Universiteit Leiden

Investigating the learnability of the perfective auxiliary:

Does the interlanguage grammar of English speakers of Dutch reflect a selectional hierarchy?

Author: Wendy van Ruiten-Toye
Student number: 1054376
E-mail: ruiten-translations@hetnet.nl
Master thesis (20 ECT points)
Course: Linguistics
Specialisation: English Language and Linguistics
Leiden University
Supervisor: Dr. J. Caspers
Second reader: Dr. A.K. Lipták
August, 2015

Table of contents

Chapter **Page**

1. Introduction

The concepts within the transformational generative grammar framework put forward by Chomsky and others in the latter half of the twentieth century revolutionised the field of linguistics and insights into language (e.g. Chomsky 1965, 1973, 1980, 1981) . These theories are still widely accepted by linguists today and have provided the basis for innumerable linguistic studies. One of the propositions of the generative grammar theory is that the human species has an innate capacity for language learning, called the Language Acquisition Device, and that all languages share a common grammatical framework referred to as Universal Grammar (Chomsky, 1965). Universal Grammar is an innate cognitive system designed specifically for language acquisition, consisting of a set of general rules that apply to all languages ("principles") and variable properties that are language-specific and are set by language learners on the basis of input ("parameters") (Chomsky, 1981). The reasoning behind the proposition of an innate language faculty is that human language is so complex that it would otherwise be impossible for children to learn their native language so quickly on the basis of the often impoverished primary linguistic data. This argument is commonly referred to as the "poverty of the stimulus" – there seems to be a discrepancy between what a child actually knows and what it could be expected to know on the basis of its experience (Chomsky, 1980). Being born already equipped with a basic grammar framework and a limited number of parameter settings gives children a head start, reducing the possible grammars and making it easier for them to acquire their native language.

Although the notion of a Universal Grammar is widely accepted among nativists such as Chomsky, the nature and properties of such a grammar are the subject of much discussion and theorising, often accompanied by empirical data from different languages in support of the proposed theory. Clearly, if one seeks to establish the universal properties of language and distinguish these universal properties from specific features of a particular language, comparisons between languages can offer compelling evidence. The immature grammars of young children learning their native language can also provide clues as to the underlying structure of human language. Children often make language errors that are not made by adult native speakers and cannot therefore have been learned from primary linguistic data (Pinker, 1994; Thornton, 1990). From this we may deduce that they are creating their own

grammatical structures. It has been suggested that although these primitive grammars may violate the grammar of the native language, they always comply with Universal Grammar (Crain & Pietroski, 2002). Evidence in support of this has been found in many studies of child language, in which the ungrammatical structures produced by children closely resemble structures that are grammatical in other languages unknown to the child (e.g. Thornton, 1990; Thornton, 2008). It is not only young children who make these kinds of creative language errors. Adult learners of a second language have been found to make similar mistakes. The constantly evolving language system produced by adult second language learners, commonly known as the interlanguage, differs from the target language in a number of respects, due to influence from the native language as well as the gradual acquisition of the new language rules (Crystal, 2008). Although this interlanguage may contain structures that are generally considered ungrammatical in the target language, as with first language acquisition it is believed that these structures do not violate Universal Grammar, and parallels with existing languages have been observed (e.g. Slavkov, 2009). If a second language speaker produces ungrammatical L2 language structures that do not exist in his or her native language, it is logical to conclude that there must be an explanation for these errors that goes beyond first language interference. Experiments comparing subjects with different native languages, investigating errors made while the subject is distracted or placed under time pressure, or, as in this paper, investigating the correct use of language features that do not exist in the native language, may isolate transfer effects and enable conclusions to be drawn on aspects such as universality.

In short, the examination of interlanguage grammars can provide a wealth of information on language learning, language structure and language representation and processing. If we assume that second language learners have continued access to Universal Grammar (although this is disputed, as discussed in the next section), language structures considered to be part of or governed by Universal Grammar can be expected to be more robust in the interlanguage of a second language speaker and more easily learnable than structures that are specific to a particular language. Examining the error-sensitivity and learnability of certain phenomena can therefore offer an indication of the universality of underlying language features. After all, Universal Grammar by definition should not have to be learned,

so a language learner should have an instinctive sense of what is grammatical and what is not beyond the language-specific parametric differences.

This paper examines the phenomenon of unaccusativity on the basis of the treatment of unaccusative verbs by English native speakers of Dutch as a second language. The term unaccusativity, initially identified by Perlmutter and formulated in the *Unaccusative Hypothesis* (Perlmutter, 1978), characterises a subclass of intransitive verbs that have certain syntactic and semantic properties (see section 3.2 for a more detailed account). This phenomenon has since been extensively studied and investigated for a range of different languages and seems to be valid in general across languages. This cross-linguistic validity suggests that the inherent properties of unaccusative verbs constitute part of Universal Grammar, although the specific features and the diagnostics differ between languages (Alexiadou et al., 2004). In some languages, such as Italian and French (Sorace, 1993), German (Keller & Sorace, 2003), Dutch (Perlmutter, 1978) and Danish (Allan et al., 2000), unaccusative verbs are distinguished by the use of the perfective auxiliary. Unaccusative verbs generally take the perfective auxiliary BE, whereas other verbs generally take the auxiliary HAVE. English is an exception in this regard, as in English all perfectives are formed with the auxiliary HAVE. This means that English native speakers learning a second language with auxiliary selection have to learn the correct form without the assistance of positive transfer – in other words, they cannot use their knowledge of their native language as a learning aid due to the difference between the L1 and L2 in this regard. However, under the premise that unaccusativity constitutes part of UG, it should not be necessary for second language learners to learn the correct auxiliary for each verb individually. Simply acquiring the knowledge that a particular perfective auxiliary applies to a particular class of verb should suffice. The situation is not quite as simple as this, however, as the notion of unaccusativity is not categorical and clearly delineated, and there is considerable variation between languages in their use of the perfective auxiliary. For example, Italian applies the auxiliary BE to a far wider range of verbs than French (Sorace, 1993). This difference between languages is addressed by Sorace in a series of papers (Sorace, 1993, 2000, 2004). She suggests that this variation in auxiliary selection is not random but follows specific patterns, with verbs that fall into certain categories being far more likely to select the auxiliary BE or HAVE respectively, both within and across languages. She summarises this

patterning in the Auxiliary Selection Hierarchy (Sorace, 2000), developed primarily on the basis of the behaviour of these verbs in Italian. Dividing intransitive verbs into groups, she demonstrates in a series of experiments that native speaker intuitions on the correct use of perfective auxiliaries are stronger for some verb groups than for others. She also suggests that auxiliary selection for these verbs is easier to acquire by both first and second language speakers. One of the questions investigated in this paper is therefore whether these so-called "core" unaccusative verbs are easier to learn in terms of auxiliary selection than other verbs. If the auxiliary selection hierarchy can also be applied to learnability, auxiliary selection for specific verb classes should be acquired earlier and show less variation within and among learners than for other verb classes. A second question is whether the choice of auxiliary made by second language learners reflects this hierarchy – in other words, are the language users more likely to select the auxiliary BE or HAVE respectively depending on the position of the verb in the hierarchy.

This paper is organized as follows. The next section discusses certain aspects of second language acquisition that bear a relation to the subject under consideration. Chapter 3 gives a theoretical account of the structure of the verb phrase and the phenomenon of unaccusativity. In chapter 4, Sorace's Auxiliary Selection Hierarchy is examined in greater detail and the specific considerations in relation to the Dutch language are discussed. Chapter 5 defines the research question and makes predictions on the basis of Dutch language-specific features. Chapter 6 describes an experiment conducted among English native speakers of Dutch as a second language, after which the results are presented and analysed. The paper concludes with a general discussion on the findings.

2. Second language acquisition

It is a widely researched and generally accepted fact that learning a second language is not the same as learning one's native language as a young child. A second language learner rarely or never achieves the same level of proficiency as a native speaker, and the age of acquisition plays a defining role in the ultimate attainment, which is the highest level of competence achievable by the language learner (Abrahamsson & Hyltenstam, 2008; DeKeyser, 2000; White & Genesee, 1996, among others). The general consensus is that there is a critical period, or sensitive period, during which language is easier to acquire, and that a second language acquired after this period will never be learned to the same level of competence as a native language (e.g. Lenneberg, 1967). Various explanations have been suggested for this perceived difficulty in learning a second language, such as a decrease in brain plasticity after puberty (Lenneberg, 1967), or the negative influence of the first language (Odlin, 1989). Another theory put forward by Bley-Vroman (1989) is that adult second language learners no longer have access to an innate domain-specific language acquisition device based on a universal grammar that is hard-wired into the brain, and that these adult learners consequently have to rely on general cognitive learning abilities to acquire a second language. Bley-Vroman suggests that this explains not only the general inability to attain native speaker competence but also the great variation between learners, the influence of both individual aptitude and explicit instruction (neither of which play an influential role in first language acquisition), and the perceived stagnation in development generally referred to as fossilisation. However, this is disputed by other researchers who believe that Universal Grammar is fully or partially available to adult second language learners (e.g. Epstein et al., 1996; Coppieters, 1987; White and Genesee, 1996). For example, Coppieters (1987) suggests that some syntactic language features and grammatical constraints considered to be associated with Universal Grammar enjoy protected status (i.e. he found a smaller discrepancy between native and near-native speakers with respect to these formal features than with respect to language-specific features such as the use of tense). In a study by White and Genesee (1996), they conclude that the native-like performance of some second language learners regarding the recognition of ungrammatical

sentences that violate principles of Universal Grammar demonstrates continued access to UG.

A further indication that second language learners do in fact have access of some kind to an innate language learning device can be derived from Slavkov (2009) in his dissertation on the elicited production of English two-clause questions by Canadian-French and Bulgarian native speakers. His test subjects produced a range of structures that are ungrammatical in English as well as in their respective native languages, although such structures are grammatical in various other languages with which the speakers were unfamiliar. Furthermore, the structures produced by the two groups of test subjects were very similar to one another, despite the fact that question structure differs in French and Bulgarian with respect to the syntactic movement of the question word. The structures produced by Slavkov's subjects were also similar to two-clause questions produced by English children in a study by Thornton (1990) and Dutch children in a study by Van Kampen (1997). The suggestion by Van Kampen is that the children are using default structures that are less complex and therefore easier to process, in keeping with their limited cognitive capacity. The fact that adult second language learners fall back on similar default structures when unsure of the correct form seems to point to the application of an innate language acquisition tool.

The assumption that Universal Grammar is available in some form to adult second language learners comes with an implication that languages features considered to be governed by Universal Grammar may be easier for second language learners to acquire and use than non-universal, language-specific features. As mentioned in the first chapter, the syntactic phenomenon of unaccusativity investigated in this paper is found cross-linguistically and the inherent properties of unaccusative verbs may therefore be considered to constitute part of Universal Grammar (Alexiadou et al., 2004). The following chapter discusses this phenomenon in greater detail. The chapter begins with a general description of the verb phrase within the framework of Chomsky's Government and Binding theory of grammar (Chomsky, 1981), to provide a theoretical foundation for the notion of unaccusativity. The second part of the chapter takes a closer look at unaccusativity and its analysis and diagnostics in different languages.

3. Verb properties

3.1. The verb phrase

The Government and Binding theory of grammar (Chomsky, 1981) has been highly influential and widely accepted among linguists for many decades. It provides insights into many characteristics of natural language, it is supported by empirical data in terms of what is and what is not considered grammatical by native speakers of a particular language, and it can be seen to apply to a certain extent across languages. Although a detailed examination of this theory goes far beyond the scope of this paper, some of the basic principles of the theory have a bearing on the language aspects being investigated here and will therefore be described briefly in this section.

The terms “government” and “binding” relate to the hierarchies and dependencies between elements within a sentence. Government and Binding (GB) distinguishes four levels of grammatical representation: deep structure (DS), surface structure (SS), logical form (LF) and phonetic form (PF) (Chomsky, 1981). Deep structure is the starting point of the syntactic derivation before any syntactic movement has taken place, and is considered to be the base structure, or underlying representation. Transformational rules are then applied to the base structure to move words to different positions in a sentence, producing the surface structure, this being what is actually pronounced (Carnie, 2007). Syntactic movement is considered to have taken place if a word or phrase is interpreted in a different place than its linear position in the sentence would suggest. Perhaps the most salient example of syntactic movement (in English) is the movement of questions words and phrases. Comparing the sentences below, in 1a and 1b the object phrases (*Mary* and *an elephant* respectively) come after the verb, which is the canonical position of objects in English. In sentences 1c and 1d, the object of the sentence is now a question word, but it has moved to the front of the sentence. If the question takes the form of a phrase, as in 1e and 1f, the entire phrase moves to the initial sentence position. The object phrases in the sentences below are shown in bold type. (The insertion of the dummy auxiliary *do* is a specific feature of the English language, required in questions and negatives, but it bears no further relevance to the phenomenon of *wh*-movement discussed here.)

1. a) Bill saw **Mary**.

   b) Bill saw **an elephant**.

   c) **Who** did Bill see?

   d) **What** did Bill see?

   e) **Which zoo animal** did Bill see?

   f) **How many elephants** did Bill see?

Under GB theory, movement is triggered by the need to check certain features of a word or phrase for compatibility with another part of the sentence with which it bears a close relation. Syntactic movement is language-specific. For example, Chinese does not have overt *wh*-movement[1]. In the Chinese language *wh*-expressions remain in the canonical position, which for objects is after the verb, as shown in sentence 2 (taken from Hornstein et al, 2006).

2. Bill mai-le shenme?

   Bill buy-ASP what?

   "What did Bill buy?"

The remaining two levels of grammatical representation, PF and LF, correspond respectively to the sound and meaning components of language, representing the interface with the phonological system on the one hand and with the semantic or interpretive system on the other hand. The aforementioned four levels of grammatical representation and their relationship to one another are depicted in the model in Figure 1 (taken from Burzio, 1986).

[1] Languages such as Chinese with no visible, or overt, *wh*-movement are widely considered to have covert movement (Cheng, 2009). This refers to movement for interpretation (scope) purposes that takes place between the SS and LF levels and is therefore not reflected in the spoken language. The motivation and evidence for this proposal are highly complex and go beyond the scope of this paper.

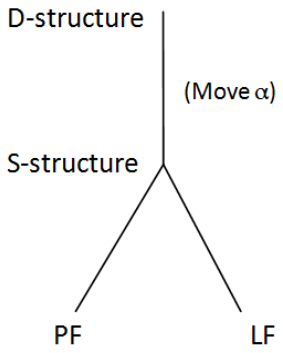

Fig. 1 The GB model of grammar

At the level of deep structure, the lexical items (i.e. the words) are combined with the phrase structure rules to form the input for the derivation (Davies & Dubinsky, 2004). The phrase structure rules within GB are based on X-bar theory (Chomsky, 1970; Jackendoff, 1977). X-bar theory aims to capture the hierarchy and recursivity of natural language, allowing phrases to be linked together with increasingly deep embedding and, in principle, indefinitely. X is a variable that can stand for any lexical category such as a noun, verb or preposition, or for a functional category such as tense. A simple phrase according to X-bar theory consists of a head (the lexical or functional element on which the phrase is based) plus, optionally, a specifier and a complement. The specifier and complement are also phrases, which are linked to the X-phrase in a specific way by virtue of their position. This is depicted in the tree structure in Figure 2.

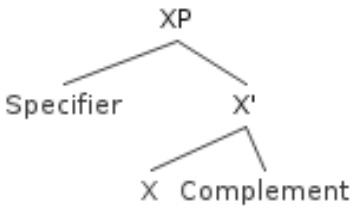

Fig. 2 Tree structure for simple phrase

The specifier position is the position in which the sentence subject is found. This position may be vacant, as not all phrases have a subject. The complement position may be occupied by an object or by a phrase or clause, as illustrated by the expressions in bold type in the sentences below. In sentence 3a, the noun phrase *fish* occupies the complement position of

the verb *like*, in 3b this position is occupied by the infinitival tense phrase *to play tennis*, and in 3c the complementiser phrase *that Bill is angry* is the complement of the verb *know*.

3. a) John likes **fish**.
   b) John likes **to play tennis**.
   c) John knows **that Bill is angry**.

Different lexical items have different requirements with respect to how these can be combined with other items on the basis of the phrase structure rules. These requirements are generally referred to as subcategorisation properties (Davies & Dubinsky, 2004). In the case of the verb, all verbs must have one or more arguments. An argument is often a noun phrase (NP). For example, the subject and object in sentence 3a above (*John* and *fish* respectively) are arguments of the verb *like*. Depending on the type of verb and the structure of the sentence, an argument may also be a tense phrase (TP) or complementiser phrase (CP) (as in the complements of sentences 3b and 3c) or a prepositional phrase (PP). The arguments bear some kind of relation to the activity or event described by the verb and are necessary in order for the meaning of the verb to be properly expressed. Depending on the semantics of the verb, arguments have specific semantic roles, commonly referred to as thematic or theta roles. Examples of these roles, among others, are the agent (initiates or performs an action), the experiencer (feels or perceives an event), the theme (undergoes an action or is moved or perceived) and the patient (is affected in some way by the action) (Marantz,1984). According to the Theta Criterion (Chomsky, 1981), all theta roles of a verb must be assigned, with each theta role being assigned to only one argument and each argument being assigned only one theta role. As a generalisation for the English language, the agent roles are the logical subjects in a sentence, whereas the theme/patient roles (these roles share common ground and may overlap depending on the semantics of the verb) are the logical objects in a sentence (Marantz,1984). However, not all sentences follow a pattern in which the subject has the agent role and the object has the theme/patient role. Subjects and objects are syntactic rather than semantic concepts, being distinguished by their case and not by their semantic role. In nominative-accusative languages such as the Germanic and Romance languages, the subject receives nominative case and the object receives accusative case. Case is a syntactic phenomenon that signifies the grammatical

function of an NP and its relation to a verb or preposition. In some languages, such as German, case is marked morphologically, but in a morphologically poor language like English only pronouns display case overtly. However, even if case is not morphologically marked, NPs are still considered to have been assigned abstract case by virtue of their position within the sentence structure. In summary, case can be seen as a syntactic property relating to the grammatical function of an NP in the sentence, whereas theta roles are semantic concepts relating to the semantic properties of the verb.

As mentioned above, theta roles do not always coincide with their logical syntactic equivalent. Whereas at the DS level a theme/patient is generated in the complement (object) position, a transformation such as passivisation turns the theme or patient into the sentence subject. For example, in the sentences 4a and 4b below (from Carnie, 2007) the thematic roles are the same but the syntactic positions are not.

4. a) The policeman kissed the puppy.
   b) The puppy was kissed by the policeman.

In both sentences *the policeman* is the agent and *the puppy* is the theme. *The puppy* is the object in sentence 4a, but in sentence 4b *the puppy* is the subject. *The policeman* is the subject in 4a, but in 4b this NP is contained within a prepositional phrase and is no longer an argument of the verb (the prepositional phrase is an adjunct rather than a complement as it is a non-compulsory addition – in other words, it can be omitted without the sentence becoming ungrammatical). Although the topics of the two sentences are different, the truth conditions are the same (∃x [x = policeman], ∃y [y = puppy], kissed [xy]). It is therefore widely believed within the framework of transformational grammar that the deep structure of the two sentences is the same. Burzio (1986), for example, states that at the DS level there is a one-to-one correspondence between thematic relations and grammatical function. This perceived underlying correspondence between semantic role and syntactic function is integral to the notion of unaccusativity as described in the next section.

Within GB the subject is generally referred to as the external argument and the object is referred to as the internal argument. The terms “external” and “internal” relate to the fact

that the subject is assumed to be generated outside the verb phrase (VP), in the specifier position of the tense phrase (TP), whereas the object is assumed to be generated inside the VP, in the complement position. The TP (sometimes called IP, or inflectional phrase) is an additional phrasal level to accommodate the tense inflection, auxiliary or modal verb. This is depicted in Figures 3 and 4 below.

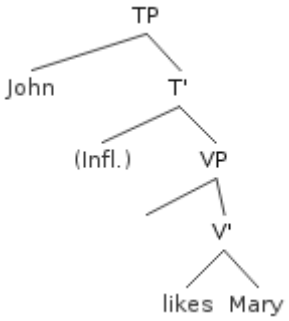

Fig. 3 TP and VP with tense inflection

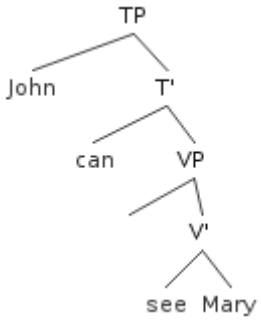

Fig. 4 TP and VP with modal verb

It should be pointed out here that syntactic theories have evolved since the Government and Binding model. For example, according to the principles of the widely accepted VP internal subject hypothesis (e.g. Kuroda, 1988), the subject is now assumed to originate inside the VP and then move to the TP. This means that in Figures 3 and 4 above, the specifier position of the VP would contain a trace of the subject (*John*). The Minimalist Program (Chomsky, 1995) led to further developments in syntactic theory. One of the components to be re-evaluated was the structure of the VP and the assignment of the internal and external arguments. Under the new ideas, it was suggested that the VP has a more complex structure, comprising an inner and an outer "shell". Within the Minimalist Program, the external argument is considered to be introduced by the outer VP shell rather than by a functional tense category. The finer details go beyond the scope of this paper, but the revised conceptualisation does not fundamentally affect the theories relevant to unaccusativity discussed in the following section. Furthermore, much of the literature relating to unaccusativity is based on the GB model. In the remainder of this paper we shall therefore continue to use the terms and verb phrase structure as set out in this section.

One of the principles of the GB model is that all NPs must have case, which can only be assigned by a verb or by a preposition (Chomsky, 1981). Nominative case is assigned to the

subject within the TP by the finite tense of the tensed verb, with which it has to agree in number and person. Accusative case is assigned to the object within the VP by the lexical verb. If an NP is generated in DS in a position where no case can be assigned it must move to a case position. According to the highly influential and widely quoted work by Burzio (1986), passive verb forms and some intransitive verbs (the unaccusatives, discussed in the following section) do not have an agentive, or active, semantic role and cannot therefore assign an external theta role to the subject of the verb. He further suggests that only transitive verbs can assign accusative case. This led to the formulation of what is known as Burzio's Generalisation, which states that if a verb does not assign an external theta role, it cannot assign accusative case to its internal argument. However, as all NPs must have case, an NP generated in the complement position of a passive verb form must move to a case position. Furthermore, according to Chomsky's Extended Projection Principle, all clauses must have subjects (Chomsky, 1981). Taken together, these conditions provide motivation for the syntactic movement of NPs in passive constructions. This is depicted in Figure 5. The object *Mary* is generated in the complement position and then moves to the subject position (specifier of TP) in order to satisfy the case and sentence subject requirements, leaving a trace in the underlying complement position.

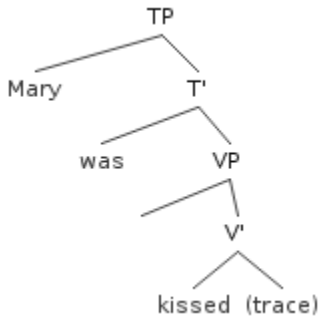

Fig. 5 Passive construction with movement to subject position

Within the GB framework the subcategorisation requirements of the verbs are considered to be an integral part of its lexical properties. In other words, these properties are stored in the mental lexicon together with the semantic (meaning) and phonetic (sound) information. According to Chomsky's Projection Principle (Chomsky, 1981) these subcategorisation properties are preserved at all syntactic levels. If a verb is stored in the lexicon with the requirement that it must have two arguments (e.g. a subject and an object), the verb must

have a subject and an object at all levels of representation. However, this lexicalist, or projectionist, view has been challenged by some studies that analyse the interaction between the mental lexicon and the syntactic structure of sentences (e.g. Folli & Harley, 2005; Sorace, 2004). One persuasive objection to the idea that all information pertaining to argument structure is already present at the lexical level comes from the existence of verbs that display variable behaviour. For example, the sentences 5a-5e below (taken from Folli & Harley, 2005) show one verb used in five different sentence structures. 5a is intransitive, 5b is transitive, and 5c could be said to be ditransitive (the addition of the PP *off the table* is necessary in order for the sentence to make sense, as otherwise the crumbs themselves would be being cleaned, which is clearly not what is meant). 5d is a resultative construction (indicating the result of the action described by the verb), and 5e shows an idiomatic use of *clean* together with the particle *out*.

5. a) Mary cleaned.
   b) Mary cleaned the table.
   c) Mary cleaned the crumbs off the table.
   d) Mary cleaned the table spotless.
   e) Mary cleaned out her savings.

Sentences 6a-6d (also from Folli & Harley, 2005) show another kind of variable verb behaviour. The verbs *open* and *break* can be used either transitively, as in 6b and 6d, or intransitively, as in 6a and 6c, the so-called transitivity alternation. In the case of the transitive verbs, the causer of the action or event is specified, this being the external argument of the verb, whereas the internal argument is the theme/patient. The intransitive verbs are change-of-state verbs that have just one argument, which is the theme/patient, now in the subject position.

6. a) The door opened.
   b) John/The wind opened the door.
   c) The glass broke.
   d) Mary/The stick broke the glass.

Based on the lexicalist approach, these variable verbs would require multiple entries in the mental lexicon, one for each syntactic structure. This seems to be inefficient. After all, although the word *clean* in sentences 5a-5e has a slightly different meaning in each sentence depending on the complement of the verb, the meanings are related (dirt of some kind or something in an idiomatic sense is being removed from something else). Multiple entries are therefore unnecessary as the meaning can be derived from the syntax. The same can be said for the verbs in 6a-6d. It is therefore reasonable to assume that these are not completely different lexical entities, otherwise the similarity of form would be highly coincidental.

The opposing view to the lexical view is the constructionalist approach, which argues that the lexical entry contains bare lexical information and does not incorporate constraints such as the number of arguments and the permissible syntactic structures into which a verb can be inserted (Sorace, 2004). The different meanings are derived from the syntactic structure in which a word is found. In other words, the syntax determines the way in which a variable verb such as *clean* in sentences 5a-5e is to be interpreted (Borer, 2004). However, the problem with the constructionalist approach, as pointed out by Folli & Harley (2005), is that with no lexical specification at all to restrict the syntactic behaviour of verbs, all verbs could be expected to appear in all positions. Such flexibility is not witnessed, however. The transitivity alternation of verbs such as those in 6a-6d, for example, does not apply universally, and some verbs are more flexible than others (Sorace, 2004).

Various theories have been put forward to explain the variable flexibility of verb behaviour. For example, Levin & Rappaport Hovav (1995) distinguish four so-called linking rules, which determine the argument structure of the verb on the basis of its semantic properties. For example, the Immediate Cause linking rule states that the external argument of a verb is the immediate cause of the event described by the verb, and the Directed Change linking rule states that the direct internal argument is the entity that undergoes change. This approach shares some common ground with the projectionalist approach, as the lexical entry contains certain information on argument structure, but the application of different linking rules to different verbs drastically reduces the possible structures in which a verb can appear, while still allowing for some variation. Arad (1996) similarly proposes that a combination of syntactic and lexical properties is responsible for the projection of the verb. According to

Arad, the lexical information contained in the lexical entry constrains rather than determines the syntactic structures in which the verb can appear. Some verbs have flexible meanings and are therefore compatible with more than one syntactic structure – like the sentences in 5 and 6. Another approach that suggests an interaction between the syntax and semantics is described by Sorace (2004, 2006) in her examination of split intransitivity. She suggests that some verbs are more rigid in their syntactic behaviour and are unaffected by context and aspect (grammatical or lexical aspect refers to the perspective on a state or action in relation to its temporal structure, such as its inception, duration or completion), whereas others show variable syntactic behaviour depending on the meaning of the predicate. The former seem to fit the projectionist approach, whereas the latter pattern according to the constructionist approach (Sorace, 2006). With respect to unaccusativity, described in more detail in the following section, she suggests that both a syntactic and a lexical-semantic characterisation are needed to explain the inflexible distributional properties of some verbs on the one hand and the variation of some verbs on the other hand (Sorace, 2004). This would also go some way to explaining the differences between languages highlighted in the following section.

3.2. Unaccusativity

There is a widely accepted view that intransitive verbs can be divided into two types, each with different syntactic and semantic properties. This is often referred to as split intransitivity. Perlmutter (1978) was one of the first to address this distinction in his observation of Dutch verbs. He noted that some intransitive verbs in Dutch allow the so-called impersonal passive construction, whereas others do not. A simple example of the Dutch impersonal passive is given in the sentences in 7a and 7b.

7.	a)	Er	werd	gedanst.
		There	was	danced.
	b)	* Er	werd	gearriveerd.
		There	was	arrived.

This construction is grammatical with the verb *dansen* (dance) in sentence 7a, which can be roughly paraphrased as “dancing took place”. The sentence in 7b using the verb *arriveren*

(arrive) is ungrammatical. (The insertion of the expletive *er* in the subject position in 7a is required to satisfy the principle that all clauses must have subjects: Chomsky, 1981.) In Perlmutter's detailed analysis of the distribution of the impersonal passive in Dutch, he concluded that there are two different types of intransitive verbs, one of which allows the impersonal passive and one of which does not. He called these two verb types unergative and unaccusative respectively, and his analysis led to the formulation of the highly influential and extensively quoted Unaccusative Hypothesis. Perlmutter observed that the subjects of unaccusative intransitive verbs have certain semantic properties in common with the direct objects of transitive verbs (which take accusative case in many languages), such as a lack of agentivity and a high level of affectedness by the action of the verb. In other words, the subjects of these intransitive verbs have a theme or patient theta role instead of the agent theta role that is more typically associated with the verb subject. On the other hand, the subjects of unergative intransitive verbs pattern with the subjects of transitive verbs, as the subjects of these verbs have an agent theta role, playing an active and usually volitional role in the action or event denoted by the verb.

This distinction is illustrated in the sentences below. In sentence 8a, the subject John has an agent theta role, as he is performing the action of dancing, presumably voluntarily. In 8b, John has a theme theta role, as he is undergoing the act of falling. Although it is conceivable that he could fall deliberately, perhaps for dramatic or comic effect, the more common interpretation is that this action is not volitional (Zaenen, 1988).

8. a) John danced.
   b) John fell.

On the basis of these thematic considerations, Perlmutter concluded that the argument of an unaccusative verb is actually an underlying direct object, or internal argument, whereas the argument of an unergative verb is an underlying subject, or external argument. In the words of Levin & Rappaport Hovav (1995: p. 3), this can be summed up as follows:

> " (….) in argument structure terms, an unergative verb has an external argument but no direct internal argument, whereas an unaccusative verb has a direct internal argument but no external argument."

Further work carried out by Burzio (1986) based on observations of these verbs in Italian (described in more detail below) developed this concept further, leading to the formulation of Burzio's Generalisation. As mentioned in the previous section, Burzio's Generalisation states that if a verb does not assign a theta role to its external argument, it cannot assign accusative case to its internal argument (hence the term unaccusative for these verb types). Burzio's Generalisation applies to unaccusative verbs as well as passive verb forms, both of which lack an external argument. Unaccusative verbs are considered to be inherently passive, as they have no external agentive theta role. As with the passive verbs described in section 3.1, the single argument of the unaccusative verb is an internal argument that is a semantic patient or theme, and it is base-generated in the object position. In other words, at the beginning of the derivation, on the deep structure level, this argument occupies the complement (object) position of the VP. It then moves to the subject position to satisfy the case and sentence subject requirement (the Extended Projection Principle, which states that all clauses must have a subject). Burzio's Generalisation relating to passives and unaccusatives is in line with Perlmutter's analysis of the impersonal passive construction. Perlmutter suggests that the subjects of unaccusative verbs have been promoted to the subject position from the underlying object position. Consequently these verbs cannot be passivised, as in passive constructions the underlying object is also raised to the subject position, and according to Perlmutter this promotion of the underlying object can only take place once. This is illustrated by the (simplified) tree structures below.

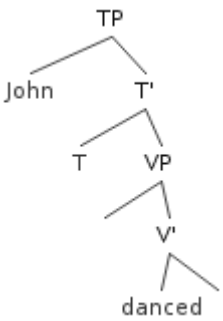

Fig. 6 Unergative structure

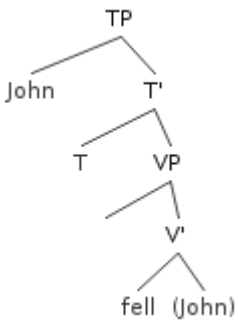

Fig. 7 Unaccusative structure

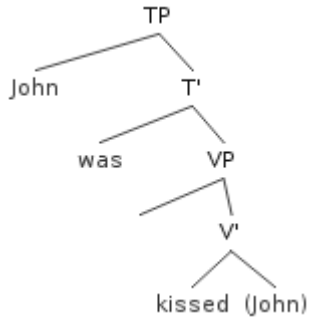

Fig. 8 Passive structure

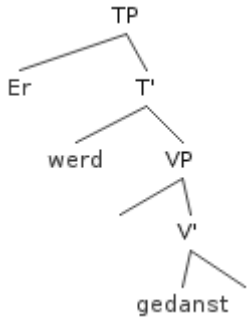

Fig. 9 Dutch impersonal passive

Figure 6 shows a simplified structure with the unergative verb *dance*. Being an intransitive verb, there is no object and the agentive subject occupies the canonical subject position. Figure 7 gives the structure of the unaccusative verb *fall*. According to the general consensus on the structure of unaccusative verbs, the subject is generated in the object position and moves to the subject position, leaving behind a trace (indicated by the bracketed *John*). Figure 8 shows a passive verb construction. The similarity with the unaccusative structure is evident. Here, too, a trace is left in the underlying object position. Figure 9 gives an impression of the Dutch impersonal passive *er werd gedanst* ("there was danced"). Due to the passivisation there is no agentive subject, but being an intransitive verb there is no object either. The dummy subject *er* is inserted to fill the subject position.

Since the analyses of Perlmutter and Burzio, unaccusativity has been studied in many languages and the Unaccusative Hypothesis is generally considered to apply across languages. A number of theories and syntactic diagnostics have been put forward and there is considerable empirical evidence for a distinction between the two verb types (e.g. Zaenen, 1988, 2006 for Dutch; Burzio, 1986 for Italian; Sorace, 1993, 2000 for Italian). For Dutch, Zaenen (1988) presents a detailed account of unaccusativity diagnostics. In addition to the above-mentioned impersonal passives, she gives examples of other factors that may distinguish between unergative and unaccusative verbs. Some of these are presented in sentences 9a-9h, with the literal English translations in brackets (sentences 9a-9d and 9g-9h are from Zaenen, 1988).

9. a) Jan heeft getelefoneerd. (John has phoned)
   b) Jan is gearriveerd. (John is arrived)
   c) * De gewerkte man. (The worked man)
   d) Het gevallen blad. (The fallen leaf)
   e) De geplante boom. (The planted tree)
   f) * De geplante man. (The planted man)
   g) De werker. (The worker)
   h) * De valler. (The faller)
   i) De lezer. (The reader)

The first diagnostic shown here in sentences 9a and 9b is auxiliary selection. The unergative verb *telefoneren* (phone) takes the perfective auxiliary *hebben* (have), whereas the unaccusative verb *arriveren* (arrive) takes the auxiliary *zijn* (be). Sentences 9c and 9d show a distinction in the use of the prenominal past participle. This is acceptable with the unaccusative verb *vallen* (fall) but not with the unergative verb *werken* (work). The parallel examples in sentences 9e and 9f show how the arguments of the transitive verb *planten* (plant) behave. The internal argument *boom* (tree) can be used with a prenominal past participle, whereas the external argument *man* (understood to mean the man who is planting the tree) cannot. In other words, the single argument of the unaccusative verb patterns with the internal argument of the transitive verb, whereas the single argument of the unergative verb patterns with the external argument of the transitive verb. Finally 9g and 9h present nominalisations using the suffix *–er*, which is possible with unergative *werken* (work) but not with unaccusative *vallen* (fall). Here, too, a parallel can be seen with the arguments of a transitive verb. The nominalisation of *lezen* (read) in 9i is generally understood to mean someone who reads (external argument), not something that is read (internal argument).

Zaenen points out that these tests do not correlate perfectly with one another, as she presents many examples of verbs that are predicted to be unaccusative by one test and unergative by another test. One example she gives (originally from Perlmutter, 1978) concerns the verb *duren* (last), as shown below in 10. Although this verb takes the auxiliary

hebben (have), suggesting it is unergative, it cannot be used in an impersonal passive, suggesting it is unaccusative.

10. a) Het concert heeft een hele tijd geduurd.
    The concert has lasted a long time.
    b) * Er werd (door het concert) een hele tijd geduurd.
    There was lasted a long time (by the concert).

Zaenen suggests that the distinction is more complex than a two-way unergative-unaccusative split. She argues for a split along two dimensions, with a division between processes that can be controlled (implying a degree of intentionality) along one dimension, and an aspectual distinction between telic and atelic verbs along the other dimension. Telic verbs have an inherent end point, whereas atelic verbs refer to activities with a certain duration and no implicit end point (Velupillai, 2012). This takes the discussion beyond syntactic features, as telicity and intentionality are semantic concepts. Zaenen therefore concludes that semantic factors and sentence context cannot be ignored in discussions on split intransitivity.

In the case of Italian, Burzio's analysis of Italian verbs also uncovers distinctions between unaccusative[2] and unergative verbs. One of his observations concerns the use of the perfective auxiliary. As in Dutch, unaccusative verbs take the auxiliary *essere* (BE) whereas unergative verbs take the auxiliary *habere* (HAVE). Another important distinction made by Burzio relates to the phenomenon of so-called *ne*-cliticisation. *Ne* is a clitic which means *of them,* and it behaves differently according to the verb type. Examples of these different behaviours, taken from Burzio (1986), are given below.

11. a) Ne arrivano molti
    of-them arrive many
    "Many of them arrive"

[2] Burzio uses the terms *ergative* and *unergative* for these two verb types. However, the underlying reasoning is the same, so for the sake of consistency we will continue to use the term *unaccusative* rather than *ergative* here.

b) * Ne telefonano molti
of-them telephone many
"Many of them telephone"

c) Giovanni ne invitera molti
Giovanni of-them will-invite many
"Giovanni will invite many of them"

d) Ne saranno invitati molti
of-them will-be invited many
"Many of them will be invited"

Sentence 11a with the unaccusative verb *arrive* is grammatical, but sentence 11b with the unergative verb *telephone* is ungrammatical. Burzio then goes on to give many examples of other sentences demonstrating that the *ne* clitic can only be used with a direct object, as in sentence 11c. However, it can also be used in a passive construction as in sentence 11d. This seems to provide compelling evidence that the subjects of both passives and unaccusatives are underlying direct objects, as argued by Perlmutter.

A range of other syntactic diagnostics that can be applied to different languages can be found in Zaenen (2006). Besides those diagnostics discussed above, she mentions reflexive constructions in Italian and French, prefixation and negation phenomena in Russian, and the aforementioned auxiliary selection for several Germanic and Romance languages. English is an exception in this regard, as there is no choice of perfective auxiliary in the English language. However, this does not mean that unaccusativity cannot be tested in English. Levin & Rappaport Hovav (1995) discuss distinctions between unaccusatives and unergatives in English. They propose that one difference between these verb types concerns the possibility of appearing in a resultative construction. A resultative phrase denotes the end state of a noun as a result of the action described in the verb, and can only be used to describe a direct object (Levin & Rappaport Hovav, 1995). It cannot therefore be used with intransitive verbs as these do not have direct objects. However, it is acceptable with passives and unaccusatives. This is demonstrated below in sentences 12a-12e (all taken from Levin & Rappaport Hovav, 1995, with the exception of 12a, from Folli & Harley, 2005).

12. a) Mary cleaned the table spotless.

b) * Dora shouted hoarse.

c) Dora shouted herself hoarse.

d) She was shaken awake by the earthquake.

e) The river froze solid.

12a is a typical resultative construction with a transitive verb. The table becomes spotless due to the verb action (cleaning). In sentence 12b the unergative verb *shout* is ungrammatical in such a construction, although in 12c this sentence has been repaired with the addition of a so-called fake reflexive object (Levin & Rappaport Hovav, 1995). Sentence 12d shows a passive construction, and 12e has an unaccusative verb. Both sentences are acceptable.

Locative inversion, or *there* inversion, is also discussed as a possible diagnostic for unaccusativity in English (Carnie, 2007; Levin & Rappaport Hovav, 1995). This phenomenon in English refers to the inversion of subject and verb after a locative phrase or after the pleonastic pronoun *there*. According to Carnie (2007), this is possible with unaccusative verbs but not with unergative verbs. However, this is disputed as a diagnostic by Levin & Rappaport Hovav due to the large number of exceptions. Sentences 13a and 13b (from Carnie, 2007) show *there* inversion, in a grammatical sentence with an unaccusative verb in 13a, and in an ungrammatical sentence with an unergative verb in 13b. The sentences in 13c and 13d show similar examples with locative inversion. 13c (unaccusative) is grammatical, whereas 13d (unergative) is ungrammatical. However, if there is no subject-verb inversion, as in 13e, this sentence with the verb *shout* is acceptable.

13. a) There arrived three men at the palace.

b) * There danced three men at the palace.

c) In the room stood three man.

d) * In the room shouted three men.

e) In the room three men shouted.

Carnie (2007) also suggests another way in which unaccusatives and unergatives differ in English. Unergatives can take an optional direct object, whereas generally speaking unaccusatives cannot, as shown in the sentences below (from Carnie, 2007).[3]

14. a) Stacy danced (a jig).
    b) * Stacy arrived a letter.

If unaccusatives are analysed as having an underlying direct object, as has been argued in this chapter, this difference is not surprising. Sentence 14b cannot have a direct object as this position is occupied by *Stacy* in the underlying representation.

The diagnostics for Dutch presented above in sentences 9c-9i could also conceivably be applied to English. The use of the prenominal past participle in 15a and 15b follows the same pattern as in Dutch – grammatical for unaccusatives but ungrammatical for unergatives. The same could be said for the *er*-nominalisations in 15c and 15d. 15c is acceptable, whereas 15d is certainly questionable (in the context of referring to a person or thing that falls), if not entirely ungrammatical.

15. a) * The worked man
    b) The fallen leaf
    c) The worker
    d) ?? The faller

McCloskey (1993) puts forward another diagnostic for unaccusativity that applies uniquely to English. This is based on the vulgar expression *sod all*, which means "nothing at all". McCloskey gives examples in which this expression appears in the subject position of passives and unaccusatives and in the object position of transitives, but is ungrammatical in the subject position of transitives and unergatives.

[3] A notable exception here is the expression "die a horrible death", which allows a direct object even though *die* is an unaccusative verb according to most diagnostics.

In summary, although auxiliary selection does not exist as an unaccusativity diagnostic in English, other diagnostics can be applied to English, such as the resultative construction, locative inversion, optional direct object, prenominal past participle, *er*-nominalisation and language-specific idiosyncratic expressions as described above.

Besides the syntactic properties set out above, there does also seem to be some empirical evidence for the unaccusative-unergative split outside the field of syntax. For example, Legate (2003) observes a prosodic effect in relation to unaccusative constructions. In English, according to the Nuclear Stress Rule primary stress is assigned to the final stress-bearing element in sentences with a neutral context (Legate, 2003). In a canonical English subject-verb-object sentence this stress would consequently fall on the object. However, Legate notes that in the case of passive and unaccusative VPs the stress falls on the pre-verbal element - the subject of the verb. This is illustrated in sentences 16a-16e below (sentences 16a-16c are from Legate, 2003). The stress-bearing words are indicated in bold type.

16. a) (What happened yesterday?) My **bike** was stolen.
    b) # **John** stole my bike.
    c) (What happened this morning?) The **train** arrived.
    d) The people **danced.**
    e) The children **played.**

Sentence 16a is a passive sentence in a neutral context in which the primary stress falls on the passive subject. In the corresponding active sentence in 16b, primary stress on the subject *John* is only possible in a marked context, in which some kind of contrast or correction is being made (e.g. not Bill but **John** stole my bike). Sentence 16c has an unaccusative verb, and here, too, primary stress falls on the subject *train*. In sentences 16d and 16e, which have unergative verbs, the primary stress falls on the final element of the sentence, in this case the verb. Legate does not make this distinction between unergatives and unaccusatives herself in this paper, as the objective of her paper is to find evidence for the phase (a subsection of a syntactic derivation). However, this effect could be interpreted as a further indication that the subject of the unaccusative verb has moved from the underlying object position. In any case, the patterning is notable as it provides another

example in which the subjects of unaccusative verbs behave in the same way as the direct objects of transitive verbs (and the subjects of passive verbs).

Evidence has also been found for the syntactic encoding of unaccusativity in the field of psycholinguistics. Friedmann et al. (2008) used a cross modal lexical priming technique to show that the subjects of unaccusatives are generated in the underlying object position. Lexical priming is a technique used to detect the activation of a word in the mental lexicon. A lexical prime is a word that is presented to a test subject shortly before the presentation of a target word. The test subject then has to quickly make a decision about whether the target word is a real word or a non-word. If the prime and target words are semantically related, this speeds up lexical access and with this the lexical decision process. By manipulating the position of the prime and analysing the results, conclusions on lexical activation can be drawn. Friedmann et al. (2008) showed that the subjects of unaccusative verbs are activated twice: once in the position in which they actually appear and again in the post-verbal position from which according to the Unaccusative Hypothesis the underlying object has moved. In the case of unergative verbs, no post-verbal reactivation was found.

In summary, there is considerable empirical evidence for the phenomenon of unaccusativity with reference to a class of verbs that have certain characteristics and whose subjects seem to be object-like with regard to both syntactic behaviour and semantic qualities. The question of whether unaccusativity is a semantic or a syntactic phenomenon is one that has been discussed extensively in the literature. It has been argued that the difference between the two verb types is semantic rather than syntactic (e.g. Van Valin, 1990), but the general consensus seems to be that the situation is actually less clear cut and that the distinction between unaccusative and unergative verbs is a complex interaction between syntactic properties and semantic aspects. Perlmutter (1978) concluded on the basis of the Dutch data that the unaccusativity-unergativity split seemed to depend on the semantic relation between the noun phrase and the predicate. According to Levin and Rappaport-Hovav (1995), split intransitivity is semantically determined and syntactically encoded. In other words, the semantic properties of the verb determine the category into which it falls, and this category is then reflected in the syntax of the sentence. Belien (2012), in her study of auxiliary selection for Dutch motion verbs, suggests that rather than a rigid unaccusative-

unergative split the distinction depends on compatibility between the lexical semantics of the verb and the sentence context. Another study relating to auxiliary selection in Dutch by Lieber & Baayen (1997) makes a similar suggestion. Lieber & Baayen propose that auxiliary choice is based on a feature of semantic structure which they call the "inferable eventual position or state". In other words, verbs that select BE provide information of some kind on the final position or final state of the argument. For example, in the case of a verb such as *come*, the end position would be a position closer to an implied reference point (e.g. *home* in the predicate *come home*), whereas the final state of the argument of a verb such as *grow* would be a size larger than the initial size (Lieber & Baayen, 1997). Hoekstra (1999), on the other hand, disputes this analysis, claiming that a syntactic analysis of unaccusativity is preferable as this relates auxiliary selection to other syntactic properties.

As already discussed in this section, auxiliary selection is widely considered to be an important diagnostic for unaccusativity. However, this viewpoint is problematic due to the great variation between languages. For example, Italian applies the auxiliary BE to a wider range of verbs than French, German or Dutch. (Sorace, 1993, 2000). Languages such as English and Spanish have uniform auxiliary use, selecting only HAVE, and Welsh is one example of a language that has a completely different perfective construction, using an aspect marker rather than an auxiliary to express the perfect aspect (Borsley et al., 2007). Sorace's view is that auxiliary selection variation is orderly and predictable, and this is underpinned by a number of studies investigating this phenomenon in Italian, French and German (Sorace, 1993; Keller & Sorace, 2003). This has resulted in the formulation of the Auxiliary Selection Hierarchy (ASH) (Sorace, 2000). The ASH subdivides unaccusative and unergative verb types into different classes that follow a hierarchy, ranging from "core" (categorical) unaccusative to "core" (categorical) unergative. The core properties of unaccusatives are telic and dynamic change affecting a theme argument, whereas the core properties of unergatives are agentive non-motional activity performed by a volitional participant (Sorace, 2000; Sorace & Keller, 2005). The ASH is presented by Sorace as an explanation for the variation between languages. The core verb classes are more consistent and robust in their auxiliary use, both within and between languages, than the peripheral verb classes. The use of the correct auxiliary is also easier to learn for these core verb

classes, and this learnability is the focus of the experiment described in chapter 6. The ASH and the different verb classes are described in more detail in the next chapter.

4. The auxiliary selection hierarchy

4.1. Definition

One of the most salient diagnostics for the discrimination of unaccusatives and unergatives is the selection of the perfective auxiliary in languages that have a choice between two auxiliaries, with unaccusative verbs selecting BE and unergative verbs selecting HAVE. Although striking similarities can be found in many Romance and Germanic languages, there is also considerable variation across languages. As mentioned in the previous section, Italian applies the auxiliary BE to a wider range of verbs than French, German or Dutch (Sorace, 1993, 2000), whereas modern English and modern Spanish do not use the auxiliary BE at all (Legendre, 2007). Some verb classes are fairly consistent with regard to auxiliary selection, but other verb classes show selection discrepancies both within and across languages (Sorace, 2000).

Sorace (1993, 2000) maintains that this variation in auxiliary selection is orderly and follows certain patterns based on a hierarchy of auxiliary selection, with verbs that fall into certain syntactic and semantic categories being more likely to select either BE or HAVE as perfective auxiliary. She refers to this as the Auxiliary Selection Hierarchy (ASH, Sorace, 2000), derived partly from the previously formulated Unaccusative Hierarchy (Sorace, 1993). This hierarchy is shown below, with the verb class most likely to select BE at the top and the verb class most likely to select HAVE at the bottom.

Change of location	(core unaccusative verbs)
Change of state	
Continuation of a pre-existing state	
Existence of a state	
Uncontrolled process	
Controlled process (motional)	
Controlled process (non motional)	(core unergative verbs)

Table 1. Auxiliary selection hierarchy (from Sorace, 2000)

The hierarchy reflects the notion that unaccusativity is underlain by dynamic change, the most concrete form of which is change of location (Sorace, 1993). The change of state verbs

also express change, but this is by definition not a dynamic change. At the other end of the scale, the most concrete form of unergativity is volitional and agentive. The verbs in the various classes are described in more detail in the next section and examples are given.

The core verbs at either end of the hierarchy are the categorical unaccusative and categorical unergative verbs, and these are the most consistent in their use of BE or HAVE respectively in those languages that have a choice of perfective auxiliary (e.g. Italian, French, German, Dutch, Danish). The verbs that fall in the intervening classes in the hierarchy are more susceptible to variation, both across and within languages and also across time. Cross-language comparisons made by Legendre (2007) demonstrate the rigidity of the core categories, revealing a subset relationship between languages. For example, French has just a small group of verbs that take BE, but this group is a subset of the Dutch and German verbs that take BE, which in turn is a subset of the Italian verbs. This is depicted in the table below, taken from Legendre (2007).

ASH (Sorace, 2000)	Example	French	Dutch	German	Italian
Change of location	John came	E	E	E	E
Change of state	John died	E	E	E	E
	John went up	E	E	E	E
	John disappeared	A	E	E	E
	John lost weight	A	E	E	E
Continuation of state	His worry lasted	A	A	A	E
Existence of state	Dinosaurs existed	A	A	A	E
Uncontrolled processes	John shivered	A	A	A	A
Motional processes	John ran	A	A	E	A
Non-motional processes	John worked	A	A	A	A

(E = BE, A = HAVE)

Table 2. Split auxiliary selection in French, Italian, German, and Dutch (Legendre, 2007)

This table shows that there is a cut-off point for auxiliary use and that this is language specific. In the case of French the Change of State verbs take either HAVE or BE, indicating that the cut-off point in French seems to occur within this verb class. In Dutch, the cut-off point is less clearly delineated as depicted in the above table, as exceptions such as the Continuation of State verb *blijven* (remain/stay) and the Existence of State verb *zijn* (be) also take the auxiliary BE. Furthermore, the motional verbs in Dutch take BE rather than HAVE in

the presence of a telic modifier. This is discussed in more detail in section 4.2.6. Legendre further points out that English and Spanish verbs all use HAVE, whereas Slavic languages also show uniformity, but in these languages all verbs use BE.

Variation in auxiliary selection within a language has also been found to reflect the ASH. In a study of the acceptability of auxiliary use in Italian, Sorace found that Italian native speaker intuitions on auxiliary selection were less rigid for verbs in the peripheral categories than for verbs in the core categories. In particular, the native speakers found peripheral unaccusative verbs taking *avere* (HAVE) to be more acceptable than core unaccusative verbs taking *avere* (Sorace, 1993). Another study by Keller & Sorace (2000) comparing auxiliary use in two dialects of German revealed a similar pattern. The core categories at either end of the hierarchy showed very little variation between dialects, but in the peripheral verb categories, differences in auxiliary selection were found between the dialects. Furthermore, variation within the dialects was also perceived for the peripheral verbs.

With regard to diachronic language change, Sorace (1993) observes that Italian is conservative in its preservation of perfective auxiliary use, whereas in many other Romance languages there has been a tendency for BE to be superseded by HAVE. However, this evolution has taken place systematically, with the core unaccusative verbs being more resistant to change than the verbs in the peripheral category. In French, the peripheral verb types are less robust in their use of auxiliary than the verbs in the core categories and are more open to diachronic change, with a tendency for *être* (BE) to be replaced with *avoir* (HAVE) (Sorace, 1993). In the case of English, the modern language does not have a choice of auxiliary, but there is evidence of historical use of the auxiliary BE that has now been lost (Rydén & Brorström,1987). The same applies to Spanish, which also had a split auxiliary system in the past that is no longer seen in the modern language (Legendre, 2007). However, Legendre points out that if diachronic change occurs this always seems to take place in the same direction, with BE being replaced by HAVE. According to her there are no cases in which this change goes in the other direction. HAVE therefore seems to be the default choice.

Sorace suggests that the hierarchy of verb classes in the ASH may be applied not just to auxiliary selection but also to the other defining features of the unaccusativity/unergativity split: the diagnostics described in section 3.2. Core unaccusative verbs satisfy unaccusativity diagnostics more consistently across languages and less ambiguously within languages than verbs in the peripheral classes (Sorace, 2006). Along the same lines, the lack of correlation between diagnostic tests for unaccusativity pointed out by Zaenen (1988), as referred to in section 3.2, seems to apply to verbs in the peripheral verb classes in many of the examples she gives. In the case of the core verbs there is greater correlation between tests.

Another area in which the ASH is reflected concerns learnability. It is claimed that there is a tendency for the correct auxiliary to be acquired more easily in the core categories than in the peripheral categories. This applies to both first and second language learners (Sorace, 2000). In Sorace (1993), the acceptability judgements of Italian native speakers were compared with those of French and English native speakers who spoke Italian as a second language. This study examined the top four verb classes in the above hierarchy (those on the "unaccusative" side of the hierarchy), in addition to two other unaccusative verb classes with transitive and unergative alternative forms. All verb classes in this study use *essere* (BE) as auxiliary in standard Italian (Sorace, 1993). The Italian, French and English native speakers were asked to give acceptability judgements on a sliding scale on Italian sentences containing one or other of the perfective auxiliaries. The core unaccusative verbs showed the highest level of acceptability for *essere* (BE) and the lowest level of acceptability for *avere* (HAVE). This pattern was seen with all three language groups. The use of *avere* (HAVE) was considered increasingly acceptable for verbs lower down in the hierarchy, even by the native speakers, although this is generally considered ungrammatical in Italian.

The notion that the correct auxiliary is easier to learn for the core categories than for the peripheral categories in the ASH can be tested by investigating the interlanguage of second language learners whose native language differs from this second language with respect to auxiliary selection. This is the focus of the experiment described in chapter 6, which examines the use of the auxiliary in Dutch sentences by English native speakers.

The ASH is therefore an attempt to explain the variation within and between languages in the use of the perfective auxiliary for specific classes of verbs, by identifying orderly patterns within this variation based on the syntactic phenomenon of unaccusativity. According to Sorace (2000), these differences between languages reflect different "cut-off" points along the hierarchy for the use of BE and HAVE. Depending on the precise position of this cut-off point, verbs in the middle of the hierarchy would fall either on the BE side or on the HAVE side, but the core verbs would not be affected. It has been suggested by Levin & Rappaport Hovav (1995) that the position of this cut-off point could conceivably be considered a language parameter in the sense of Chomsky's principles and parameters of language (Chomsky, 1981). Sorace (2000) maintains that this variation between verb classes and between languages can be explained by the fact that the non-core verbs have differing degrees of lexical ambiguity, whereas the core verbs do not have this ambiguity and therefore do not show variability in auxiliary selection. However, the hierarchy does not necessarily imply that all intermediate categories will show variation, but only that if there is variation this will occur in the intermediate rather than the core verb categories (Keller & Sorace, 2003).

In summary, the core verbs in the Auxiliary Selection Hierarchy can be described as having consistent behaviour across and within languages regardless of the sentence context. Furthermore they are more resistant to diachronic change, show little variation in native speaker intuitions, and also show primacy in first and second language acquisition (Sorace, 2004). The different verb classes are described in greater detail below.

4.2. Verb classes

The seven different verb classes in the auxiliary selection hierarchy (ASH) proposed by Sorace (2000) are set out below. The first four verb classes are described as conditions or states and can be generally considered to be on the unaccusative side of the hierarchy. The last three verb classes are described as processes and can be considered to be on the unergative side of the hierarchy (Bard, Frenck-Mestre & Sorace, 2010; Sorace, 2000). However, as stated in the previous section and depicted in Table 2, the ambiguity increases towards the middle of the hierarchy and the distinction between the two verb groups becomes increasingly fuzzy and variable. Furthermore the ASH has been derived largely from

the behaviour of Italian verbs. Whereas the core unaccusative verbs and core unergative verbs seem to be fairly consistent across languages, this is not the case with the intermediate categories. As pointed out by Zaenen (2006), a verb may not mean exactly the same as its generally accepted translation in another language. As this paper examines the treatment of Dutch verbs in the interlanguage of English native speakers on the basis of the ASH, the classification needs to take into account the specific characteristics of Dutch verbs. In each of the verb classes described below, sample verbs are given. These have been taken partly from the comprehensive list of Dutch verbs in Lieber & Baayen (1997), partly from Sorace (2000), and partly from the German verbs used in the experiment in Keller & Sorace (2003), as German is closely related to Dutch and has similar auxiliary selection properties. The verb groups distinguished by Lieber & Baayen are slightly different from those distinguished by Sorace, and the classification method is not quite the same either. Lieber & Baayen's study focuses on the semantic aspects of the verbs rather than the syntactic distinction between unergatives and unaccusatives, and their verb classification is based first and foremost on auxiliary selection. This means that the categories that have variable auxiliary selection in Dutch do not match up precisely with the categories in Sorace's ASH. However, for the most part the grouping overlaps with that proposed by Sorace and therefore provides a useful source of relevant Dutch verbs. Where different classifications are given for particular verbs, consideration is given in this paper to the principles of the ASH as well as the way the verb in question is used in the Dutch language. The Dutch translations are based on entries in the English-Dutch *Handwoordenboek* by Van Dale (Hannay & Schrama, 1988).

4.2.1. Change of location

This core unaccusative verb class contains verbs that are inherently telic and dynamic (Sorace, 2000). Telic verbs have an inherent end point, and dynamic verbs have an inherent element of change (Velupillai, 2012). These verbs describe concrete movement from one place to another with a specific or inferable final location (Sorace, 2000; Lieber & Baayen, 1997). In languages with a choice of auxiliary, this verb type seems to be the most consistent in its use of BE. Examples below from Sorace (2000) show the use of BE in Italian (17a), French (17b) and Dutch (17c).

17. a) Maria e venuta alla festa
Maria is come to the party
"Maria came to the party."

b) Marie est arrivée en retard
Marie is arrived late
"Marie arrived late."

c) De brief is met de tweede post gekomen
The letter is with the second post arrived
"The letter arrived with the second post."[4]

The implication made by Sorace (2000) is that if a language does have auxiliary selection, this verb class will take BE in all cases. She further suggests that the use of the auxiliary BE is acquired first with these verbs and also that the intuitions of native speakers show the least variation with these verbs. Examples include *come, go, flee, arrive, rise, fall, escape* (Keller & Sorace, 2003; Lieber & Baayen, 1997). The semantic equivalents in Dutch (*komen, gaan, vluchten, arriveren, stijgen, vallen, ontsnappen*) all take the auxiliary *zijn* (BE).

4.2.2. Change of state

The verbs in this class all denote a change in state or condition. They may not be inherently telic in the same way as the previous verb class, but they generally indicate a direction that gradually approaches an implied end point (Sorace, 2000). This verb class differs between languages as to its auxiliary selection. In Italian, Dutch and German, for example, the auxiliary BE is selected, whereas in French these verbs take either HAVE or BE (Legendre, 2007). However, according to Sorace (2000) the auxiliary selection intuitions of Italian native speakers are slightly less rigid and consistent for this verb class than for the Change of Location verb class. Examples of verbs in this class include *die, grow, appear, happen, burst, become, succeed*. The Dutch semantic equivalents *sterven, groeien, verschijnen, gebeuren, barsten, worden, slagen* all take the auxiliary *zijn* (BE).

[4] The English translations given by Sorace have a simple past tense instead of a present perfect. The simple past would generally be used in English an unmarked context.

4.2.3. Continuation of (pre-existing) state

These stative verbs clearly do not have the dynamic character of the two preceding verb classes as they denote continuation rather than change. However, according to Sorace (2000) they do make a reference to change, namely the specific negation of any change. This in itself infers a final position or state. After all, if nothing changes the final position/state is the same as the initial position/state. These verbs primarily take *essere* (BE) in Italian, although there are exceptions, with *avere* (HAVE) being more acceptable if the subject is agentive (Sorace, 2000). Lieber & Baayen's Dutch list contains only one verb in this class, namely *blijven* (remain/stay), which has the auxiliary BE. However, their classification is based in the first instance on auxiliary selection, and the verb class with this name comes under "verbs selecting only BE". Other verbs that could be considered to fall in this class are *survive, persist, endure* (Keller & Sorace, 2003), *continue, rest* (author's own suggestions). Dutch translation equivalents are *overleven, persisteren, duren, continueren, rusten* respectively. All these Dutch verbs take the auxiliary *hebben* (HAVE), in contrast with *blijven*. Sorace (2000) notes that the verb *remain* shows exceptional auxiliary selection behaviour in a number of languages, including French and German as well as Dutch. She suggests that this exceptional behaviour is related to the inherent meaning of the verb, in which the final location is inferred (in common with the Change of Location and Change of State verbs).

4.2.4. Existence of state

In common with the previous verb class, these are stative, non-dynamic verbs. The difference between this verb class and that in the previous section lies in the fact that generally speaking these verbs make no reference at all to change or the lack of it, as this is simply not relevant to the meaning being conveyed (Sorace, 2000). These verbs represent physical or psychological states or physical positions (Keller & Sorace, 2003). Lieber & Baayen again list only one verb: *zijn* (to be), this being the only verb in this class to use the perfective auxiliary *zijn* (BE) in Dutch according to them.[5] Other verbs that may fall in this category as suggested by Keller & Sorace (2003), Sorace (2000) and the author of this paper are *exist, seem, suffice, kneel, sit, stand, live, resemble*, with the respective Dutch translations *bestaan, blijken, voldoen, knielen, zitten, staan, wonen, lijken*. Only the verbs *zijn* and *blijken* take the

[5] Lieber & Baayen classify the verb *blijken* (seem) as a Change of State verb, in contrast with the other authors referred to in this paper, who classify this as an Existence of State verb.

auxiliary *zijn* (BE) in Dutch, the rest taking *hebben* (HAVE) in most situations. However, there seems to be some discrepancy regarding the status of the "position" verbs such as sit, lie and kneel. Lieber & Baayen categorise these verbs as selecting only HAVE, but in other reference works (e.g. Koenen & Drewes, 1982) both auxiliaries are mentioned as being possible. These same verbs also show variation in German (Keller & Sorace, 2003).

4.2.5. Uncontrolled process

The types of processes included in this group tend to be non-volitional, non-agentive, stative, non-dynamic activities, such as involuntary reactions (motional or non-motional), bodily functions, or emissions of sound, light or smell (Sorace, 2000; Keller & Sorace, 2003). The subjects of these verbs are generally affected involuntarily in some way by the activity conveyed by the verb. According to Sorace (2000), these verbs vary in auxiliary selection in Italian, generally taking *avere* (HAVE) but showing variable behaviour under the influence of factors such as animacy and agentivity and the degree to which the subject is affected. In Dutch, French and German these verbs tend to select HAVE. Some examples (Keller & Sorace, 2003; Sorace, 2000; Lieber & Baayen, 1997) are *yawn, cough, sweat, tremble, rattle, shine, stink* (in Dutch respectively *gapen, hoesten, zweten, beven, ratelen, schijnen, stinken*).

4.2.6. Controlled motional process

The main verb sort in this class concerns manner of motion verbs, which encode the movement of the subject of the verb. This verb subject has a double role, as it may play an agentive, volitional role but at the same time it is an experiencer of the effect of the motion (Sorace, 2000). For example, in the sentence *John ran home*, the running action is agentive, but the subject John is also affected by the action, as his position changes as a result of the motion. These verbs display great variation in auxiliary selection, both across languages and within languages, depending on the context of the activity being referred to. According to Sorace (2000), these verbs generally select HAVE in Italian, French and Dutch but often select BE in German, although she states that native speaker intuitions are more variable for these verbs than for the core unergative verbs described in the following section. However, as she points out, the context of the sentence has a strong influence on auxiliary selection. For example, in Italian the auxiliary *essere* (BE) is more likely to be selected if the subject is non-volitional or non-agentive. This is illustrated in the sentences below, taken from Sorace

(2000). Sentence 18a has a volitional subject and the preferred auxiliary is *avere* (HAVE), whereas in sentence 18b with a non-volitional subject there is a preference for *essere* (BE).

18. a) Il pilota ha/?e atterrato sulla pista di emergenza
The pilot has/is landed on the runway of emergency
"The pilot landed on the emergency runway."

b) L'aereo e/?ha atterrato sulla pista di emergenza
The plane is/has landed on the runway of emergency
"The plane landed on the emergency runway."

Besides volition and agentivity, these verbs are also affected by telicity. Sentences containing a telic modifier tend to select BE in some languages, including Dutch. Sorace states that in Italian this only applies to a small subset of motion verbs, whereas in Dutch all motion verbs systematically switch to *zijn* (BE) in the presence of a telic modifier. This is illustrated in the sentences below. Sentence 19b has the telic modifier *naar huis* (to home), and the auxiliary consequently changes from *hebben* (HAVE) to *zijn* (BE).

19. a) John heeft snel gelopen.
John has quickly walked.
"John walked quickly."

b) John is naar huis gelopen.
John is to home walked.
"John walked home."

This phenomenon has been widely studied and described in Dutch (Zaenen, 1988; Belien, 2012, among others) and the general consensus seems to be that telic adverbials, either temporal ("until three o'clock") or positional ("into the river"), telicise the predicate and trigger a switch in auxiliary from HAVE to BE. Only directional positional modifiers telicise the predicate and trigger this change in auxiliary selection. Non-directional positional modifiers denote a continuous rather than a telic activity and consequently no auxiliary change is triggered, as shown below in the contrast between 20a (with directional modifier) and 20b (with non-directional modifier).

20. a) John is in de sloot gesprongen.
John is in the ditch jumped.
“John jumped into the ditch.”
b) John heeft in de sloot gesprongen.
John has in the ditch jumped.
“John jumped up and down in the ditch.”

Examples of manner of motion verbs (Keller & Sorace, 2003; Lieber & Baayen, 1997) are *swim, run, climb, jump, fly* (in Dutch *zwemmen, rennen, klimmen, springen, vliegen*).

4.2.7. Controlled non-motional process

The verbs in this group are considered to be the core unergative verbs. They can be characterised as having a subject with a high degree of volition and agentivity that tends to be unaffected by the process denoted by the verb (Sorace, 2000). These verbs are fairly consistent in auxiliary choice both within and between languages, systematically selecting the auxiliary HAVE (Keller & Sorace, 2003) without this being dependent on the semantic content of the verb or the context of the sentence. Contrary to the manner of motion verbs, adding a telic adverbial does not trigger a change in auxiliary from HAVE to BE. This is illustrated for Italian in the sentence below, taken from Sorace (2000).

21. I poliziotti hanno lavorato fino all'alba.
The policemen have worked until the dawn.
“The policemen worked until dawn.”

Examples of controlled non-motional process verbs include *talk, work, play, dance, laugh, shout, think* (in Dutch *praten, werken, spelen, dansen, lachen, schreeuwen, denken*). Of course, verbs such as *dance* and *play* may involve motion of some kind, but the point is that these verbs generally describe an agentive and volitional process rather than provide information on the relocation of the verb subject.

In summary, the verb classes on the unaccusative side of the hierarchy gradually increase in telicity, dynamism and subject affectedness from the periphery to the core, whereas the verb classes on the unergative side of the hierarchy gradually increase in agentivity and volition and decrease in subject affectedness from the periphery to the core. All verbs with a single argument (intransitive verbs) can be classified within this hierarchy (Sorace, 2000).

4.3. Dutch verb characteristics

The hierarchy described above has been derived from the behaviour of verbs in Italian. Although clear parallels can be drawn with Dutch verbs, there are also significant differences. Only the first two verb types (Change of Location and Change of State) systematically take the auxiliary BE in Dutch. The peripheral unaccusative categories (Continuation/Existence of State) take HAVE (with very few exceptions, as mentioned in the previous section), and the peripheral unergative category (Uncontrolled Process) takes HAVE with no exceptions. However, in Dutch the manner of motion verbs in the Controlled Motional Process category systematically take BE when the motion is telicised (Belien, 2012). This contrasts with Italian, where the shift in auxiliary only applies to a subset of motion verbs, and with French, where the auxiliary is not affected by aspectual changes such as telicity (Sorace, 2000).

In Dutch it seems that aspectual difference is crucial to auxiliary selection and applies systematically. There are many verbs that can be used with either auxiliary depending on the sentence context. Lieber & Baayen (1997) provide a fairly exhaustive list of these verbs. Besides the manner of motion verbs described above, the list includes means of motion verbs (e.g. *schaatsen,* “skate”), path of motion verbs (e.g. *draaien*, “turn”), speed of motion verbs (e.g. *spoeden*, “hasten”), manner of position verbs (e.g. *zweven*, “float”), a range of verbs that have a transitive alternant (e.g. *smelten*, “melt”), and a few exceptional transitive verbs that can be used with the auxiliary BE (e.g. *volgen*, “follow”; *vergeten,* “forget”). These are exceptional as all other Dutch transitive verbs always take HAVE. (See sentences 23a and 23b for examples.) Lieber & Baayen argue that auxiliary selection is based on a semantic principle rather than syntactic properties such as argument structure, calling the relevant semantic principle IEPS (Inferable Eventual Position or State), a feature that a predicate either possesses or does not possess. As described in section 3.2, a verb with this feature

provides information on the final position or state of the argument. According to Lieber & Baayen, if the eventual position or state of the subject is implied by the lexical meaning of the verb or by the context of the sentence, the auxiliary BE is selected in Dutch. The parallel with the telicity requirement for Sorace's core unaccusatives is clear. If an eventual position or state can be inferred, this has the effect of telicising the act or event communicated by the verb. With regard to the various motion-related verbs in Lieber & Baayen's list, these could be analysed in the same way as the manner of motion verbs described above, taking BE if there is a telic modifier and HAVE if there is not. These motion verbs are widely assumed to be unergative in a general sense but if they are telicised by adding an end point they are unaccusative (Belien, 2012). This makes sense in relation to the verb classes described above. With the addition of an end point these verbs could be considered to be transformed into the core unaccusative Change of Location verb types, whereas in a neutral context the meaning of the verb encodes the type or means of the motion itself, leaning more towards the volitional, agentive properties of the core unergative verbs. It could therefore be concluded that a motion verb actually has two different forms with different meanings, depending on the context. In sentences 19a and 19b, repeated below as 22a and 22b, sentence 22a describes the act of walking (quick), whereas sentence 22b says nothing about the act itself but places the focus on the eventual location (home).

22. a) John heeft snel gelopen
John has quickly walked
"John walked quickly."

b) John is naar huis gelopen
John is to home walked
"John walked home."

In summary, in the case of Dutch motion verbs HAVE is used when the motion is viewed as a type of act, whereas BE is used when the motion is viewed as a change of location (Belien, 2012).

Some of the verbs in the Existence of State verb class, namely the "position" verbs such as lie, sit and kneel, show a similar variation in auxiliary selection. As pointed out by Keller &

Sorace (2003) these verbs may be interpreted in various different ways. For instance, in English the verb *to sit* could be used to refer to the activity of becoming seated ("assuming position") or to the situation of being seated ("maintaining position"). In the case of Dutch, Renkema (2002) makes a distinction between verbs that encode an activity or event, which select HAVE, and verbs that encode a situation or change, which select BE. This distinction seems to be applicable to the position verbs considered here. Renkema uses the same analysis to explain the deviant auxiliary selection of the few transitive verbs in Dutch that take both HAVE and BE, such as *verliezen* (lose) and *volgen* (follow). The sentences below illustrate this. The Dutch sentences are from Renkema (2002); the translations are provided by the author of this paper.

23. a) Ik heb al mijn geld verloren (in de casino).
I have all my money lost (in the casino).
"I lost all my money at the casino."

b) Ik ben mijn portemonnee verloren.
I am my wallet lost.
"I have lost my wallet."

Sentence 23a describes the actual event of losing money, which took place in the casino. Sentence 23b does not describe the losing event itself, as the speaker would probably be unaware of exactly what had happened, but instead the emphasis is placed on the situation as it currently stands ("I now have no wallet"). A parallel can be seen here with Sorace's distinction in the ASH between the unaccusative "condition" or "state" verbs, which encode the current situation resulting from the change and relate to subject affectedness, and the unergative "process" verbs, which encode the activity or event taking place and relate to agentivity.

The groups of intransitive verbs with a transitive alternant in Lieber & Baayen's list are also widely discussed in the literature. These are the verbs such as *break* and *open*, which have already been discussed in section 3.1. They are either used transitively with an external and an internal argument (*John opened the door*) or intransitively with one internal argument (*The door opened*). Sorace (2000) refers to these verbs as anticausatives, from the

assumption that in the languages discussed in her paper they are derived from their transitive (causative) alternatives. Based on overgeneralisation errors made by children as well as adult second language speakers, the transitive structure is considered to be the default structure (Montrul, 2000). An example of such an overgeneralisation error is given below (from Montrul, 2000).

24. a) The rabbit disappeared.
    b) * The magician disappeared the rabbit.

The intransitive verb *disappear* has no transitive alternant in English. Sentence 24b is consequently ungrammatical, as here the verb is used transitively. According to Montrul, these errors are not made in the other direction, with transitive verbs used intransitively, suggesting that the intransitive form is derived from the transitive form. The behaviour of these verbs differs between languages, according to Sorace (2000). In Dutch the transitive forms of these verbs always use HAVE (as do all transitive verbs, with the exception of those mentioned in the previous paragraph), whereas the intransitive forms always use BE, as shown in sentences 25a and 25b below (taken from Algemene Nederlandse Spraakkunst, 2013, translated by the author of this paper). In 25a the verb is used transitively, with a subject (*the doctor*) and an object (*him*). In sentence 25b the verb is intransitive and has only a subject (*he*) with a patient thematic role.

25. a) De dokter heeft hem genezen.
       The doctor has him cured.
       "The doctor has cured him."
    b) Hij is genezen.
       He is cured.
       "He has recovered (from his illness)."

Lieber & Baayen (1997) characterise these verbs in their intransitive form as Change of State verbs, and they seem to match Sorace's description of the Change of State verb class in the Auxiliary Selection Hierarchy, therefore the choice of BE as perfective auxiliary seems fitting. Typical examples of these verbs are *melt, break* and *dry*, which have both transitive and

intransitive forms in both Dutch and English. However, in English of course there is no difference in auxiliary selection as English uses only HAVE. Due to the possible ambiguity of these verb types, they will not be used in the experiment. As intransitive Change of State verbs they would fall in the category of unaccusatives that take the auxiliary BE. However, any errors in auxiliary selection made by the participants may be due to their double role - as transitive verbs they appear with HAVE - rather than due to the fact that they are not core unaccusatives. This would make it difficult to make firm claims regarding the nature of any errors made.

In summary we can say that Dutch is fairly systematic in its auxiliary use. In the case of alternating verbs that have a choice of auxiliary, the selection always depends on the semantic aspect of the verb and the context of the sentence. With these ambiguous verbs, the distinction between HAVE and BE seems to be based on whether the verb encodes an activity or a situation; in other words, a process versus a condition/state. This is the same distinction made by Sorace in her description of unaccusative and unergative verb types. However, in contrast with the Italian verbs discussed by Sorace (2000), in Dutch almost all the peripheral unaccusative verbs, which also encode a state or situation, select the auxiliary HAVE. In the experiment described in the following chapters, we examine the interlanguage of English native speakers to see whether and how the verb classification based on the ASH is reflected in the choice of auxiliary by these second language speakers in Dutch.

The similarities and differences between English and Dutch with respect to auxiliary use in the various verb classes are summarised in Table 3.

Verb class	English	Aux	Dutch	Aux
Change of location				
	John has arrived	H	Jan is gearriveerd	B
	The lion has escaped	H	De leeuw is ontsnapt	B
	The plane has landed	H	Het vliegtuig is geland	B
Change of state				
	John has died	H	Jan is gestorven	B
	The balloon has burst	H	De ballon is gebarsten	B
	The shirt has shrunk	H	De hemd is gekrompen	B
	The child has grown	H	Het kind is gegroeid	B
Continuation of state				
	John has stayed	H	Jan is gebleven *	B
	The bear has survived	H	De beer heeft (het) overleefd	H
	The party has lasted for hours	H	Het feest heeft uren geduurd	H
Existence of state				
	John has been ill	H	Jan is ziek geweest **	B
	Dragons have never existed	H	Draken hebben nooit bestaan	H
	John has lived there	H	Jan heeft daar gewoond	H
Uncontrolled processes				
	John has sneezed	H	Jan heeft geniesd	H
	The tap has leaked	H	De kraan heeft gelekt	H
	The hinge has creaked	H	Het scharnier heeft gepiept	H
Controlled motional processes				
	John has walked quickly	H	Jan heeft snel gelopen	H
	John has swum in the sea	H	Jan heeft in zee gezwommen	H
	John has walked home	H	Jan is naar huis gelopen ***	B
	John has swum to the shore	H	Jan is naar de kust gezwommen ***	B
Controlled non-motional processes				
	John has worked	H	Jan heeft gewerkt	H
	John has spoken	H	Jan heeft gesproken	H
	The children have played	H	De kinderen hebben gespeeld	H

H = HAVE, B = BE
*Exceptional use of BE within this category
** Exceptional use of BE within this category
*** Auxiliary use with telic modifier

Table 3. A summary of auxiliary selection in Dutch and English

5. Research question

In this paper we aim to investigate the learnability of the perfective auxiliary in Dutch by examining the interlanguage of English speakers of Dutch as a second language for compatibility with the Auxiliary Selection Hierarchy (Sorace, 2000). As English is a language with no choice of perfective auxiliary, English native speakers learning Dutch have to learn the correct auxiliaries for the different verbs without being able to use features from their native language as a learning aid. The ASH suggests that there are different types of intransitive verbs, of which some types (unergatives) are more likely to select HAVE whereas other types (unaccusatives) are more likely to select BE in languages that have a choice of auxiliary. The hierarchy also suggests that preferences are gradient, with the core categories at each end of the hierarchy showing the least variation both within and across languages, and the peripheral categories showing the most variation. As mentioned above (Chapters 1 and 2), the notion of unaccusativity is widely considered to be a syntactic property of certain verb types and a feature of Universal Grammar. This suggests that the core unaccusatives should be the easiest to learn for first language and second language learners (Sorace, 2004).

The research question being investigated here is whether the perfective auxiliary selection of English speakers of Dutch as a second language supports the Auxiliary Selection Hierarchy. In other words, will the perfective auxiliary be easier to acquire for the core unaccusative and unergative verbs than for the verbs in the intervening verb classes in the hierarchy? A secondary question is whether the language learners will be more inclined to use the auxiliary BE for unaccusatives across the board and the auxiliary HAVE for unergatives across the board. To answer these questions we are comparing Dutch auxiliary selection by English native speakers with varying levels of proficiency. Their task is to fill in the correct auxiliary in a number of Dutch sentences with a range of present perfect verb forms. On the basis of the ASH, the characteristics of Dutch intransitive verbs and the features of the English language, we can make certain predictions with respect to the findings. These predictions relate on the one hand to the choice of auxiliary and on the other hand to the extent to which these choices differ from common usage in Dutch.

5.1. General prediction

If the preferences for the perfective auxiliary are gradient, as suggested by Sorace and summarised in the ASH, we would expect to see a gradual transition from the more frequent selection of BE to the more frequent selection of HAVE along the verb classes in the hierarchy. However, a transfer effect may affect this gradience, with a bias towards HAVE in accordance with auxiliary usage in English. Furthermore, taking into account the ambiguity of the Controlled Motional Process verbs in Dutch, as described in section 4.3, this intermediate unergative verb class may deviate from the gradient pattern.

5.2. Unaccusative verbs

1) Across the board the selection of the auxiliary BE will follow the hierarchy, with the most frequent selection in the core verb class (Change of Location) and the least frequent selection in the peripheral verb classes (Continuation/Existence of State).
2) Following on logically from the above, the English subjects will make fewer auxiliary selection errors with the core unaccusative (Change of Location) verbs than with the intermediate unaccusative (Change of State) verbs. Errors in the peripheral categories (Continuation/Existence of State) will be limited due to a transfer effect (almost all of these verbs select HAVE in both English and Dutch).
3) The less advanced Dutch speakers will learn the correct auxiliary for the core unaccusative (Change of Location) verbs before the correct auxiliary for the intermediate unaccusative (Change of State) verbs. In the case of the advanced Dutch speakers, this may be impossible to establish due to a ceiling effect. The discrepancy between these two verb classes will therefore be greater for the less proficient Dutch speakers than for advanced and near-native speakers.

5.3. Unergative verbs

1) Across the board there will be little variation between the three unergative verb classes, with a strong preference for HAVE in all cases. There is no reason for the subjects to choose BE as this is not the correct auxiliary for these verbs in either language. Should the subjects be in any doubt, they will probably be influenced by their native language and will resort to the default auxiliary HAVE.

2) Contrary to the ASH, the intermediate unergative class (Controlled Motional Process) may show a slightly more frequent use of BE as auxiliary, due to the fact that the motion verbs in this class vary in Dutch according to the context of the sentence, as described in section 4.2.6. The greatest variation in auxiliary selection (in other words, a greater tendency to select BE rather than HAVE) is therefore expected in this intermediate unergative verb class.

If these predictions are borne out, this would seem to provide support not only for the ASH but also for the claim that the hierarchy applies to learnability as well as to cross-language variation. Evidence of enhanced learnability would in turn lend further support to the widely held notion that unaccusativity is indeed a syntactic phenomenon and a feature of UG (Alexiadou et al., 2004). However, the findings are unlikely to enable us to make any claims regarding unergative verbs, as the similarities between English and Dutch will probably neutralise any possible effect. Perhaps a similar experiment conducted among native speakers of a Slavic language that uses only the auxiliary BE could shed more light on the perception of these verbs.

6. Experiment

This chapter describes an experiment carried out to test the use of the perfective auxiliary in Dutch by English native speakers of Dutch as a second language.

6.1. Subjects

A total of 31 people took part in the experiment (7 male and 24 female), all adult English native speakers who lived in the Netherlands at the time of the experiment. They were recruited via the British School in The Netherlands (12), Leiden University (3), social media sites (1) and personal acquaintances (15). The British School in The Netherlands is a group of associated schools located in the vicinity of The Hague offering primary and secondary education based on the British curriculum. The participants recruited from this school were all secondary school teachers. They were approached through a contact person, who also participated in the experiment. The participants recruited via Leiden University were students (one Bachelor and two PhD students), and the remaining participants had a range of different backgrounds. Participation was voluntary and the subjects did not receive any payment for taking part.

Prior to completing the experimental part of the study, the participants were asked to fill in a questionnaire requesting certain personal details and relevant background information, such as their use of Dutch, knowledge of other languages and country of birth. This personal details questionnaire can be found in Appendix II. The youngest participant was 28 years of age and the oldest was 73, with an average age of 52. The age of arrival in the Netherlands ranged from 18 to 50 with an average AOA of 31. The length of residence in the Netherlands ranged from a minimum of 1 year to a maximum of 51 years, with an average of 21 years. 7 of the participants had a secondary school level of education, 14 had a university level of education and 10 had followed or were following postgraduate education. 4 participants had never taken Dutch classes, 22 had taken beginners' classes, 4 had followed lessons at an intermediate level and 1 at an advanced level. 7 of the participants used Dutch in their day-to-day life more than half the time, 7 approximately half the time, and 17 less than half the time.

In order to establish the participants' level of proficiency in the Dutch language, they were also asked to fill in a self-assessment form after completing the experimental part of the study. This form contained a table in which the participants were asked for their opinion of their ability in speaking, understanding and grammar as well as their general level of proficiency. The form also included a number of so-called can-do statements, which are statements that begin with the words "I can ...." and then go on to describe competencies and activities of varying difficulty. These statements are based on the reference levels set out in the Common European Framework of Reference for Languages (CEFR, 2001), a framework providing comprehensive descriptions of the knowledge and skills required to use language in practical situations and to communicate effectively at various levels. Broadly speaking, six common reference levels are distinguished. These fall into the three categories of basic (A), intermediate (B) and advanced (C), with each category having two levels: low and high. The CEFR labels these levels A1 (Breakthrough), A2 (Waystage), B1 (Threshold), B2 (Vantage), C1 (Effective Operational Proficiency) and C2 (Mastery). The can-do statements used for this experiment have been derived from the global scale in the Common European Framework of Reference for Languages (CEFR, 2001: p. 33), which is a summarised version of the self-assessment grid provided in the CEFR. The self-assessment form used in the experiment can be found in Appendix IV. The first six can-do statements on the form correspond to level A (3 x A1 and 3 x A2), the following seven statements correspond to level B (4 x B1 and 3 x B2), and the final seven statements correspond to level C (4 x C1 and 3 x C2). Although this self-assessment procedure may not provide conclusive information on proficiency levels, it can be used as a general guide. Research suggests that self-assessment is an effective way of establishing general proficiency levels as long as nothing is at stake for the subject, such as a job offer or admittance to a course (CEFR, 2001).

6.2. Materials

This experiment was based on a sentence completion test, in which the participants were presented with Dutch sentences with a missing word and asked to fill the gap with the word that they considered to be the most fitting within the context of the sentence. A sentence completion test was considered preferable to the magnitude estimation test used to elicit acceptability judgements in Sorace (1993) and Sorace & Keller (2005). This is because the test subjects in those studies were natives or near-natives, whereas the proficiency of the

participants in this experiment is expected to be lower and more diverse. Subjects with a limited knowledge of the target language cannot be expected to have keen enough judgements to be able to grade the acceptability of a sentence.

The experiment consisted of 36 test sentences and 36 filler sentences. The 36 test sentences were divided into the two categories of unaccusative and unergative, with 18 sentences in each category. Each category was further subdivided into three verb classes: core, intermediate and peripheral. This gives the six verb classes listed below.

Unaccusative

Core = Change of location (COL)

Intermediate = Change of state (COS)

Peripheral = Continuation/existence of state (CES)

Unergative

Core = Controlled non-motional process (CNP)

Intermediate = Controlled motional process (CMP)

Peripheral = Uncontrolled process (UCP)

These verb classes reflect the verb classes in the ASH as described in section 4.2 above, the exception being that the two classes Continuation of State and Existence of State from Sorace's hierarchy have been combined into one category of stative verbs (Continuation/Existence of State, or CES). The reason for this is that there are few verbs to choose from in these two classes and also that the dividing line is difficult to draw in some cases. All these verbs share the feature of being stative verbs, and although the former group specifically encodes the denial of change, the latter group also conveys a lack of change. Semantically, therefore, these verbs could all be considered to belong to a combined "stative" class. Sorace states that these verb classes are commonly combined into a single class as they are all non-dynamic (Sorace, 2000). Perlmutter also suggests that the durative verbs could be considered a sub-class of the verbs of existence (Perlmutter, 1978).

For the test sentences, six verbs were chosen from each of the six verb classes listed above. The verbs in each class were balanced with regard to the frequency and length of the past

participle in order to ensure similarity between verb classes and eliminate any word frequency bias. The Subtlex database was used to establish word frequency (Subtlex, 2015). This database is based on word counts in Dutch subtitles in films and television programmes, and has been shown to be an effective measure of word frequency for the Dutch language (Keuleers, E., Brysbaert, M. & New, B., 2010). The frequencies for the 36 chosen verbs can be found in Appendix V. These verbs were used to form 36 sentences with a present perfect verb construction in the main clause. This verb construction consists of a perfective auxiliary and a past participle. The reason for placing the test construction inside a main clause is that Dutch is a V2 language, in which the inflected verb (in this case the auxiliary) raises to the second position in the main clause, whereas the participle remains in the canonical clause-final position. This creates the maximum distance between the auxiliary and the participle, thus avoiding collocation effects that may aid statistical learning or formulaic learning (in a subordinate clause the auxiliary does not raise to the V2 position and therefore appears next to the participle). In all 36 test sentences the main clause is preceded by a subordinate clause or an adverbial clause, causing the inversion of the subject and inflected verb and thus creating an even greater distance between the auxiliary and the participle. This is illustrated in sentences 26a and 26b, which are two of the test sentences used in the experiment. The omitted word - the inflected auxiliary – is given in brackets. The auxiliary and the participle are shown in bold type.

26. a) Sinds de loodgieter is geweest **(hebben)** de oude waterleidingen in de badkamer niet meer **gelekt.**
"Since the visit of the plumber the old water pipes in the bathroom have no longer leaked"

b) In de loop van vorige week **(zijn)** de soldaten met alle hulpgoederen naar het oorlogsgebied **vertrokken**.
"During the course of last week the soldiers departed for the war zone with all their equipment."

In all test sentences the perfective auxiliary is the missing word. In the core and intermediate unaccusative verb classes this is the auxiliary *zijn* (BE), and in the peripheral unaccusative

class and all three unergative verb classes this is the auxiliary *hebben* (HAVE). The sentences are designed so that the perfective auxiliary is the only possible choice to form a grammatical sentence. This is because the past participle needs to be paired with a perfective auxiliary in order to complete the perfect verb construction. Although past participles can also be used in passive constructions, passivisation is not possible with intransitive verbs (with the exception of the impersonal passive described in section 3.2, but this would then require the insertion of an expletive). The purpose is therefore to compel the participant to select one of the two possible auxiliaries.

Verbs with exceptional auxiliary selection behaviour within their verb class, as highlighted in Table 3 in section 4.3 (*zijn* and *blijven*), have not been used for the test sentences, in order to ensure uniformity within the verb class and thus facilitate comparisons between verb classes. For the same reason, all verbs in the Controlled Motional Process class (CMP) are used in sentences with an atelic modifier in order to preserve the unergative character of the verbs and thus prevent the telicisation of the predicate and the corresponding switch in auxiliary from HAVE to BE, as explained in section 4.2.6. An example of this is given in sentence 27, one of the test sentences from the CMP class. Again, the omitted auxiliary is in brackets and the auxiliary and participle are shown in bold type. The action described by the verb is atelic due to the atelic temporal modifier *maandenlang* (for many months). With the addition of this modifier, the sentence is describing the act of driving rather than the direction of the journey.

27. Wegens zijn nieuwe baan **(heeft)** Bart maandenlang tussen Groningen en Leiden heen en weer **gereden**.
"Because of his new job, Bart drove back and forth between Leiden and Groningen for many months."

As mentioned above, the six verbs used for the sentences in the unaccusative CES class all select the auxiliary *hebben* (HAVE), as do all 18 verbs in the three unergative categories. Consequently there are 24 test sentences with the auxiliary HAVE and 12 sentences with the auxiliary BE, creating an imbalance between the two auxiliaries. In order to address this imbalance, 12 of the filler sentences also have a form of the verb BE, either as a copula or in

a passive construction. The aim of the 36 filler sentences is to distract attention from the purpose of the experiment by including sentences in which different language problems are presented. Like the test sentences, the filler sentences all have two clauses. Besides the 12 filler sentences with the auxiliary BE, there are 12 sentences in which the definite article is omitted and 12 sentences in which a conjunction is omitted. The sentences with the missing article present the problem of gender choice. Dutch has two genders, common and neuter, each requiring a different definite article in the singular form (*de* and *het* respectively). Learning the correct article is notoriously difficult for second language learners, as there are very few discernible patterns and there often seems to be no semantic motivation for gender assignment. The filler sentences with missing conjunctions suggest that the Dutch word for *because* is missing. However, in Dutch this can be translated either with a coordinating conjunction (*want*) or with a subordinating conjunction (*omdat*). This affects the word order of the clause introduced by the conjunction, as the inflected verb raises to the second position in main clauses only. This difference in word order is often tackled in Dutch language classes, so should this grammatical feature capture the attention of the participants, it may serve to distract attention from the present perfect constructions in the experiment and thus prevent a specific focus on auxiliary selection. These 36 filler sentences together with the 36 test sentences give a total of 72 sentences. The test sentences in each class with English translations are given in Appendix I.

6.3. Procedure

The test sentences and filler sentences were quasi randomised, ensuring an even distribution of auxiliaries, verb classes and fillers throughout the test, and numbered from 1 to 72. The test was presented to the participants in two ways. The group of participants from the British School in The Netherlands plus one other participant recruited through personal acquaintances were given the test on paper and they completed it in controlled conditions. The remaining participants completed the test online in their own environment. The decision to present the test online as well as in controlled conditions was necessitated by the difficulty in recruiting sufficient numbers of suitable participants. All participants were instructed to fill the blank space in each sentence with the word that they considered to be the most appropriate in combination with the remainder of the sentence. No further instructions were given in order to avoid attention being drawn to the perfective

constructions. The online participants received slightly different instructions, asking them to refrain from using grammar books or requesting help from others. The instructions for both participant groups together with the sentence completion form can be found in Appendix III. The participants first answered the personal details questionnaire (Appendix II), then they completed the sentences (Appendix III), and finally they filled in the self-assessment form (Appendix IV). The test in its entirety took between 20 and 45 minutes.

A control group of 10 Dutch native speakers was also asked to complete the 72 sentences. The purpose of this control group was to establish a benchmark for auxiliary use in Dutch to enable comparisons with the answers given by the English test subjects. The results from this control group and from the test subjects are presented in the following chapter.

7. Results

7.1 Proficiency

Based on the responses given on the self-assessment form, the English participants were divided into three proficiency groups corresponding to the CEFR proficiency categories A (basic), B (intermediate) and C (advanced) as described in section 6.1. This was done as follows. If the answer *yes* was given for more than half of the can-do statements in each of the six CEFR sub-categories (A1-C2), that level was assumed to have been achieved. The highest level to be achieved was then taken as the proficiency level for the subject in question. In the case of subjects who did not fall into a definitive category, their own assessment of their abilities as given in the proficiency table was used as the decisive factor. This division placed four subjects in category A (basic), fifteen subjects in category B (intermediate) and twelve subjects in category C (advanced). A significant correlation was found between the level of proficiency of the subjects and their age of arrival in the Netherlands (rho = 0.655, $p < .001$) as well as between proficiency level and years of residence in the Netherlands (rho = 0.631, $p < .001$), and proficiency level and use of Dutch in day-to-day life (rho = 0.523, $p = .001$). Age of arrival and years of residence are generally considered to be influential factors in successful second language acquisition (e.g. Johnson & Newport, 1989), and this correlation seems to endorse the effectiveness of the self-assessment form in establishing proficiency. No correlation was found between proficiency level and the level of Dutch classes followed by the subjects (rho = .283, $p = .062$). Considering that only a small minority of the subjects had followed classes at an intermediate or advanced level, this is not really surprising, and it suggests that the Dutch language knowledge of the subjects is based primarily on implicit learning methods, which are natural and subconscious, rather than explicit methods, which are conscious and deliberate (Ellis, 1994).

All four subjects in category A and one subject in category B failed to provide appropriate responses for over half of the test sentences. An appropriate response was considered to be any form of the auxiliary *zijn* (BE) or *hebben* (HAVE). These five subjects either failed to fill in any word at all or used a word that was not possible within the context of the sentence, such as a modal verb. As explained in section 6.2, the past participle in the test sentences needs

to be paired with a perfective auxiliary in order to form a grammatical sentence, whereas a modal verb would require an infinitive rather than a participle. However, the sentences proved to be too difficult for participants with a limited knowledge of Dutch, and they were consequently unable to use the information provided within the sentences to deduce the appropriate response. Due to the lack of usable data, these five subjects were excluded from further analysis, leaving two remaining proficiency levels: intermediate (B1/B2) with fourteen subjects and advanced (C1/C2) with twelve subjects.

7.2 Auxiliary selection

In order to establish typical auxiliary selection in Dutch, the responses given by the Dutch native speakers were first analysed. The selected auxiliaries for the verbs in the test sentences, grouped according to verb class, are summarised in the graph in Figure 10.

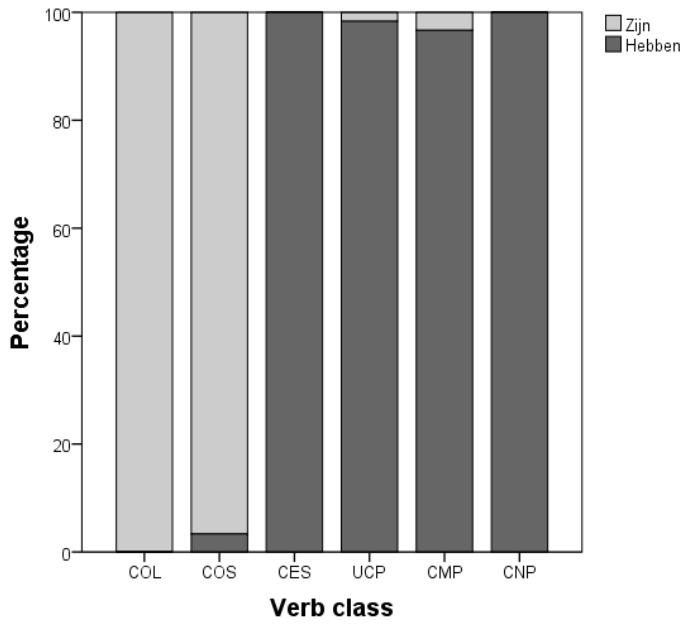

Fig. 10 Distribution of auxiliary selection per verb class by Dutch native speakers

As the figure shows, consensus among the ten Dutch native speakers is high. There are very few deviant responses, and the choice of auxiliary is in agreement with the general consensus in Dutch grammar books and dictionaries, with the first two unaccusative categories COL (Change of Location) and COS (Change of State) selecting BE, the peripheral unaccusative category CES (Continuation/Existence of State) selecting HAVE, and all three unergative categories also selecting HAVE. The absolute numbers and percentages are given

in Table 1 in Appendix VI. On the basis of this outcome we are taking these auxiliary choices as a benchmark for auxiliary use in Dutch.

The responses given by the twenty-six English test subjects in proficiency levels B (intermediate) and C (advanced) were then analysed in further detail. Any form of the verb *zijn* (BE) or *hebben* (HAVE) was accepted as an appropriate answer. No attention was paid to the tense of the auxiliary (in fact, in many of the sentences a pluperfect construction with an auxiliary in the past tense would also be grammatical). Spelling errors and inflection errors in terms of number or person were not taken into consideration either. After all, the matter under investigation is the choice of auxiliary between HAVE or BE, not the ability of the test subjects to produce grammatically correct Dutch. For each sentence the selected auxiliary was noted for each subject. Inappropriate responses as described in section 7.1 were treated as missing data (N = 65, 6.9%). The selected auxiliaries per verb type for all twenty-six English native speakers are summarised in Figure 11.

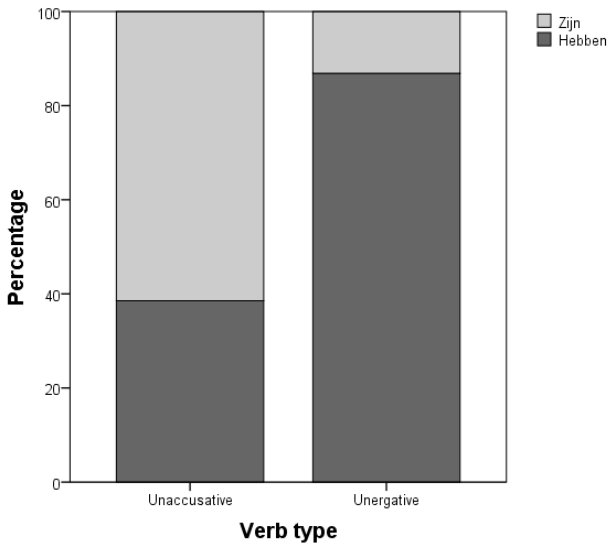

Fig. 11 Distribution of auxiliary selection per verb type by all English native speakers

The graph shows a clear distinction between the auxiliary selection for the unaccusative verbs and for the unergative verbs. For the unaccusatives, 61% of all responses gave *zijn* (BE) and 39% gave *hebben* (HAVE). For the unergatives these percentages are 13% and 87% respectively. A chi-square test shows this to be a significant difference (Pearson X2 = 218,640, df = 1, $p < .001$).

In order to enable finer distinctions to be made between the tested verbs and to obtain an indication of the perception of these verbs by the English test subjects, we then examined the auxiliary selection behaviour for the different verb classes within each verb type. The collective findings for all twenty-six English native speakers are depicted in Figure 12.

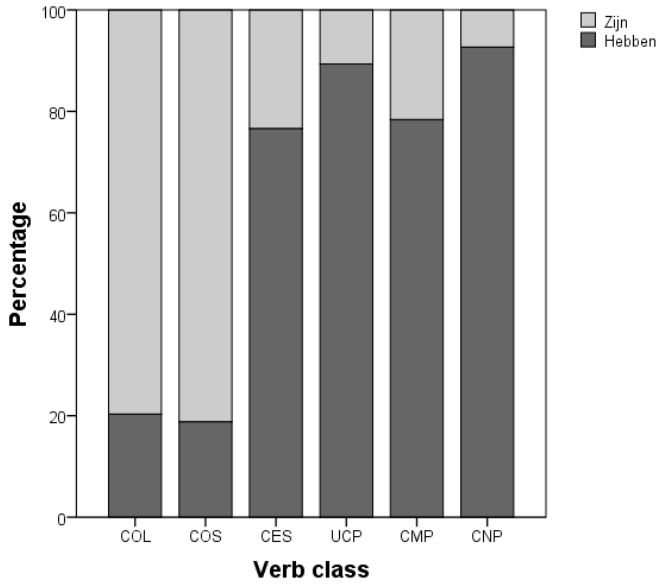

Fig. 12 Distribution of auxiliary selection per verb class by all English native speakers

Here we see that there is little difference in auxiliary selection between the first two unaccusative verb classes, COL (Change of Location) and COS (Change of State), with approximately eighty percent of responses opting for BE, the preferred choice of the Dutch native speakers. The remaining four verb classes – the peripheral unaccusative CES (Continuation/Existence of State) class and the three unergative classes – all show a preference for HAVE, which is also the preferred choice of the Dutch control group. The correct tendency to select HAVE increases gradually from the peripheral unaccusative class to the core unergative class, with the exception of the CMP (Controlled Motional Process) class, where there is a drop in the selection of HAVE. The absolute numbers and percentages are given in Table 2 in Appendix VI.

Post hoc pairwise comparisons (Bonferroni) of auxiliary selection within the verb classes were conducted for the English test subjects collectively. The core unaccusative COL verb class differed significantly from the peripheral unaccusative CES class and all three unergative verb classes ($p < .001$ in all cases), but not with the intermediate unaccusative

COS class. Similarly, the COS class differed significantly from all other verb classes with the exception of the COL class ($p < .001$ in all cases). This contrast is not really surprising, as the verbs in these two unaccusative classes are the only ones to select BE in Dutch. However, the lack of differentiation between the COL and COS verbs contradicts the predictions based on the ASH. We return to this point in the next chapter. With regard to the other verbs, the only significant difference was between the peripheral unaccusative CES verbs and the core unergative CPN verbs ($p = .024$). Of the verbs that select HAVE in Dutch, these are the verbs that are farthest apart on the hierarchy. The pairwise comparisons for all English native speakers are given in Table 10 in Appendix VI.

The responses given by the twenty-six test subjects broken down by proficiency level are depicted in Figure 13. Here, too, the same tendencies can be seen. The most frequent selection of BE is found in the unaccusative COL and COS verb classes, and for both proficiency groups the difference between these two classes is very slight. In the other four verb classes, the selection of HAVE is lower in the peripheral unaccusative CES verb class and higher in the core unergative CNP verb class. For the medium proficiency group, the same gradient increase as in Figure 12 is seen, from peripheral unaccusative to core unergative. However, for the advanced proficiency group this gradience is not evident. This is probably due to a ceiling effect, as the results from this advanced group converge on the results from the Dutch native speakers. The absolute numbers and percentages for the two proficiency levels are given in Tables 3 and 4 in Appendix VI.

A repeated measures ANOVA on the data from the two English proficiency groups, with verb class as the within-subjects factor and proficiency as the between-subjects factor, revealed a highly significant main effect of verb class ($F_{(3,036,\ 120)} = 71,995$, $p < .001$). There was also a significant interaction between verb class and proficiency ($F_{(3,036,\ 120)} = 5,149$, $p < .005$).

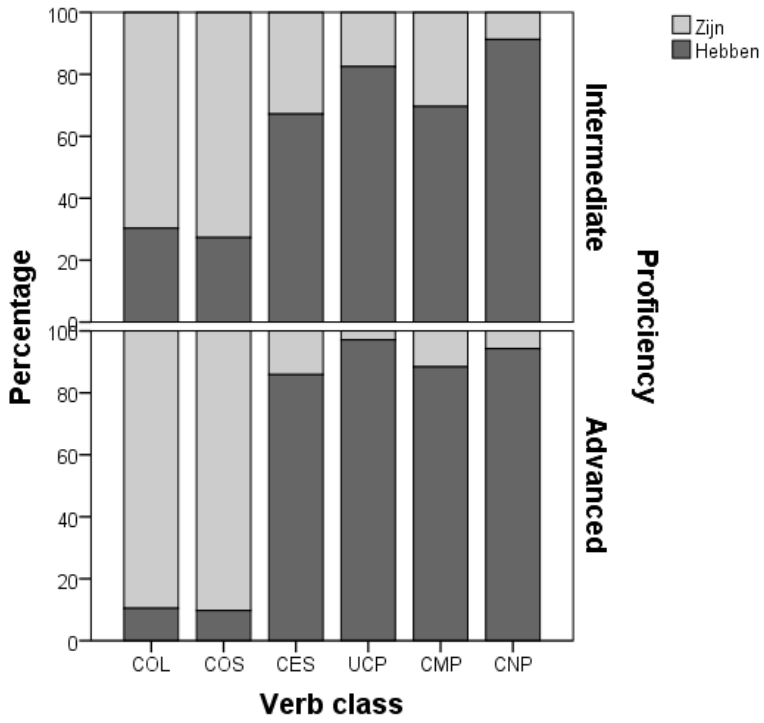

Fig. 13 Distribution of auxiliary selection per verb class by English native speakers, broken down by proficiency level

The line chart in Figure 14 compares the selection of BE in the six verb classes by all Dutch and English native speakers, demonstrating that the results from the advanced group converge on the results from the Dutch native speakers.

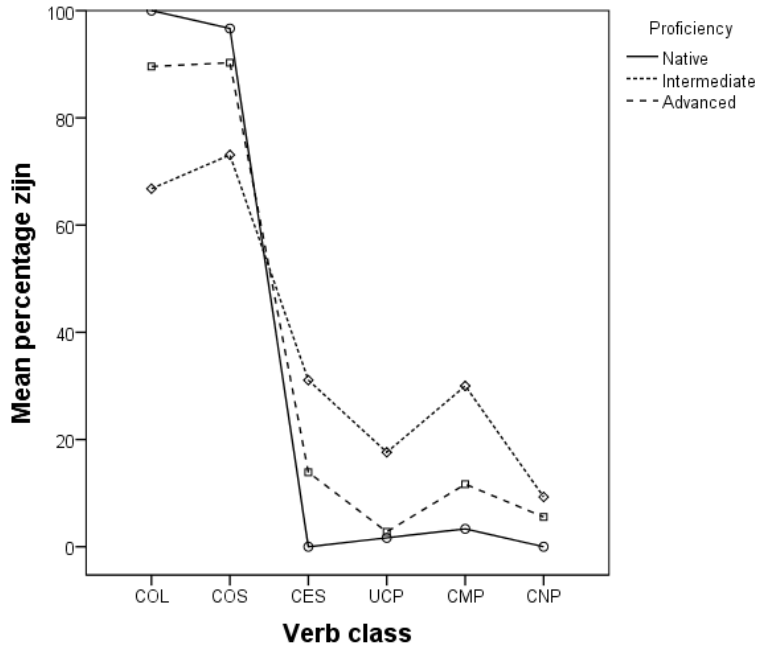

Fig. 14 Mean percentage of *zijn* selection according to verb class and proficiency level

In consideration of the near-native performance of the advanced proficiency group, post hoc pairwise comparisons (Bonferroni) of auxiliary selection within the verb classes were carried out on the medium proficiency group separately, to see whether any differences may have been obscured by a ceiling effect. Again, a significant main effect of verb class was found ($F_{(2,867,\ 65)}$ = 15,523, p < .001). The post hoc pairwise comparisons did not reveal any other relevant differences. In fact the differences were smaller than for the combined proficiency groups. The COS verbs still differed significantly from all other classes except the COL verbs, but to a lesser extent (COS-CES, p = .036; COS-UCP, p = .005; COS-CPM, p = .009, COS-CPN, p < 001). The COL verbs no longer differed significantly from the CES verbs (p = .051) or the CPM verbs (p = .105). These smaller differences may be due to the smaller number of cases (452 compared to 871 for the two groups combined). Furthermore, the less proficient Dutch speakers incorrectly select HAVE more often than the advanced speakers, thus reducing the discrepancy between the core and intermediate unaccusatives and the other four verb classes.

7.3 Correct responses

The responses given by the Dutch control group were used to establish a benchmark for correct auxiliary selection. The Dutch native speakers consistently selected BE for the core unaccusative (COL) verb class and the intermediate unaccusative (COS) verb class, HAVE for the peripheral unaccusative (CES) verb class, and HAVE for all three unergative verb classes. Taking these as the correct auxiliaries in Dutch, the responses from the English native speakers were then examined for correctness. Figure 15 shows the percentage of correct responses in each verb class given by all English test subjects combined. Here we see no gradient differentiation. The unergative verb classes UCP and CNP have the highest percentage of correct answers, whereas the other four classes differ only very slightly from one another. The absolute numbers and percentages can be found in Table 5 in Appendix VI.

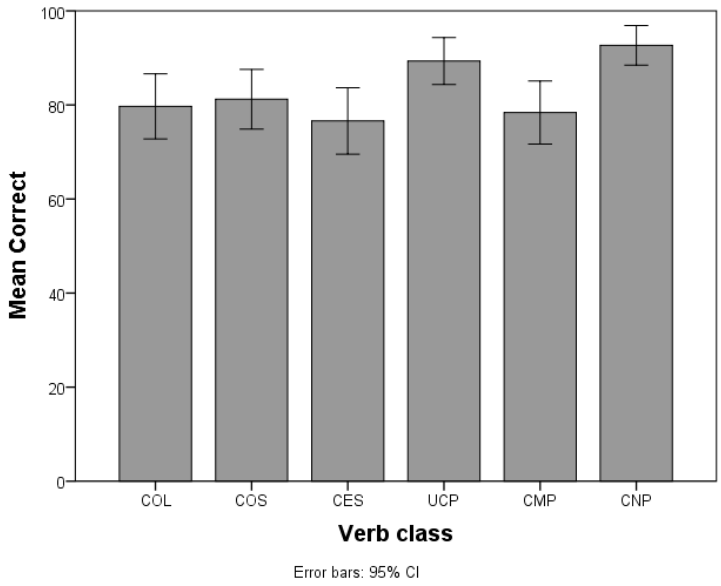

Fig. 15 Distribution of correct auxiliary use per verb class

A repeated measures ANOVA with verb class as the within-subjects factor and proficiency as the between-subjects factor revealed a significant main effect of verb class $F_{(3,417,\ 120)}$ = 3,639, $p < .05$). There was no significant interaction between verb class and proficiency ($F_{(3,417,\ 120)} < 1$, ins). Post hoc pairwise comparisons (Bonferroni) of correct responses within the verb classes revealed that the unergative verb classes UCP and CNP both differed significantly from the other four verb classes but not from each other. These verbs therefore seem to be the easiest to learn. This conclusion is discussed further in the next chapter.

Figure 16 shows the results of the two proficiency levels next to those of the Dutch native speakers for each verb class. This indicates that the advanced speakers are approaching native speaker proficiency. Post hoc multiple comparisons (Bonferroni) show no significant difference between the advanced proficiency group and the Dutch native speakers ($p = .285$). However, there is a highly significant difference between the medium proficiency group and the native speakers ($p < .001$) and also between the medium proficiency and advanced proficiency groups ($p = .002$). This seems to confirm the near-native attainment of the advanced group.

The difference between the two proficiency levels and the lack of a significant interaction between verb class and proficiency suggests that, contrary to the prediction, there is no learnability distinction between the verb classes. Figure 17 depicts the relations between the

two English proficiency groups and the Dutch control group, and indicates that the differences between the medium and advanced proficiency level are additive.

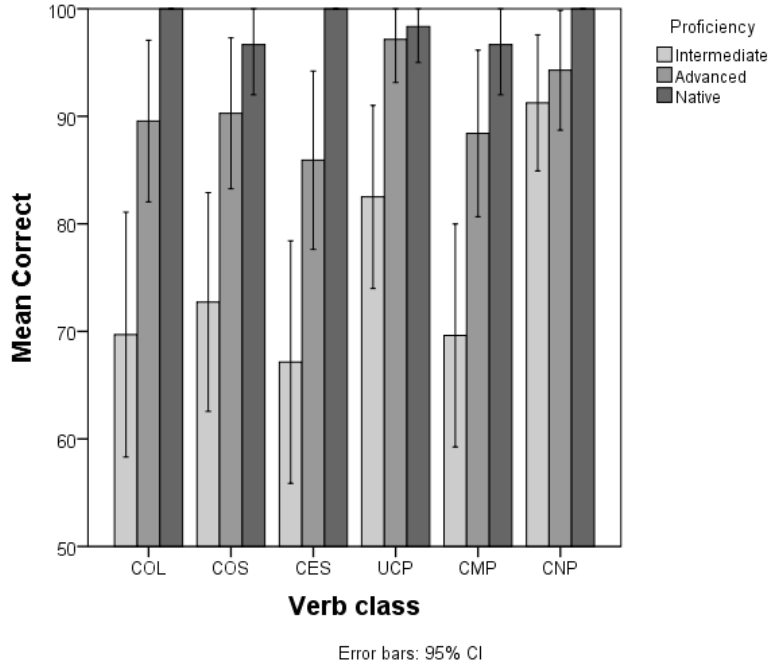

Fig. 16 Distribution of correct auxiliary use per verb class, broken down by proficiency level

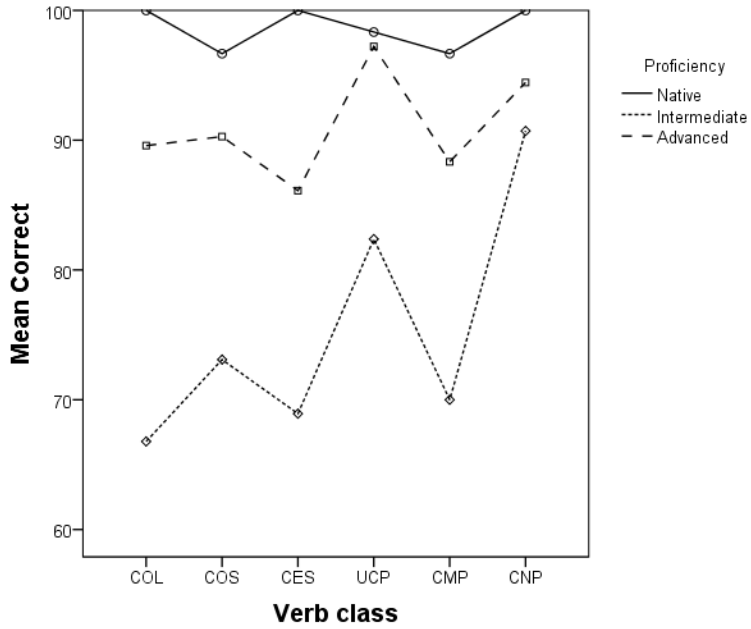

Fig. 17 Correct auxiliary use per verb class for Dutch and English native speakers

In chapter 8 these results and their implications for the research questions are discussed in more detail.

8. Conclusion

The research question under investigation is whether the perfective auxiliary selection of English learners of Dutch as a second language reflects the Auxiliary Selection Hierarchy, with gradient preferences for BE or HAVE respectively depending on the classification of the lexical verb along the hierarchy. The general prediction that there will be a gradual transition from BE to HAVE along the verb classes of the hierarchy, from the core unaccusatives to the core unergatives respectively, is partly borne out by the findings of the experiment. Figure 12 shows a trend for the increasing use of HAVE in accordance with the hierarchy, although, as predicted, the ambiguous Controlled Motional Process (CMP) verb class deviates from this pattern, with a more frequent selection of BE for these verbs in relation to the other unergative verb classes (UCP and CNP). We return to this in the discussion on the unergative verbs below. Answering the secondary research question of whether the language learners will be more inclined to use the auxiliary BE for unaccusatives across the board and the auxiliary HAVE for unergatives across the board, we can conclude that there is a highly significant effect of verb type on auxiliary selection, as depicted in Figure 11 in section 7.2.

Beyond these general tendencies we now take a closer look at the specific preferences for the two verb types of unaccusative and unergative. The first prediction for the unaccusative verbs is that selection of the auxiliary BE will follow the ASH. However, contrary to this prediction there is no distinction between the core Change of Location verb class and the intermediate Change of State verb class, although in line with the prediction we do see the least frequent selection of BE in the peripheral CES verb class (Continuation/Existence of State). As shown in Figure 12, the selection of the auxiliary BE is almost the same for the COL and COS groups.

The second prediction is that fewer auxiliary selection errors will be made for the core unaccusative verbs than for the intermediate unaccusative verbs. However, again the findings contradict this prediction. As we see in Figure 15, there is very little difference between the core Change of Location and the intermediate Change of State verb classes. There is also no evidence of the predicted transfer effect in the peripheral CES verb class. Despite the fact that all the tested verbs in this combined stative verb class take HAVE in

both English and Dutch, the incorrect auxiliary BE was chosen in 23% of the responses. This is a conspicuous result considering that the English native speakers will not have heard this combination of perfective auxiliary and participle in the ambient language, as this is ungrammatical in Dutch and was not witnessed at all in the Dutch native speaker control group. Instead of being influenced by L1 transfer and resorting to the default option of HAVE in line with the corresponding feature in their native language, the test subjects were presumably more strongly influenced by the unaccusative character of the verbs, leading to the incorrect selection of BE in almost a quarter of the responses. They therefore seem to be overgeneralising the Dutch language feature of the selection of BE as perfective auxiliary for unaccusative verbs. There seems to be no other logical explanation.

The third prediction that the subjects will learn the correct form for the core unaccusative (COL) verbs before the intermediate unaccusative (COS) verbs is not endorsed either. If this were the case we would expect to see a higher incidence of BE selection for the COL verbs than for the COS verbs, and we may also expect to see a greater discrepancy between these two verb classes for the less advanced (intermediate) learners than for the advanced learners. However, the findings contradict these expectations. As already stated, and as shown in Figure 12, we see that the two verb classes produce very similar results. Figure 13 compares the responses of the two proficiency groups. In both groups we see very similar results for the two verb classes COL and COS. The advanced group has a higher incidence of BE selection than the less advanced group for both verb classes, but the effect is additive and the predicted discrepancy based on the premise that the COL verbs would be easier to learn is not witnessed. Figure 14 further shows that the advanced group is approaching native speaker proficiency levels.

The lack of differentiation between the unaccusative verb classes COL and COS calls for closer scrutiny. Looking at the responses for the individual verbs, we see that in the COL verb class there is a disproportionately low mean selection of BE for the verb *emigreren* (emigrate) (52% compared with an overall mean of 80% for this verb class), and in the COS verb class there is a disproportionately high mean selection of BE for the COS verb *sneuvelen* (perish) (100% compared with an overall mean of 81% for this verb class). If we compare the findings for the remaining five verbs in each class, we see a slight difference between the

two verb classes in accordance with the ASH, as depicted in Figure 18. However, post hoc pairwise comparisons (Bonferroni) of auxiliary selection within the verb classes after removal of the two deviant verbs reveal that this is not significant (p = 1.000). The pairwise comparisons after removal of the deviant verbs are given in Table 11 in Appendix VI.

Figure 19 further shows that the discrepancy between the two proficiency levels is very slightly larger for the COL verbs than for the COS verbs, in accordance with the prediction, although this difference is negligible. The absolute numbers and percentages can be found in Tables 6, 7 and 8 in Appendix VI.

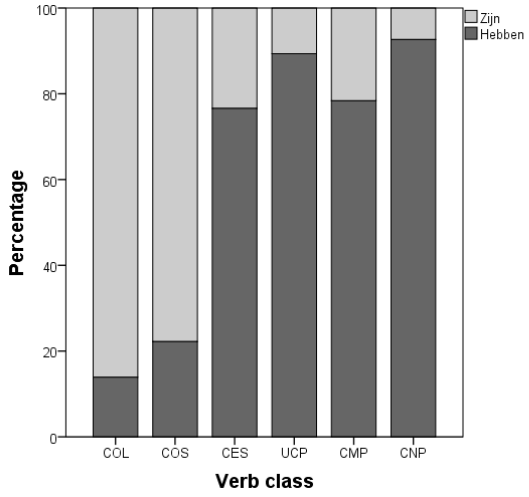

Fig. 18 Distribution of auxiliary selection for all English natives (after removal of deviant verbs)

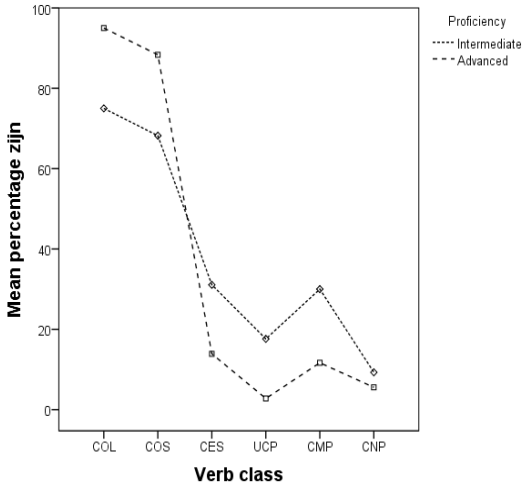

Fig. 19 *Zijn* selection according to proficiency level (after removal of deviant verbs)

Of course, the verbs in each class were chosen randomly, the only consideration being that word frequency and word length should be balanced between classes, and it is clearly not the intention to simply disregard verbs that disprove the theory. However, the substantial discrepancy displayed by these two verbs within their respective classes does raise certain questions. With regard to the COL verb *emigreren*, although *emigreren* seems to be an intrinsic Change of Location verb – the very nature of the action encoded in the verb relates to the relocation of the verb subject from one place to another – it could perhaps in usage be considered a Change of State verb in the sense that the most salient aspect of emigration may not be the location itself but the change in environment of the experiencer subject. The element of volition may also play a role. Emigration could be seen as a voluntary act, whereas volition is generally considered to be a feature of unergative verbs, as stated in

section 3.2. On the other hand, if the verb is describing the actual physical process of emigration, this verb could conceivably be classified as a Controlled Motional Process verb, similar to the motion verbs such as *run* and *walk* that have ambiguous status in Dutch as described in section 4.3. A similar argument could be made for other core unaccusative verbs such as *escape* and *flee* if the semantic import of the verb activity within the sentence context could be considered to affect the interpretation. From this viewpoint the classification of such verbs would therefore depend on both the context of the sentence and the perspective taken on the activity. Taking into consideration the exact context of the test sentence containing the verb *emigreren*, shown below in sentence 28, it could be said to have an atelic quality in that the emigration took place more than once over a period of time.

28. Al vind ik verhuizen niet zo leuk, ik (ben) de laatste tien jaar al drie keer geëmigreerd.
"Even though I don't like moving house, I have emigrated three times in the last ten years."

The 36 test sentences were designed to contain a varied range of telic and atelic predicates and animate and inanimate subjects. However, it was difficult to do this systematically, partly due to the small number of available verbs in each class and partly due to the attempt to ensure that the sentences sounded reasonably natural in both Dutch and English, even though the present perfect is used in a slightly different way in the two languages. Furthermore, *emigreren* is the lowest frequency verb in the core unaccusative COL verb class. It is therefore possible that the combination of a low frequency verb with a non-telic modifier influenced the subjects in their auxiliary choice. However, this did not seem to affect the Dutch native speakers in any way, as in spite of these considerations the Dutch control group was unanimous in its choice of BE as auxiliary. Clearly, having learned the auxiliary-participle pairing *zijn + geëmigreerd* the Dutch native speakers are not inclined to modify the auxiliary depending on the context, as they do with the prototypical motion verbs.

Sentence 29 gives the test sentence containing the COS verb *sneuvelen*, which produced a disproportionately high selection of the auxiliary BE.

29. Vandaag (is) de bril van meester Hans gesneuveld door een ongelukje met een voetbal.
"Today the glasses of teacher Hans perished due to an accident with a football."

As shown in Table 9 in Appendix VI, there are only three verbs for which the English native speakers have a 100% score, and of these three verbs *sneuvelen* is the only one that takes the auxiliary BE. This word literally means *perish* or *die* in Dutch, and it has no transitive alternant - in other words, it is not possible to *sneuvel* something or someone. In the test sentence this verb is used in combination with a pair of glasses, an inanimate object. It is therefore being used in an idiomatic sense, as glasses cannot die. This is a very common use of this verb in Dutch, although if this sentence were translated informally into English, *gesneuveld* may perhaps be more naturally translated with the word *broken*. *Break* is a transitivity alternation verb in both English and Dutch (*breken*). It is therefore possible that some of the test subjects were treating this word as a transitive verb in a passive construction with the auxiliary BE, in line with an English passive construction such as *the glasses were broken.*

Moving to the unergative verbs, the first prediction that there will be little variation between the three unergative verb classes, with a strong preference for HAVE in all cases, is not fully borne out by the findings. Figure 12 in section 7.2 shows that there is a discrepancy between the peripheral unergative verb class UCP and the core unergative verb class CPN. Although the difference is not significant, this gradient effect is in accordance with the ASH and the predicted transfer effect is not witnessed. Again, the question arises as to why this figure should be so high for the peripheral unergative UCP verbs, as the English test subjects will not have heard this ungrammatical combination of perfective auxiliary and participle from their surroundings, and any L1 transfer effect would lead them to select the default option HAVE. However, as with the peripheral unaccusatives, the unergative character of these peripheral verbs seems to be less robust than that of the core unergative verbs. In Figure 13

we see that this difference comes entirely from the less advanced (intermediate) Dutch speakers, with 18% of BE responses for the UCP verb class and 9% of BE responses for the CPN verb class. The advanced group chose HAVE in almost all cases, having mastered both verb classes to near-native proficiency levels. The absolute numbers and percentages can be found in Tables 2, 3 and 4 in Appendix VI.

The second prediction for the unergative verbs is that the intermediate unergative Controlled Motional Process (CMP) verbs would deviate from the ASH, showing a greater tendency to select BE rather than HAVE. As seen in Figure 12, this prediction is confirmed by the data, with 22% of all responses for this verb class selecting the auxiliary BE. As expected, the ambiguous character of the motion verbs seems to have an influence on auxiliary selection. Depending on the context, these verbs can select either HAVE or BE in Dutch. Dutch second language learners hearing motion verbs used in combination with the auxiliary BE may conclude that this is the correct form in all situations, or they may not be able to recognise the subtle aspectual distinctions between the different contexts. Considering that only five of the subjects had followed Dutch classes at an intermediate or advanced level, most participants would have to rely primarily on implicit learning methods, based on input and statistical learning. Furthermore, some of the subjects with advanced knowledge of German may have been influenced by the behaviour of the motion verbs in German. In German, locomotion rather than telicity is the main trigger for the selection of BE as auxiliary (Randall et al., 2004), and these motion verbs, as well as other verbs used in a context in which a change of position is indicated, tend to select BE in German (Keller & Sorace, 2003). A reasonable assumption is therefore that the test subjects have learned the auxiliary BE for these verbs but have not learned the relevant subtleties of meaning that affect its use. However, this does seem to be learnable, as the percentage of BE responses for this verb class drops from 30% for the less proficient group to 12% for the advanced group. The ambiguity of these verbs is also supported by the fact that 3% of the responses from the Dutch native speakers also select BE for this verb class.

In conclusion, besides the highly significant unaccusativity effect, a slight gradient effect reflecting the verb classes of the ASH can be seen. The differences are not significant, but there is a clear tendency to select respectively BE or HAVE more frequently for the

unaccusative and unergative core categories than for the unaccusative and unergative peripheral categories. For the unaccusatives, the intermediate COS verb class is indistinguishable from the core COL verb class, but when the outlying verbs with disproportionate responses in each class are removed, a small, non-significant difference can be perceived. For the unergatives, the results of the intermediate CMP verb class are affected by the ambiguous nature of these motion verbs in Dutch. The tendency to incorrectly select BE for both peripheral categories is notable, as there would be no evidence for this in the input and any L1 transfer effect would incite the use of HAVE. The only explanation seems to be a difference in the robustness of the unaccusative or unergative character of these verbs in accordance with the gradience of the ASH. No learnability effect between verb classes was found. The advanced learners outperformed the less advanced learners to a similar extent in all verb classes.

The results of the experiment therefore seem to support the notion of gradience in accordance with a selectional hierarchy as proposed by Sorace (1993, 2000, 2004). Even though no significant differences between verb classes were found, the discrepancy between the two intransitive verb types and also the differences between the core and peripheral categories within the two verb types display a variation that can be captured by the ASH.

9. Discussion

The clearest and perhaps most notable outcome of the experiment described in this paper is the tendency among English native speakers to select BE for unaccusatives and HAVE for unergatives. Even though the perfective auxiliary BE is not the correct auxiliary for peripheral unaccusative verbs in Dutch (with a few exceptions, as stated), some learners seem to treat this as the default option. As this ungrammatical option is not present in the primary linguistic data, a reasonable conclusion is that the English native speakers have learned that some verbs take BE as the perfective auxiliary and have overgeneralised this to all unaccusative verbs. The syntactic manifestations of unaccusativity in English - the diagnostics discussed in section 3.2 - are far less salient than the perfective auxiliary distinction in many Germanic and Romance languages, or the Dutch impersonal passive construction. There is consequently little conspicuous positive evidence for unaccusativity in English. However, in spite of this English natives must have some underlying sense of a distinction between unaccusative and unergative verbs, otherwise this tendency for incorrect auxiliary selection would be difficult to explain. This seems to provide strong support for the notion that the inherent properties of unaccusative verbs constitute part of Universal Grammar. The test subjects have an innate knowledge of unaccusativity, they have learned that BE is the perfective auxiliary of choice for unaccusative verbs in Dutch, but they have not yet learned the language-specific application rules. As pointed out by Sorace & Shomura (2001), the main difficulty lies in working out how a language links the complex interaction of lexical-semantic properties such as telicity, agentivity and volition to the binary syntactic unaccusative-unergative division.

The results of the experiment suggest that auxiliary selection by Dutch second language learners is influenced by the context of the lexical verb within the sentence and also by semantic and aspectual factors such as telicity and animacy. Other factors such as L1 transfer and language-specific idiomatic use may also play a role. However, designing an experiment in which variables including animacy and telicity are factored in systematically would require a prohibitively large number of sentences. Perhaps this could be achieved by focusing on fewer verb classes, such as just the unaccusatives. This may enable finer distinctions to be made.

For the peripheral unaccusative CES verbs, it is difficult to evaluate the relative impact of the transfer effect on the responses, as almost all of these verbs select HAVE in both English and Dutch. In other words, whereas the incorrect choice of BE must be influenced by the unaccusative character of the verb, choosing HAVE may be caused by a transfer effect, or it may simply be that the test subject has learned the correct verb form. The only way to tease apart the transfer effect would be to make comparisons between verbs selecting HAVE and the few verbs in the CES category that select BE in Dutch, such as *blijven* (stay) and *zijn* (be). However, the high frequency of these verbs would make it difficult to make sound comparisons. Furthermore, the verb *to be* is clearly exceptional in its usage, as it also functions as a copular and as a passive auxiliary, and is marked when appearing on its own (e.g. *I am*).

As discussed in the previous chapter, the results of the experiment offer no evidence of enhanced learnability for certain verbs, as the advanced learners outperformed the intermediate learners in all verb classes. However, some of these advanced learners displayed near-native competence, making it difficult to distinguish between verb classes due to a ceiling effect. On the other hand, the test proved too difficult for test subjects with a limited knowledge of Dutch, preventing their results from being used in the analysis. Not knowing the proficiency of the test subjects in advance makes it very difficult to design a suitable test, and it would be almost impossible to design a test that suits all proficiency levels. The chosen sentence completion test seems appropriate for the intermediate test subjects, but it may have been too easy for the advanced Dutch speakers. For this group, an acceptability judgement test may have revealed finer distinctions between the verb classes. The beginners, on the other hand, would have been able to provide a greater percentage of usable responses if they had been presented with two auxiliaries and asked to choose. Responses from elementary language learners may offer a better insight into whether the COL verbs are learned first.

Returning to the question of the variable flexibility of verbs discussed in section 3.1, the different treatment of a verb such as *emigreren* by English and Dutch native speakers is notable. The Dutch native speakers are unanimous in their auxiliary choice and seem to have stored the combination of auxiliary and participle as a cohesive unit in their mental lexicon.

Regardless of the context they are not inclined to vary the auxiliary, suggesting that the lexical representation is fixed. With typical Dutch motion verbs the representation is not fixed. The addition of a telic modifier - which presumably takes place as the sentence is being formed, after the word has left the lexicon - determines the syntactic structure into which the verb is slotted and consequently whether it has an internal or external argument and whether it takes BE or HAVE. Sorace (2004, 2006) suggests that verbs with rigid behaviour seem to fit the projectionist (lexicalist) approach, in which the lexical entry contains information that specifies how the verb can be used, whereas the flexible verbs pattern according to the constructionist approach, in which the lexical entry is bare and the verb can be slotted into various syntactic configurations. However, the English second language learners seem to have incomplete representations. They construct the perfect form from its constituent parts during processing, using unaccusativity as an aid for auxiliary selection. They are deceived by the peripheral unaccusative verbs, tending to overgeneralise the use of BE, although the more advanced learners have overcome this overgeneralisation (similar to the U-shaped overgeneralisation seen in children: Pinker, 1994). As native speakers have more detailed representations, less complex processing is required, thus facilitating spontaneous speech and minimising errors. Second language speakers, on the other hand, have less determinate, less detailed representations, calling for more complex processing. Perhaps the need for more complex processing due to incomplete or indeterminate representations goes some way towards explaining the problems of second language acquisition.

Finally, the findings of this experiment raise a number of questions that may be worth investigating in future research. For example, this experiment examines the responses given by English native speakers. In order to gain a better understanding of how second language learners acquire a language-specific property such as the perfective auxiliary, it would be interesting to carry out the same test on a range of Dutch language learners with various native languages in which the perfective is created differently. Spanish, like English, uses only the auxiliary HAVE in perfect constructions and lacks robust syntactic evidence for unaccusativity, although it has been suggested that word order differences exist between unaccusative and unergative verbs (Parafita Couto et al., 2015). It would be interesting to compare responses from Spanish native speakers with those from English native speakers to

check for relevant similarities or differences. We have already mentioned the possibility of conducting a similar experiment among native speakers of a Slavic language that uses only the auxiliary BE, such as Croatian (Knežević & Brdar, 2012). The task of these Slavic speakers would be the opposite of the English native speakers - they would have to learn a different auxiliary for the unergative verbs rather than the unaccusative verbs. This may consequently reveal greater differentiation between the unergative verb classes, as L1 transfer between English and Dutch may neutralise possible effects with unergative verbs. In Russian, perfective aspect is generally expressed through a prefix (Janda & Lyashevskaya, 2012). Although Russian has no auxiliary selection to distinguish between verb types, other syntactic manifestations of unaccusativity in Russian have been investigated and described in detail (e.g. Schoorlemmer, 2004). It would be interesting to see whether any form of split intransitivity hierarchy can be applied to these diagnostics and to compare this with the results from a sentence completion test carried out among Russian learners of Dutch. There are many other languages in which the perfect aspect is expressed not with an auxiliary and participle but in a completely different way (Dahl, 1985). For example, in Welsh the perfect aspect is expressed with the aspect marker *wedi,* which means “after” (Borsley et al., 2007), and in Japanese the perfect is formed by using a suffix (Nishiyama, 2006). How would Dutch language learners with Welsh or Japanese as their L1 differentiate between the verb classes within the ASH? Comparisons of the treatment of the Dutch perfect construction by native speakers of typologically different languages could offer revealing insights into the phenomenon of unaccusativity as well as into second language acquisition in general.

References

Abrahamsson, N. & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in second language acquisition,* 30, pp. 481-509.

Alexiadou, A., Anagnostopoulou, E. & Everaert, M. (2004). *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface*. Oxford : Oxford University Press.

Allan, R., Holmes, P. & Lundskær-Nielsen, T. (2000). *Danish. An essential grammar.* Routledge: Oxon.

Algemene Nederlandse Spraakkunst (2013). Retrieved 27-03-2015 from http://ans.ruhosting.nl

Arad, M. (1996). *A Minimalist view of the Syntax-lexical semantics Interface*. UCL Working Papers in Linguistics, No. 8, pp. 215-242.

Bard, E. G., Frenck-Mestre, C., & Sorace, A. (2010). Processing auxiliary selection with Italian intransitive verbs. *Linguistics,* 48 (2), pp. 325-361.

Belien, M. (2012). Dutch manner of motion verbs: Disentangling auxiliary choice, telicity and syntactic function. *Cognitive Linguistics*, 23 (10, pp. 1–26. Retrieved 2-10-2014 from http://dx.doi.org/10.1515/cog-2012-0001.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S.Gass & J.Schachter (eds.), *Linguistic perspectives on second language acquisition,* pp. 41-68. Cambridge: Cambridge University Press.

Borer, H. (2004). The grammar machine. In Alexiadou, A., Anagnostopoulou, E. & Everaert, M. (Eds.), *The Unaccusativity Puzzle: Explorations of the Syntax- Lexicon Interface,* pp. 288-331. Oxford: Oxford University Press.

Borsley, R., Tallerman, M, & Willis, D. (2007). *The Syntax of Welsh (Cambridge Syntax Guides).* Cambridge: Cambridge University Press.

Burzio, L. (1986). *Italian syntax: a government-binding approach*. Dordrecht: Reidel.

Carnie, A. (2007). *Syntax: A generative introduction*. Malden, MA: Wiley-Blackwell.

Cheng, L. L. (2009). Wh-in-situ, from the 1980s to Now. *Language and Linguistics Compass, 3*, pp. 767-791.

Chomsky, N. (1965). Aspects of the Theory of Syntax, Cambridge, Massachusetts: MIT Press.

Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds.) *Reading in English Transformational Grammar*, pp. 184-221. Waltham: Ginn.

Chomsky, N. (1973). Conditions on transformations. In Stephen Anderson and Paul Kiparsky (eds.), *A Festschrift for Morris Halle,* pp. 232-286. New York: Holt, Rinehart, and Winston.

Chomsky, N. (1980). Rules and Representations. New York: Columbia University Press.

Chomsky, N. (1981). Lectures on government and binding. Dordrecht: Foris Publications.

Chomsky, N. (1981). Lectures on government and binding. Dordrecht: Foris Publications.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass: MIT Press.

CEFR (2001). Common European Framework of Reference for Languages. Retrieved 18-03-2015 from http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Key_reference/CEFR_EN.pdf

Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language, 63*, pp. 544−573.

Crain, S. & Pietroski, P. (2002). Why language acquisition is a snap. *The Linguistic Review,* 19, pp. 163-183.

Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics (6th edition*). MA: Blackwell Publishers.

Dahl, Ö. (1985). *Tense and Aspect Systems*. Oxford: Basil Blackwell.

Davies, W.D. & Dubinsky, S. (2004). *The Grammar of Raising and Control: A Course in Syntactic Argumentation*. Malden, MA: Blackwell Publishing Ltd.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, pp. 499–533.

Ellis, N. (1994). Implicit and explicit language learning - An overview. In N.C. Ellis (Ed.), *Implicit and explicit learning of languages,* pp. 1-32. London: Academic Press.

Epstein, S., Flynn, S. & Martohardjono, G. (1996). Second language acquisition: theoretical and experimental issues in contemporary research. *Brain and Behavioral Sciences, 19*, pp. 677-758.

Folli, R. and H. Harley (2005) Consuming results in Italian and English: flavors of v. In P. Kempchinsky and R. Slabakova (eds.) *Aspectual inquiries*, pp. 95-120. Dordrecht: Springer.

Friedmann, N., Taranto, G., Shapiro, L. & Swinney, D. (2008). The leaf fell (the leaf): the online processing of unaccusatives, *Linguistic Inquiry,* 39, pp. 355-377.

Hannay, M. & Schrama, M. H. M. (1988). *Van Dale handwoordenboek Engels-Nederlands (3rd edition)*. Utrecht: Van Dale Lexicografie.

Hoekstra, T. (1999). Auxiliary selection in Dutch. *Natural Language and Linguistic Theory,* 17, pp. 67-84.

Hornstein, N., Nunes, J. and K. Grohmann (2006). *Understanding Minimalism*. Cambridge: Cambridge University Press.

Hout, van, A. (2004). Unaccusativity as telicity checking. In Alexiadou, A., Anagnostopoulou, E. & Everaert, M. (Eds.), *The Unaccusativity Puzzle: Explorations of the Syntax- Lexicon Interface*. Oxford University Press: Oxford.

Jackendoff, R. (1977). *X-bar-Syntax: A Study of Phrase Structure*. Linguistic Inquiry Monograph 2. Cambridge, MA: MIT Press.

Janda, L. A. & Lyashevskaya, O. (2012). Semantic Profiles of Five Russian Prefixes: po-, s-, za-, na-, pro-. *Journal of Slavic Linguistics,* 20(2).

Johnson, J. & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology,* 21, pp. 60-99.

Keller, F. & Sorace, A. (2003). Gradient auxiliary selection and impersonal passivization in German: an experimental investigation. *Journal of Linguistics* , 39, pp. 57-108.

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

Knežević, B. & Brdar, I. (2012). Unaccusatives and unergatives: Evidence from Croatian. *Folia Linguistica,* 48 (1), pp. 61–84.

Koenen, M.J. & Drewes, J.B. (1982). *Verklarend Handwoordenboek der Nederlandse Taal*, (*26th edition).* Groningen: Wolters-Noordhoff.

Kuroda S. Y. (1988). *Whether We Agree or Not: A Comparative Syntax of English and Japanese*. In W. J. Poser (Ed.), *Papers from the Second International Workshop on Japanese Syntax,* pp. 103-143. Stanford, CA: CSLI.

Legate, J. (2003). Some interface properties of the phase. *Linguistic Inquiry, 34*, pp. 506-516**.**

Legendre, G. (2007). On the typology of auxiliary selection. *Lingua,* 117 (9), pp. 1522-1540.

Lenneberg, E. (1967). The biological foundations of language. New York: Wiley.

Levin, B. & Rappaport Hovav, M. (1995). *Unaccusativity: At the syntax-semantics interface.*

Cambridge, MA: MIT Press.
Lieber, R. & Baayen, H. (1997). A Semantic Principle of Auxiliary Selection. *Natural Language and Linguistic Theory,* 15, pp. 789-845.
Marantz, A. (1984). *On the nature of grammatical relations*. Cambridge: MIT Press.
McCloskey, J. (1993). A crude test for unaccusativity in English. In Pullum, G. & Potsdam, E. (Eds.), *Syntax at Santa Cruz* 2, pp. 21-24. Linguistic Research Center, University of California, Santa Cruz.
Montrul, S. (2000). Transitivity alternations in L2 acquisition: Toward a modular view of transfer. Cambridge University Press: Cambridge.
Nishiyama, A. (2006). The Meaning and Interpretations of the Japanese Aspect Marker *tei*. *Journal of Semantics*, 23 (2), pp. 185-216.
Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge, Mass: Cambridge University Press.
Parafita Couto, M.C., Mueller Gathercole, V.C. & Stadthagen-González, H. (2015). Interface strategies in monolingual and end-state L2 Spanish grammars are not that different. *Frontiers in Psychology: Language Science* 5: 1525.
Perlmutter, D. (1978). Impersonal passives and the unaccusative hypothesis. In *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*, Berkeley, pp. 157–189.
Pinker, S. (1994). *The Language Instinct*. New York: William Morrow.
Randall, J., van Hout, A., Weissenborn, J. & Baayen, H. (2004). Acquiring unaccusativity: a cross-linguistic look. In Alexiadou, A., Anagnostopoulou, E. & Everaert, M. (Eds.), *The Unaccusativity Puzzle: Explorations of the Syntax- Lexicon Interface*. Oxford University Press: Oxford.
Renkema, J. (2002). *Schrijfwijzer.* The Hague: SDU.
Rydén, M. & Brorström, S. (1987). The Be/Have variation with intransitives in English, with special reference to the late modern period. Stockholm: Almqvist & Wixell International.
Schoorlemmer, M. (2004). Syntactic Unaccusativity in Russian. In Alexiadou, A., Anagnostopoulou, E. & Everaert, M. (Eds.), *The Unaccusativity Puzzle: Explorations of the Syntax- Lexicon Interface*. Oxford University Press: Oxford.
Slavkov, N. (2009). The acquisition of complex wh-questions in the L2 English of Canadian French and Bulgarian speakers: Medial wh-constructions, inversion phenomena, and avoidance strategies. *Doctoral dissertation, University of Ottawa.*
Sorace, A. (1993). Incomplete vs. divergent representations of unaccusativity in non-native grammars of Italian. *Second Language Research,* 9, pp. 22-47.
Sorace, A. (2000). Gradients in Auxiliary Selection with Intransitive Verbs. *Language*, 76 (4), pp. 859-890.
Sorace, A. (2004). Gradience at the Lexicon–Syntax Interface: Evidence from Auxiliary Selection and Implications for Unaccusativity. In Alexiadou, A., Anagnostopoulou, E. & Everaert, M. (Eds.), *The Unaccusativity Puzzle: Explorations of the Syntax- Lexicon Interface*. Oxford University Press: Oxford.
Sorace, A. (2006). Gradedness and optionality in mature and developing grammars. In Fanselow G**,** Fery, C., Schlesewsky, M., & Vogel, R. (Eds.), *Gradience in grammars: Generative perspectives*, pp. 106-123**.** Oxford**:** Oxford University Press.
Sorace, A. & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research,* 22, pp. 339-368.

Sorace, A. & Keller, F. (2005). Gradience in Linguistic Data. *Lingua*, 115 (11), pp. 1497–1524.

Sorace, A. & Shomura, Y. (2001). Lexical constraints on the acquisition of split intransitivity. *Studies in Second Language Acquisition,* 23, pp. 247-278.

Subtlex (2015). Retrieved 28-02-2015 from http://crr.ugent.be/isubtlex.

Thornton, R. (1990). Adventures in long distance moving: The acquisition of complex wh-questions. Doctoral dissertation, University of Connecticut, Storrs.

Thornton, R. (2008). Why Continuity. *Natural Language & Linguistic Theory,* 26 (1), pp. 107-146.

Van Kampen, J. (1997). *First steps in wh-movement*. Delft: Eburon.

Van Valin, R. (1990). Semantic parameters of split intransitivity. *Language,* 66, pp. 221–260.

Velupillai, V. (2012). *An introduction to linguistic typology*. Amsterdam: John Benjamins.

White, L. & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research, 12*, pp. 233-265.

Zaenen, A. (1988). Unaccusative Verbs in Dutch and the Syntax-Semantics Interface. CSLI Report 123, Stanford, CA.

Zaenen, A. (2006). Unaccusativity. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, pp. 217-224. Cambridge University Press: Cambridge.

Appendix I Test sentences grouped according to verb class

The lexical verbs are given in capitals and the missing auxiliaries are indicated in brackets. The English translations are close approximations of the Dutch sentences. These translations are not literal as the present perfect is used in a slightly different way in Dutch than in English.

Unaccusative

Change of Location

DALEN (DESCEND)
Binnen enkele minuten (was) het toestel van 35,000 voet naar ongeveer 11,000 voet gedaald.
"Within several minutes the plane descended from 35,000 feet to approximately 11,000 feet."

VERTREKKEN (DEPART)
In de loop van vorige week (zijn) de soldaten met alle hulpgoederen naar het oorlogsgebied vertrokken.
"During the course of last week the soldiers departed for the war zone with all their equipment."

VLUCHTEN (FLEE)
Omdat de kinderen zich weer misdroegen (was) de geplaagde docent gistermiddag de klas uit gevlucht.
"Because the children were misbehaving again the badgered teacher fled from the classroom yesterday afternoon."

STIJGEN (RISE)
Toen ze uiteindelijk het vissersdorpje binnenreden (was) de zon al boven de horizon gestegen.
"When they finally drove into the fishing village the sun had already risen above the horizon."

EMIGREREN (EMIGRATE)
Al vind ik verhuizen niet zo leuk, ik (ben) de laatste tien jaar al drie keer geëmigreerd.
"Even though I don't like moving house, I have emigrated three times in the last ten years."

STRANDEN (+/-STALL)
Wegens het slechte weer (zijn) vijf vrachtwagens uit Litouwen op de Nederlandse snelweg gestrand.
"Due to the bad weather five lorries from Lithuania stalled on the Dutch motorway."

Change of state

SLAGEN (SUCCEED)
Na de derde poging (is) Lucas eindelijk voor zijn tentamens geschiedenis en aardrijkskunde geslaagd.
"After the third attempt Lucas finally succeeded in his history and geography exams."

SLINKEN (SHRINK)
In de loop van de middag (was) de bankrekening van Kevin behoorlijk geslonken.
"During the course of the afternoon Kevin's bank balance shrunk considerably."

FUSEREN (MERGE)
Omdat zij failliet dreigden te gaan (zijn) de twee bedrijven na een lang onderhandelingstraject gefuseerd.
"Because they were in danger of going bankrupt, the two companies merged after protracted negotiations."

GROEIEN (GROW)
In deze warme, natte zomer (zijn) de naaldbomen in het duingebied veel harder gegroeid dan normaal.
"During this hot, wet summer the conifers in the dune region have grown much faster than usual."

SNEUVELEN (PERISH)
Vandaag (is) de bril van meester Hans gesneuveld door een ongelukje met een voetbal.
"Today the glasses of teacher Hans perished due to an accident with a football."

ONTPLOFFEN (EXPLODE)
Al was de dreiging reëel, de bommen van de terroristen (zijn) uiteindelijk niet ontploft.

"Even though the threat was real, the terrorist bombs didn't explode in the end."

Continuation/existence of state

LIJKEN (RESEMBLE)
Tot nu toe (heeft) de discussie over de begroting vooral op een ordinaire ruzie geleken.
"Up to now the discussion about the budget has resembled a vulgar squabble."
LIGGEN (LIE)
Voordat Jan gisteren naar huis mocht (heeft) hij drie weken in het Rijnland ziekenhuis gelegen.
"Before Jan was allowed to go home yesterday he lay in the Rijnland hospital for three weeks."
OVERLEVEN (SURVIVE)
Wegens de langdurige sneeuwval (hebben) veel knaagdieren en insecten de winter niet overleefd.
"Due to the long period of snow many rodents and insects have not survived the winter."
DUREN (LAST)
Tot opluchting van de ministers (hebben) de slechte economische omstandigheden niet al te lang geduurd.
"To the relief of the ministers the unfavourable economic conditions have not lasted long."
BESTAAN (EXIST)
In tegenstelling tot de beweringen van sommige gelovigen (hebben) mensen en dinosauriërs nooit gelijktijdig bestaan.
"Contrary to the claims of some religious believers, humans and dinosaurs have never coexisted."
WONEN (LIVE)
Voordat Marie naar het buitenland ging (heeft) zij enkele jaren in een huurhuis in Enschede gewoond.
"Before Marie moved abroad she lived in a rented house in Enschede for several years."

Unergative

Uncontrolled process

HUILEN (CRY)
Sinds vanochtend (heeft) het zieke kind tot wanhoop van zijn moeder onophoudelijk gehuild.
"Since this morning the sick child has cried unceasingly to the despair of its mother."
BRANDEN (BURN)
Volgens de servicemonteur (heeft) het rode waarschuwingslicht op het instrumentenpaneel enige tijd gebrand.
"According to the service technician the red warning light on the instrument panel has been lit for some time."
PIEPEN (SQUEAL)
Voordat het in slaap viel (heeft) het kleine zwerfhondje in het asiel heel even gepiept.
Before it fell asleep the little stray dog in the animal rescue centre squealed briefly."
LEKKEN (LEAK)
Sinds de loodgieter is geweest (hebben) de oude waterleidingen in de badkamer niet meer gelekt.
"Since the visit of the plumber the old water pipes in the bathroom have no longer leaked."
BLOEDEN (BLEED)
Voordat de dokter de wond kon hechten (heeft) het been van Sjaak behoorlijk gebloed.
"Before the doctor could stitch the wound, Sjaak's leg bled profusely."
DROMEN (DREAM)
Na het kijken naar de enge film (hebben) Jack en zijn vrienden de hele nacht onrustig gedroomd.
"After watching the scary film Jack and his friends dreamed fitfully all night long."

Controlled motional process

SPRINGEN (JUMP)
Elke dag in de vakantie (heeft) het kleine meisje hiernaast op haar trampoline gesprongen.
"Every day in the holidays the little girl next door has been jumping on her trampoline."

VLIEGEN (FLY)
Wegens de dreiging van de separatisten (hebben) de vliegtuigen vorige maand op tien kilometer hoogte gevlogen.
"Due to the threat by the separatists the planes flew at a height of ten kilometres last month."

REIZEN (TRAVEL)
Voordat zij met haar vervolgopleiding begon (heeft) de vriendin van mijn oudste broer veel gereisd.
"Before continuing with her studies my oldest brother's girlfriend travelled extensively."

RIJDEN (DRIVE)
Wegens zijn nieuwe baan (heeft) Bart maandenlang tussen Groningen en Leiden heen en weer gereden.
"Because of his new job, Bart drove back and forth between Leiden and Groningen for many months."

ZWEMMEN (SWIM)
Op vakantie in Florida (heeft) mijn zus in een waterpark met dolfijnen gezwommen.
"On holiday in Florida my sister swam with dolphins in an aquatic park."

WANDELEN (WALK)
Aan het einde van de zomer (hebben) de pelgrims drie dagen lang door de Spaanse bergen gewandeld.
"At the end of het summer the pilgrims walked through the Spanish mountains for three days."

Controlled non-motional process

ZWAAIEN (WAVE)
Nadat haar zus was ingestapt (heeft) Maria zo lang mogelijk naar de vertrekkende trein gezwaaid.
"After her sister had alighted Maria waved at the departing train for as long as possible."

LUISTEREN (LISTEN)
Sinds de vierde klas (hebben) de drie lastige jongens niet meer naar de leraar geluisterd.
"Since the fourth class the three troublesome boys have not listened to the teacher."

LACHEN (LAUGH)
Tot opluchting van de jonge comédienne(heeft) het publiek heel hard om haar grappen gelachen.
To the relief of the young comedian the audience laughed uproariously at her jokes."

DANSEN (DANCE)
Opgezweept door de swingende muziek (hebben) de bruiloftgasten op elk liedje enthousiast gedanst.
"Stimulated by the music the wedding guests danced enthusiastically to every song."

KLAGEN (COMPLAIN)
In Londen (hebben) de socialisten over het gebrek aan zendtijd op de staatstelevisie geklaagd.
"In London the socialists complained about the lack of broadcasting time on the national television. "

SPELEN (PLAY)
Dankzij de inzet van Robben (heeft) het Nederlandse elftal na enkele slechte wedstrijden eindelijk eens goed gespeeld.
"Thanks to Robben's efforts the Dutch football team finally played well after a series of poor performances."

Appendix II Personal details questionnaire

Please fill in the following details before proceeding with the study. (The study will be anonymous, but these details are required for analysis purposes.) Write on the dotted line or tick the boxes.

Native language: ……. Gender: …….. Age: ……

Country of birth: ……

Age of arrival in the Netherlands: …….

Number of years of residence in the Netherlands: .....

Have you ever lived in another country besides your country of birth and the Netherlands?

No [ ]
Yes [ ]

If you have answered Yes, please specify below which country, at what age, and for how long.

Country 1. I lived in …….. from the age of …….. for …….. years
Country 2. I lived in …….. from the age of …….. for …….. years
Country 3. I lived in …….. from the age of …….. for …….. years

Do you know any other languages besides English and Dutch?

French:	Basic knowledge	[ ]
	Advanced	[ ]
	Fluent	[ ]
German:	Basic knowledge	[ ]
	Advanced	[ ]
	Fluent	[ ]
Other (1): ……..	Basic knowledge	[ ]
	Advanced	[ ]
	Fluent	[ ]
Other (2): ……..	Basic knowledge	[ ]
	Advanced	[ ]
	Fluent	[ ]

Highest level of education:	Secondary school	[ ]
	University	[ ]
	Postgraduate	[ ]

Have you ever taken Dutch classes or received any type of instruction in Dutch?

No [ ]
Yes [ ]

If you have answered Yes, please specify below the type of classes/instruction received and the period of time over which the lessons were followed. More than one answer is possible.

Classes with teacher, beginner's course	[ ]	for ------- months/years
Classes with teacher, intermediate course	[ ]	for ------- months/years
Classes with teacher, advanced course	[ ]	for ------- months/years
Self-learning (books/Internet), basic level	[ ]	for ------- months/years
Self-learning, intermediate level	[ ]	for ------- months/years
Self-learning, advanced level	[ ]	for ------- months/years

Did you have any knowledge of the Dutch language before moving to the Netherlands?

No [ ]
Yes [ ] (please give details below)

……
……
……..

On average, how often do you use Dutch in your day-to-day life (e.g. at home, at work and/or in social situations)?

(Almost) always [ ]
More than half the time [ ]
Approximately half the time [ ]
Less than half the time [ ]
(Almost) never [ ]

Appendix III Sentence completion form

Instructions (test on paper)
On this and the following pages are 72 Dutch sentences. In each sentence, one word has been omitted. The position of this missing word is indicated by a solid line. Please read the sentence carefully and decide which Dutch word you feel would be the most appropriate within the context of the rest of the sentence, then fill in this word in the space provided.

It is important to read the sentence in full before you make your decision regarding the missing word. Please try to complete all the sentences. If you do not know the meaning of all the words in a sentence, you could try to establish the general meaning of the sentence in order to deduce what the missing word may be.

Instructions (test online)
On this and the following pages are 72 Dutch sentences. In each sentence, one word has been omitted. The position of this missing word is indicated by a pair of brackets. Please read the sentence carefully and decide which Dutch word you feel would be the most appropriate within the context of the rest of the sentence, then fill in this word in the space provided.

It is important to read the sentence in full before you make your decision regarding the missing word. Please try to complete all the sentences. If you do not know the meaning of all the words in a sentence, you could try to establish the general meaning of the sentence in order to deduce what the missing word may be. However, it is important for you to rely on your own knowledge of Dutch, so please do not use any dictionaries or grammar books or request help from others.

		1	Sinds vanochtend ( ) het zieke kind tot wanhoop van zijn moeder onophoudelijk gehuild.
		2	Iedereen in de zaal was zeer geïnteresseerd in ( ) mening van de Nederlandse gastspreker.
		3	In de loop van de middag ( ) de bankrekening van Kevin behoorlijk geslonken.
		4	Nadat haar zus was ingestapt ( ) Maria zo lang mogelijk naar de vertrekkende trein gezwaaid.
		5	Jan en Evert zijn boos op de scheidsrechter ( ) hij het tweede doelpunt onterecht afkeurde.
		6	Binnen enkele minuten ( ) het toestel van 35,000 voet naar ongeveer 11,000 voet gedaald.
		7	De wederopbouw van Nederland kwam langzaam op gang na ( ) bevrijding door de geallieerde troepen.
		8	Elke dag in de vakantie ( ) het kleine meisje hiernaast op haar trampoline gesprongen.

		9	Jan gaat binnenkort naar Groningen verhuizen ( ) hij heeft daar een boerderij gekocht.
		10	Volgens de arts van de regionale gezondheidskliniek ( ) waterpokken niet gevaarlijk voor jonge kinderen.
		11	Voordat het in slaap viel ( ) het kleine zwerfhondje in het asiel heel even gepiept.
		12	Peter gaat met de auto naar zijn werk ( ) hij vindt de bus veel te druk.
		13	Ondanks het gebrek aan inzet ( ) de studenten op de universiteit redelijk tevreden met hun cijfers.
		14	In tegenstelling tot de beweringen van sommige gelovigen ( ) mensen en dinosauriërs nooit gelijktijdig bestaan.
		15	Het onderzoek levert geen bewijs voor ( ) gedachte dat de economische situatie dit jaar enorm is verbeterd.
		16	Volgens de servicemonteur ( ) het rode waarschuwingslicht op het instrumentenpaneel enige tijd gebrand.

		17	Na de derde poging ( ) Lucas eindelijk voor zijn tentamens geschiedenis en aardrijkskunde geslaagd.
		18	Wegens de dreiging van de separatisten ( ) de vliegtuigen vorige maand op tien kilometer hoogte gevlogen.
		19	De overvallers hadden het terrein al verlaten toen ( ) inbraakalarm van de fabriek afging.
		20	Als Griekenland uit de euro stapt, ( ) de mogelijke gevolgen heel moeilijk te voorspellen.
		21	Sinds de vierde klas ( ) de drie lastige jongens niet meer naar de leraar geluisterd.
		22	Toen ze uiteindelijk het vissersdorpje binnenreden ( ) de zon al boven de horizon gestegen.
		23	De K2 is volgens kenners de moeilijkste berg om te beklimmen ( ) hij zeer steil is.
		24	Wegens zijn nieuwe baan ( ) Bart maandenlang tussen Groningen en Leiden heen en weer gereden.

		25	Voor het uittesten van de nieuwe apparatuur ( ) het leger nu op oefening in Duitsland.
		26	Tot grote verbazing van iedereen werd ( ) tennistoernooi uiteindelijk gewonnen door de jongste deelnemer.
		27	Tot nu toe ( ) de discussie over de begroting vooral op een ordinaire ruzie geleken.
		28	Al vind ik verhuizen niet zo leuk, ik ( ) de laatste tien jaar al drie keer geëmigreerd.
		29	De woningcorporatie vond het niet nodig om ( ) verhuur van de lege appartementen te stimuleren.
		30	De rebellen willen de spoorwegbrug in handen krijgen ( ) hij van strategisch belang is.
		31	Uit onderzoek blijkt dat twee derde van de leraren weleens getuige ( ) van pesterijen op school.
		32	Sinds de loodgieter is geweest ( ) de oude waterleidingen in de badkamer niet meer gelekt.

		33	De creationisten kregen een flink meningsverschil met de aanhangers van ( ) evolutietheorie van Darwin.
		34	Omdat de kinderen zich weer misdroegen ( ) de geplaagde docent gistermiddag de klas uit gevlucht.
		35	Voordat Jan gisteren naar huis mocht ( ) hij drie weken in het Rijnland ziekenhuis gelegen.
		36	Criminaliteit en misdaadbestrijding kosten ( ) maatschappij vele miljoenen per jaar volgens een recent onderzoek.
		37	Met prachtige stranden en een warm klimaat ( ) de Filippijnen een populaire bestemming voor toeristen.
		38	Tot opluchting van de jonge comédienne ( ) het publiek heel hard om haar grappen gelachen.
		39	Het schilderij van Rembrandt moest gerepareerd worden ( ) het beschadigd was door een bezoeker.
		40	Omdat zij failliet dreigden te gaan ( ) de twee bedrijven na een lang onderhandelingstraject gefuseerd.

		41	Opgezweept door de swingende muziek ( ) de bruiloftgasten op elk liedje enthousiast gedanst.
		42	Lisette gaat vaak naar het dorpscafé ( ) zij wil graag nieuwe mensen ontmoeten.
		43	Behalve enkele Indiaanse stammen ( ) de bevolking van Argentinië merendeels van Spaanse en Italiaanse afkomst.
		44	Voordat de dokter de wond kon hechten ( ) het been van Sjaak behoorlijk gebloed.
		45	De verzekeringspremie van de taxichauffeurs is verhoogd vanwege ( ) grote aantal ongelukken in de stad.
		46	In de loop van vorige week ( ) de soldaten met alle hulpgoederen naar het oorlogsgebied vertrokken.
		47	De taalstudenten missen zelden hun literatuurklas ( ) zij vinden de Nederlandse klassiekers uiterst boeiend.
		48	In het algemeen ( ) de Russische media geneigd het beleid van President Poetin te steunen.

		49	Na het kijken naar de enge film ( ) Jack en zijn vrienden de hele nacht onrustig gedroomd.
		50	In deze warme, natte zomer ( ) de naaldbomen in het duingebied veel harder gegroeid dan normaal.
		51	Gisteren zei de minister van onderwijs dat hij ( ) wetsvoorstel over het schoolbeleid niet zou steunen.
		52	In Londen ( ) de socialisten over het gebrek aan zendtijd op de staatstelevisie geklaagd.
		53	De winkelcentra in de steden maken moeilijke tijden mee ( ) steeds meer mensen online kopen.
		54	Omdat het bedrijf een gat in de begroting heeft, ( ) de financiële data nauwkeurig onderzocht.
		55	Voordat zij met haar vervolgopleiding begon ( ) de vriendin van mijn oudste broer veel gereisd.
		56	Vandaag ( ) de bril van meester Hans gesneuveld door een ongelukje met een voetbal.

		57	De vissersboten zijn vanochtend niet uitgevaren ( ) er was een zware storm op komst.
		58	Dankzij de inzet van Robben ( ) het Nederlandse elftal na enkele slechte wedstrijden eindelijk eens goed gespeeld.
		59	Wegens het slechte weer ( ) vijf vrachtwagens uit Litouwen op de Nederlandse snelweg gestrand.
		60	Blijkbaar kon de jonge spits ( ) geduld niet opbrengen om te wachten op een kans.
		61	Wegens de langdurige sneeuwval ( ) veel knaagdieren en insecten de winter niet overleefd.
		62	Na de reeks ontslagen van drie jaar geleden ( ) de redactie van het dagblad nu weer op volle sterkte.
		63	Op vakantie in Florida ( ) mijn zus in een waterpark met dolfijnen gezwommen.
		64	De wethouder is op staande voet ontslagen ( ) hij gefraudeerd had met de onkostenvergoedingen.

		65	Getergd door de arrogantie van de tegenstanders ( ) het hockeyteam vastberaden om te winnen.
		66	Tot opluchting van de ministers ( ) de slechte economische omstandigheden niet al te lang geduurd.
		67	Vorige week kwam het bestuur van de sportvereniging bij elkaar om ( ) nieuwe reglement te bespreken.
		68	Al was de dreiging reëel, de bommen van de terroristen ( ) uiteindelijk niet ontploft.
		69	De kleine jongen speelt elke dag voetbal met zijn vrienden ( ) hij wil later profvoetballer worden.
		70	Aan het einde van de zomer ( ) de pelgrims drie dagen lang door de Spaanse bergen gewandeld.
		71	Volgens sommige consumentenorganisaties ( ) kleding aanmerkelijk goedkoper in Duitsland dan in andere Europese landen.
		72	Voordat Marie naar het buitenland ging ( ) zij enkele jaren in een huurhuis in Enschede gewoond.

Thank you for taking part in this study. In order to complete the study, please fill in the self-assessment form on the next page.

Appendix IV Self-assessment form

How would you yourself describe your proficiency in Dutch? Please tick the appropriate box in each column in the table below.

Speaking		Understanding		Grammar		General	
Basic		Basic		Basic		Basic	
Moderate		Moderate		Moderate		Moderate	
Advanced		Advanced		Advanced		Advanced	
Near-native		Near-native		Near-native		Near-native	

Please indicate by ticking the appropriate box (Yes/No) next to each statement below whether or not you can perform the activity described in the statement.

	Yes	No
I can understand and use familiar everyday expressions and very basic phrases.		
I can introduce myself and can ask and answer simple personal questions such as where I live.		
I can interact in a simple way provided the other person talks slowly and clearly.		
I can understand sentences and frequent expressions related to basic matters, such as family, shopping and work.		
I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters.		
I can describe aspects of my background and immediate environment in simple terms.		
I can understand the main points of familiar topics regularly encountered in work, school and social situations.		
I can deal with most situations likely to arise whilst travelling in an area where the language is spoken.		
I can produce simple connected text on topics which are familiar or of personal interest.		
I can describe experiences, events, hopes and ambitions and briefly give reasons and explanations for opinions and plans.		
I can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in my specialist field.		
I can interact with a degree of fluency that makes regular interaction with native speakers possible without difficulty.		
I can produce clear, detailed text on a wide range of subjects and explain and argue a viewpoint on a topical issue.		
I can understand a wide range of demanding, longer texts, and recognise implicit meaning.		
I can express myself fluently and spontaneously without much obvious searching for expressions.		
I can use language flexibly and effectively for social, academic and professional purposes.		
I can produce clear, well-structured, detailed text on complex subjects.		
I can understand with ease virtually everything heard or read		
I can summarise information from different spoken and written sources, reconstructing arguments and accounts coherently.		
I can express myself spontaneously, very fluently and precisely, differentiating finer shades of meaning in complex situations.		

Appendix V Word frequencies

Unaccusatives	Frequency
gedaald	2.40
gevlucht	9.99
vertrokken	41.92
gestrand	2.04
geëmigreerd	0.23
gestegen	5.24
geslonken	0.11
gefuseerd	0.07
geslaagd	14.04
ontploft	11.32
gesneuveld	4.55
gegroeid	25.68
bestaan	56.96
overleeft	15.34
geleken	0.43
gewoond	12.83
geduurd	7.48
gestaan	10.91

Unergatives	Frequency
geluisterd	12.44
gedanst	5.72
gezwaaid	0.46
geklaagd	2.56
gelachen	4.94
gespeeld	32.04
gevlogen	10.61
gesprongen	9.01
gewandeld	1.56
gereden	19.55
gereisd	5.76
gezwommen	2.26
gedroomd	22.00
gebrand	3.2
gehuild	5.58
gepiept	1.37
gelekt	1.05
gebloed	0.69

The frequency is the number of times the participle appears per million words according to the SUBTLEX database.

Appendix VI Tables

1. Auxiliary selection by Dutch native speakers

Included		Excluded		Total
N	Percent	N	Percent	N
346	96,1%	14	3,9%	360

Verb class	Zijn	Hebben	N	Std. Deviation
COL	100,00	0,00	46	0,000
COS	96,67	3,33	60	18,102
CES	0,00	100,00	60	0,000
UCP	1,67	98,33	60	12,910
CMP	3,33	96,67	60	18,102
CNP	0,00	100,00	60	0,000
Total	30,92	69,08	346	46,285

2. Auxiliary selection by all English native speakers

Included		Excluded		Total
N	Percent	N	Percent	N
871	93,1%	65	6,9%	936

Verb class	Zijn	Hebben	N	Std. Deviation
COL	79,70	20,30	133	40,376
COS	81,21	18,79	149	39,197
CES	23,40	76,60	141	42,491
UCP	10,67	89,33	150	30,972
CMP	21,62	78,38	148	41,306
CNP	7,33	92,67	150	26,156
Total	36,62	63,38	871	48,205

3. Auxiliary selection by English native speakers (intermediate proficiency)

Included		Excluded		Total
N	Percent	N	Percent	N
452	89,7%	52	10,3%	504

Verb class	Zijn	Hebben	N	Std. Deviation
COL	69,70	30,30	66	46,309
COS	72,73	27,27	77	44,828
CES	32,86	67,14	70	47,309
UCP	17,50	82,50	80	38,236
CMP	30,38	69,62	79	46,283
CNP	8,75	91,25	80	28,435
Total	37,61	62,39	452	48,494

4. Auxiliary selection by English native speakers (advanced proficiency)

Included		Excluded		Total
N	Percent	N	Percent	N
419	97,0%	13	3,0%	432

Verb class	Zijn	Hebben	N	Std. Deviation
COL	89,55	10,45	67	30,819
COS	90,28	9,72	72	29,834
CES	14,08	85,92	71	35,034
UCP	2,86	97,14	70	16,780
CMP	11,59	88,41	69	32,250
CNP	5,71	94,29	70	23,379
Total	35,56	64,44	419	47,927

5. Percentage correct by all English native speakers

Included		Excluded		Total
N	Percent	N	Percent	N
871	93,1%	65	6,9%	936

Verb class	Mean	N	Std. Deviation
COL	79,70	133	40,376
COS	81,21	149	39,197
CES	76,60	141	42,491
UCP	89,33	150	30,972
CMP	78,38	148	41,306
CNP	92,67	150	26,156
Total	83,12	871	37,477

6. Auxiliary selection after removal of deviant COL and COS verbs (all English speakers)

Included		Excluded		Total
N	Percent	N	Percent	N
823	93,1%	61	6,9%	884

Verb class	Zijn	Hebben	N	Std. Deviation
COL	86,11	13,89	108	34,744
COS	77,78	22,22	126	41,740
CES	23,40	76,60	141	42,491
UCP	10,67	89,33	150	30,972
CMP	21,62	78,38	148	41,306
CNP	7,33	92,67	150	26,156
Total	34,39	65,61	823	47,529

7. Auxiliary selection after removal of deviant COL and COS verbs (advanced proficiency)

Included		Excluded		Total
N	Percent	N	Percent	N
395	96,8%	13	3,2%	408

Verb class	Zijn	Hebben	N	Std. Deviation
COL	94,55	5,45	55	22,918
COS	88,33	11,67	60	32,373
CES	14,08	85,92	71	35,034
UCP	2,86	97,14	70	16,780
CMP	11,59	88,41	69	32,250
CNP	5,71	94,29	70	23,379
Total	32,66	67,34	395	46,956

8. Auxiliary selection after removal of deviant COL and COS verbs (intermediate proficiency)

Included		Excluded		Total
N	Percent	N	Percent	N
428	89,9%	48	10,1%	476

Verb class	Zijn	Hebben	N	Std. Deviation
COL	77,36	22,64	53	42,252
COS	68,18	31,82	66	46,934
CES	32,86	67,14	70	47,309
UCP	17,50	82,50	80	38,236
CMP	30,38	69,62	79	46,283
CNP	8,75	91,25	80	28,435
Total	35,98	64,02	428	48,051

9. Mean auxiliary selection per word (all English native speakers)

Verb class	Verb	Mean selection ***zijn***	Mean selection ***hebben***	N	Std. Deviation
COL	vluchten	84,00	16,00	25	37,417
	stranden	95,65	4,35	23	20,851
	stijgen	88,89	11,11	18	32,338
	emigreren	52,00	48,00	25	50,990
	vertrekken	84,62	15,38	26	36,795
	dalen	75,00	25,00	16	44,721
COS	fuseren	73,08	26,92	26	45,234
	slinken	79,17	20,83	24	41,485
	slagen	76,92	23,08	26	42,967
	sneuvelen	100,00	0,00	23	0,000
	groeien	80,77	19,23	26	40,192
	ontploffen	79,17	20,83	24	41,485
CES	bestaan	30,00	70,00	20	47,016
	wonen	4,00	96,00	25	20,000
	liggen	17,39	82,61	23	38,755
	lijken	64,00	36,00	25	48,990
	duren	12,50	87,50	24	33,783
	overleven	12,50	87,50	24	33,783
UCP	bloeden	12,00	88,00	25	33,166
	branden	26,09	73,91	23	44,898
	piepen	0,00	100,00	25	0,000
	lekken	11,54	88,46	26	32,581
	huilen	12,00	88,00	25	33,166
	dromen	3,85	96,15	26	19,612
CMP	vliegen	24,00	76,00	25	43,589
	wandelen	16,67	83,33	24	38,069
	reizen	23,08	76,92	26	42,967
	springen	24,00	76,00	25	43,589
	zwemmen	8,00	92,00	25	27,689
	rijden	34,78	65,22	23	48,698
CNP	klagen	23,08	76,92	26	42,967
	luisteren	0,00	100,00	26	0,000
	spelen	4,17	95,83	24	20,412
	zwaaien	4,17	95,83	24	20,412
	lachen	8,00	92,00	25	27,689
	dansen	4,00	96,00	25	20,000
Total		36,62	63,38	871	48,205

10 Pairwise Comparisons *zijn* selection (all English native speakers)

Verb class		Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
COL	COS	-3,502	4,522	1,000	-18,238	11,234
	CES	55,704*	6,138	,000	35,703	75,706
	UCP	67,986*	6,972	,000	45,269	90,703
	CMP	57,351*	6,942	,000	34,730	79,972
	CNP	70,764*	6,233	,000	50,454	91,074
COS	COL	3,502	4,522	1,000	-11,234	18,238
	CES	59,206*	6,985	,000	36,446	81,967
	UCP	71,488*	6,645	,000	49,835	93,142
	CMP	60,853*	5,907	,000	41,607	80,100
	CNP	74,266*	5,717	,000	55,638	92,894
CES	COL	-55,704*	6,138	,000	-75,706	-35,703
	COS	-59,206*	6,985	,000	-81,967	-36,446
	UCP	12,282	4,870	,281	-3,586	28,149
	CMP	1,647	5,622	1,000	-16,671	19,965
	CNP	15,060*	4,243	,024	1,233	28,886
UCP	COL	-67,986*	6,972	,000	-90,703	-45,269
	COS	-71,488*	6,645	,000	-93,142	-49,835
	CES	-12,282	4,870	,281	-28,149	3,586
	CMP	-10,635	5,010	,665	-26,961	5,691
	CNP	2,778	2,760	1,000	-6,214	11,770
CMP	COL	-57,351*	6,942	,000	-79,972	-34,730
	COS	-60,853*	5,907	,000	-80,100	-41,607
	CES	-1,647	5,622	1,000	-19,965	16,671
	UCP	10,635	5,010	,665	-5,691	26,961
	CNP	13,413	4,419	,086	-,985	27,810
CNP	COL	-70,764*	6,233	,000	-91,074	-50,454
	COS	-74,266*	5,717	,000	-92,894	-55,638
	CES	-15,060*	4,243	,024	-28,886	-1,233
	UCP	-2,778	2,760	1,000	-11,770	6,214
	CMP	-13,413	4,419	,086	-27,810	,985

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

11. Pairwise Comparisons *zijn* selection (all English native speakers) after removal of deviant verbs

Verb class		Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
COL	COS	6,726	4,429	1,000	-7,706	21,158
	CES	62,520*	6,158	,000	42,455	82,585
	UCP	74,802*	7,069	,000	51,769	97,834
	CMP	64,167*	6,630	,000	42,564	85,770
	CNP	77,579*	6,477	,000	56,475	98,684
COS	COL	-6,726	4,429	1,000	-21,158	7,706
	CES	55,794*	7,654	,000	30,855	80,732
	UCP	68,075*	7,462	,000	43,760	92,391
	CMP	57,440*	6,507	,000	36,239	78,642
	CNP	70,853*	6,485	,000	49,722	91,984
CES	COL	-62,520*	6,158	,000	-82,585	-42,455
	COS	-55,794*	7,654	,000	-80,732	-30,855
	UCP	12,282	4,870	,281	-3,586	28,149
	CMP	1,647	5,622	1,000	-16,671	19,965
	CNP	15,060*	4,243	,024	1,233	28,886
UCP	COL	-74,802*	7,069	,000	-97,834	-51,769
	COS	-68,075*	7,462	,000	-92,391	-43,760
	CES	-12,282	4,870	,281	-28,149	3,586
	CMP	-10,635	5,010	,665	-26,961	5,691
	CNP	2,778	2,760	1,000	-6,214	11,770
CMP	COL	-64,167*	6,630	,000	-85,770	-42,564
	COS	-57,440*	6,507	,000	-78,642	-36,239
	CES	-1,647	5,622	1,000	-19,965	16,671
	UCP	10,635	5,010	,665	-5,691	26,961
	CNP	13,413	4,419	,086	-,985	27,810
CNP	COL	-77,579*	6,477	,000	-98,684	-56,475
	COS	-70,853*	6,485	,000	-91,984	-49,722
	CES	-15,060*	4,243	,024	-28,886	-1,233
	UCP	-2,778	2,760	1,000	-11,770	6,214
	CMP	-13,413	4,419	,086	-27,810	,985

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.