also be obtained even if the participant is not attending the stimuli, which in accordance to the time-range in which the response takes place (100–200 ms) classifies the process as a pre-attentional, automatic, unconscious one. This eliminates the task contamination problem, since the MMN response can be obtained without a mediating task (e.g. in the absence of attention).

Of importance for psycholinguistics, the MMN response was found to be more pronounced (i.e. with a greater amplitude) when using existing linguistic stimuli as deviants (first syllables, and eventually words and even phrases) compared to deviants that are not part of the linguistic repertoire of the participant (e.g. pseudowords or ungrammatical words) (Näätänen et al. 2001). For example, a hypothetical MMN design that presented Dutch native speakers with a standard stimulus *ka* (a pseudoword) and a deviant one *kal* (also a pseudoword) would yield a smaller MMN response than a design using the same standard *ka*, but using as deviant the string *kat* (cat), which corresponds to an actual Dutch word.

This language-specific response was interpreted as memory traces that became activated when listening to a deviant stimulus that already exists in the participant's (lexical) memory, compared to another deviant stimulus that is not part of the participant's linguistic repertoire, and thus has only a transient memory trace. The enhanced response to words of the participant's language, when compared to the response evoked by non existing words, is known as the lexical MMN (Pulvermüller & Shtyrov 2006). Contrary to the N400, then, which reflects integration of semantic information that has already been processed, the lexical MMN seems to reflect the moment when the lexical entry is accessed, which in terms of language processing models could correspond to the notion of lexical retrieval.

The MMN has also yielded evidence that indicates the existence of a decompositional mechanism in morphological processing. Shtyrov & Pulvermüller (2002) showed that the inflected verb *comes* produced a MMN maximal at centroparietal sites, while the non-inflected stem *come* peaked in frontal sites. This suggest different memory traces for the stem and the third person suffix, which is in accordance with a decompositional model that assumes separate storage of free-standing and bound morphemes. Moreover, the latency was slower and the amplitude smaller for the inflected *comes* compared to the MMN of *come*. This would suggest that the retrieval of the constituents precedes the (de)composition process.

There is another modulation of the MMN which is specific to morphosyntactic processing. It has been hypothesized that this syntactic or combinatorial MMN response could be a result of an automatic decompositional process, triggered by the detection of complex, decomposable stimuli. Given a morphologically complex deviant stimulus, the linguistic constituents would be analyzed as parts of a larger, decomposable structure. By the time the second element is processed, the brain response is less pronounced, since the first element predicts the appearance of the second. Shtyrov et al. (2003) showed that the MMN to the word *come* was more negative (larger amplitude) than the MMN to the string *we come*. This could be explained by assuming *we come* was decomposed by participants, and thus the response to *come* was primed, and thus smaller, when preceded by *we*. This is reinforced by the fact that the MMN to *comes* did not reliably differ from the ungrammatical **we comes*, suggesting the MMN was not enhanced by ungrammatical strings, but rather reduced by the grammatical string. This syntactic MMN (sMMN) presents the opposite pattern of the lexical MMN. In the latter, the response to existing words shown as deviants are larger than the response to non-existing words shown as deviants. However, for the sMMN the smaller response is to grammatical, rather than to ungrammatical stimuli.

Hasting, Kotz & Friederici (2007) used a design with subject-verb and phrase structure agreement to study morphosyntactic processing in German. In one condition the stem was preceded by either a first or third person pronoun, yielding grammatical and ungrammatical subject-verb agreement sentences. In the other condition, the same verb stem was preceded by an indefinite article, resulting in an ungrammatical phrase structure combination. In both conditions the MMN for ungrammatical strings was larger than that elicited by grammatical ones, although they differed on their topography. These results were replicated in German with stems that were ambiguous to their word class, thus controlling for acoustic confounds when comparing phrase structure (determiner + noun) and subject-verb agreement responses (Hasting & Kotz 2008).

The sMMN was also found for morphologically complex words (i.e. stem plus affix), which suggests a common mechanism for the processing of morphologically complex words and sentences. Leminen et al. (2013) compared high and low frequency inflected and derived Finish words shown as deviants in a MMN experiment. In order to make sure that different MMN responses were not simply due to acoustic differences between standard and deviant stimuli, they had four other pairs containing pseudoword deviants that replicated the acoustic change in the inflected and derived word conditions. They found that all deviants, both inflected, derived and pseudowords, produced MMN responses, which is to be expected if the deviant is acoustically different from the standard. Crucially, they found that "inflected" pseudowords had a bigger MMN than the real inflected word. By contrast, the MMN caused by the "derived" pseudoword was smaller than that by the real derived word. A non-existing word eliciting a larger MMN than a real word is precisely the definition of the syntactic MMN. Conversely, a real word eliciting a bigger MMN than a non-existing word is the response known as lexical MMN. These results, then, are evidence that the inflected Finnish verb is decomposed, while the derived finish verb seems to be accessed lexically through its surface form. Interestingly, high-frequency deviants elicited a bigger MMN response than low frequency only in the 90–130 ms range, but not in the later 130–170 ms range. This shows how important it is to look into the time-course of early processing to disentangle early from late effects. Similar results were obtained by Bakker et al. (2013), who found a sMMN for English past tense verbs, which was independent of frequency counts of the stimuli. These results corroborate the hypothesis that decomposition processes occur very early and are thus pre-lexical.

The evidence gathered from the MMN paradigm indicates that the activation of a word's memory trace can occur very early in word recognition, that is, before 200 ms. This means response-based behavioral measures are temporally not suited to tap into this process. Moreover, the syntactic or combinatorial MMN response, which occurs at similar latencies, has provided robust evidence that the processing of morphologically complex strings (either words or phrases) is affected by their constituent elements. The MMN paradigm has also contributed to the control of possible confounding variables, like differences between stimuli and task distraction. However, these strengths are also weaknesses of the paradigm. Its limited set of stimuli makes it necessary to replicate the effects on different words and types of morphosyntactic complexity. Moreover, the effects have so far been obtained in English, Finnish and German, and so the replication of the effects on other languages is also a necessary step to confirm the hypothesis of decomposition mechanisms in morphological processing. These weaknesses, however, should be seen as an invitation for researchers working on morphological processing in Dutch. Further attempts at replicating for Dutch the sMMN found for other languages, and of extending them to the specific Dutch inflectional paradigms are necessary to have more complete models of morphological processing.

5 Discussion

We have argued that the Parallel Race Model has theoretical predictions that are inconsistent with (the interpretation of) their empirical data. Theoretically, the Parallel Race Model predicts that lexical access should be earlier than decomposition. However, a closer examination of their proposed explanations reveals that these accounts are based on the premise that decomposition is the first mechanism to act. Moreover, whereas the model allows, in theory, a separation of the lexicon into the two mechanisms, the experiments reviewed here consistently fail to find evidence of decomposition, and instead find only evidence of lexical access. We believe it is unlikely that Dutch behaves so differently from similar languages that have shown evidence for decomposition. Instead, we have argued that these results are likely due to incomplete theoretical accounts and methodological deficiencies.

Recent evidence has shown that morphological processing occurs very early and pre-attentionally, which posses a challenge to traditional experimental paradigms in psycholinguistics. The MMN component offers a way to compare early responses to linguistic stimuli that can be matched for acoustical properties, and presented outside the participant's awareness, thus shedding light on aspects of linguistic processing that are automatic and pre-attentional. It can also offer distinct responses for direct lexical activation (the lexical MMN) and combinatorial processing (the syntactic MMN). This opens the possibility of investigating which aspects of Dutch morphosyntax are processed automatically and unconsciously, and specifically via which mechanisms. A further possibility opened by the MMN design is to investigate whether the mechanisms that decompose complex words are the same, as predicted by current linguistic theories (Harley & Noyer 1999; Marantz 2013), or different from those that decompose phrases and sentences.

In sum, using the MMN paradigm to study Dutch morphosyntactic processing opens the possibility to better understand the peculiar case of Dutch morphology. The evidence gathered from such experiments could help broaden the explanatory capacity of current models, by showing how morphological processing occurs in what is perhaps the earliest time window to which current experimental methods have access to. Results that replicate the findings in Finish, English and German would provide further support to a model where decomposition plays an important role in complex word processing. However if the results of a MMN experiment in Dutch follow the line of evidence gathered from behavioral studies, the need would rise to explain why speakers of Dutch seem to dismiss the use of a decompositional mechanism, regardless of word complexity.

References

Albright, A. 2002. Islands of reliability for regular morphology: Evidence from Italian. *Language* 78(4): 684–709.

Amenta, S., & D. Crepaldi. 2012. Morphological Processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology* 3(232): 99–110.

Baayen, H., T. Dijkstra & R. Schreuder. 1997. Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language* 37(1): 94–117.

Baayen, H. 2004. Statistics in psycholinguistics: A critique of some current gold standards. *The Mental Lexicon* 1(1): 1–47.

Baayen, H., R. Schreuder, N. de Jong & A. Krott. 2002. Dutch inflection: the rules that prove the exception. In K. Nooteboom, F. Weerman & F. Wijnen (eds.), *Storage and Computation in the Language Faculty*. Dordrecht: Kluwer Academic publishers, 61–92.

Baayen, H., P. Milin, D. Đurđević, P. Hendrix & M. Marelli. 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118(3): 438–81.

Bakker, I., L. MacGregor, F. Pulvermüller & Y. Shtyrov. 2013. Past tense in the brain's time: Neurophysiological evidence for dual-route processing of past-tense verbs. *Neuroimage* (71): 187–195.

Bertram, R., R. Baayen, & R. Schreuder. 2000. Effects of family size for complex words. *Journal of Memory and Language* 42(3): 390–405.

Bertram, R., R. Schreuder & R. Baayen. 2000. The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(2): 489–511.

Bozic, M. & W. Marslen-Wilson. 2010. Neurocognitive contexts for morphological complexity: Dissociating inflection and derivation. *Language and Linguistics Compass* 4(11): 1063–1073.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Special technical report Massachusetts Institute of Technology. Research Laboratory of Electronics. Cambridge: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. & M. Halle 1968. *The Sound Pattern of English*. New York: Harper & Row.

Christiansen, M. H., C. M. Conway & S. Curtin. 2000. A connectionist single-mechanism account of rule-like behavior in infancy. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society:* 83–88.

Clahsen, H. 2006. Linguistic perspectives on morphological processing. In D. Wunderlich (ed.), *Advances in the Theory of the Lexicon*. 355–388.

Clahsen, H. & K. Neubauer. 2010. Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua* 120(11): 2627–2637.

Davis, M. H., M. van Casteren & W. Marslen-Wilson. 2003. Frequency effects in processing inflected Dutch nouns: A distributed connectionist account. In H. Baayen & R. Schreuder (eds.), *Morphological structure in language processing*. Berlin: Mouton de Gruyter. 427–462.

Diependaele, K., J. A. Ducabeitia, J. Morris & E. Keuleers. 2011. Fast morphological effects in first and second language word recognition. *Journal of Memory and Language* 64(4): 344–358.

Diependaele, K., D. Sandra & J. Grainger. 2005. Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes* 20(1): 75–114.

Feldman, L. B., A. Kostic, D. M. Basnight-Brown, D. Đurđević & M. Pastizzo. 2010. Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism: Language and Cognition.*13(2): 119–135.

Forster, K., K. Mohan, J. Hector, S. Kinoshita & S. Lupker. 2003. The mechanics of masked priming. In S. Kinoshita & S. Lupker (eds.), *Masked Priming: The State of the Art.* New York and Hove: Psychology Press, 3–37.

Friederici, A. 2002. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6(2): 78–84.

Friederici, A. & S. Kotz. 2003. The brain basis of syntactic processes: functional imaging and lesion studies. *Neuroimage* 20(1): 8–17.

Fruchter, J. & A. Marantz. 2015. Decomposition, lookup, and recombination: MEG evidence for the full decomposition model of complex visual word recognition. *Brain and Language* 143(1): 81–96.

Gor, K. & S. Cook. 2010. Nonnative processing of verbal morphology: In search of regularity. *Language Learning* 60(1): 88–126.

Harley, H. & R. Noyer. 1999. Distributed morphology. *Glot International* 4(4): 3–9.

Hasting, A., S. Kotz. & A. Friederici. 2007. Setting the stage for automatic syntax processing: The mismatch negativity as an indicator of syntactic priming. *Journal of Cognitive Neuroscience* 19(3): 386–400.

Hasting, A. & S. Kotz. 2008. Speeding up syntax: On the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of Cognitive Neuroscience* 20(7): 1207–1219.

Helenius, P., S. Maury & R. Salmelin. 2007. Localization of syntactic and semantic brain responses using magnetoencephalography. *Journal of Cognitive Neuroscience* 19(7): 1193–1205.

Leminen, A., M. Leminen, T. Kujala & Y. Shtyrov. 2013. Neural dynamics of inflectional and derivational morphology processing in the human brain. *Cortex* 49(10): 2758–2771.

Lignos, C. & K. Gorman. 2012. Revisiting frequency and storage in morphological processing. *Proceedings of CLS* (48): 447–461.

Marantz, A. 2013. No escape from morphemes in morphological processing. *Language and Cognitive Processes* 28(7): 905–916.

Morgan-Short, K., K. Steinhauer, C. Sanz & M. T. Ullman. 2012. Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience* 24(4): 933–947.

Morris, J. & L. Stockall. 2012. Early, equivalent ERP masked priming effects for regular and irregular morphology. *Brain and Language* 123(2): 81–93.

Näätänen, R., A. W. Gaillard & S. Mäntysalo. 1978. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica* 42(4): 313–329.

Näätänen, R., M. Tervaniemi, E. Sussman, P. Paavilainen, & I. Winkler. 2001. Primitive intelligence in the auditory cortex. *Trends in Neurosciences* 24(5): 283–288.

Pinker, S. 1998. Words and rules. *Lingua* 106(1): 219–242.

Pinker, S. 1999. *Words and Rules: The Ingredients of Grammar*. New York: Basic Books.

Pinker, S. & A. Prince. 1994. Regular and irregular morphology and the psychological status of rules of grammar. In S. Lima, R. Corigan & G. Iverson (eds.), *The Reality of Linguistic Rules.* Amsterdam: John Benjamins Publishing Company, 321–351.

Pinker, S. & M. T. Ullman. 2002. The past and future of the past tense. *Trends in Cognitive Sciences* 6(11): 456–463.

Pulvermüller, F. 1999. Words in the brain‘s language. *Behavioral and Brain Sciences* 22(2): 253–279.

Pulvermüller, F. & Y. Shtyrov. 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology* 79(1): 49–71.

Rastle, K., M. H. Davis & B. New. 2004. The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review* 11(6): 1090–1098.

Rumelhart, D. & J. McClelland. 1986. On learning the past tenses of English verbs. In J. McClelland & D. Rumelhart (eds.), *Parallel Distributed Processing, Volume 2: Psychological and Biological Models.* Massachusetts: The MIT Press, 216–271.

Schreuder, R. 2003. Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. *Morphological Structure in Language Processing* 151(1): 355–390.

Shtyrov, Y. & F. Pulvermüller. 2002. Memory traces for inflectional affixes as shown by mismatch negativity. *European Journal of Neuroscience* 15(6): 1085–1091.

Shtyrov, Y., F. Pulvermüller, R. Näätänen & R. Ilmoniemi. 2003. Grammar processing outside the focus of attention: an MEG study. *Journal of Cognitive Neuroscience* 15(8): 1195–1206.

Stockall, L. & A. Marantz. 2006. A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon* 1(1): 85–123.

Taft, M. 1979. Recognition of affixed words and the word frequency effect. *Memory & Cognition* 7(4): 263–272.

Taft, M. 2004. Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology* 57(4): 745–765.

Tanner, D. 2015. On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A commentary on "Grammatical agreement processing in reading: ERP findings and future directions" by Molinaro et al. *Cortex* 66(1): 149–155.

Ullman, M. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92(1); 231–270.

Ullman, M. 2007. The biocognition of the mental lexicon. In G. Gaskel (ed.), *The Oxford Handbook of Psycholinguistics.* New York: Oxford University Press, 267–286.

Van der Lely, H. 2005. Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends in Cognitive Sciences* 9(2): 53–59.

Yang, C. 2017. Rage against the machine: Evaluation metrics in the 21st century. *Language Acquisition* 24: 100–125.

Zhang, T., W. Van Heuven & K. Conklin. 2011. Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science* 22(10): 1237–1242.

Hernán Labbé Grünberg

Universiteit van Amsterdam

Spuistraat 134

1012VB Amsterdam

The Netherlands

hernanlabbeg@startmail.com