Can the late bird catch the worm?
Ultimate attainment in L2 syntax

Can the late bird catch the worm?
Ultimate attainment in L2 syntax

een wetenschappelijke proeve op het gebied van de Letteren

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de Rector Magnificus prof. dr. C.W.P.M. Blom,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op maandag 20 juni
des namiddags om 1.30 uur precies

door

Sonja Joke van Boxtel

geboren op 5 maart 1975 te Rotterdam

Promotor:
Prof. dr. P.C. Muysken

Copromotores:
Dr. Th.C.G Bongaerts
Dr. P.A.J.M. Coppen

Manuscriptcommissie:
Mw. prof. dr. A.M.C. van Kemenade (vz)
Prof. dr. D.P. Birdsong (University of Texas)
Mw. dr. M. Gullberg (Max Plack Institut für Psycholinguistik)

Dankwoord

Op deze plaats wil ik graag iedereen bedanken die op één of andere manier heeft geholpen bij de uitvoering van mijn project en het schrijven van dit proefschrift en/of heeft bijgedragen aan de plezierige omstandigheden waarin ik de afgelopen zes jaar aan mijn project heb gewerkt. Sommige van deze mensen zullen dit dankwoord niet kunnen lezen, omdat zij geen Nederlands beheersen en van sommigen is hun bijdrage te omvangrijk om hier goed te beschrijven. Een deel van hen zal ik ook persoonlijk bedanken.

In de eerste plaats wil ik mijn begeleiders, Theo Bongaerts en Peter-Arno Coppen, bedanken. Met hen heb ik van het begin tot het einde regelmatig overlegd in een zeer plezierige sfeer. Hun advies, steun, vertrouwen, positieve feedback, tijd en hulp hebben ertoe geleid dat mijn project vrij soepel en zeer aangenaam is verlopen en hebben een belangrijke bijdrage geleverd aan het resultaat. Ik kijk dan ook met veel plezier terug op onze samenwerking en vind het jammer dat die nu is afgelopen. Theo wil ik daarnaast nog bedanken voor zijn zeer gedetailleerde commentaar op mijn stukken en voor zijn steun, advies en medeleven met alle persoonlijke dingen waar ik de afgelopen jaren mee te maken heb gehad, zowel in positieve als in negatieve zin. De manier waarop hij altijd voor mij klaar stond en regelmatig van zich liet horen als er iets belangrijks speelde, heb ik als zeer bijzonder ervaren. Peter-Arno wil ik speciaal bedanken voor de programma's die hij voor mij heeft geschreven voor het afnemen van mijn zinsoordelentaak en voor zijn hulp bij het ontrafelen van de factoren die een rol spelen bij constructies met vervangende subjecten in het Nederlands.
    Vervolgens wil ik mijn opeenvolgende promotores, Kees de Bot en Pieter Muysken bedanken, voor de plezierige samenwerking, hun steun, hun vertrouwen in mij en mijn project, en hun positieve feedback. Toen Kees naar Groningen vertrok en Pieter zijn rol heeft overgenomen, kwam hij terecht in een project waar hij geen enkele zeggenschap over had gehad en bij een promovenda die hij niet goed kende. Ik waardeer het erg dat hij mijn begeleiders en mij in die situatie veel ruimte en vertrouwen heeft gegeven. Bovendien waardeer ik zijn algemene betrokkenheid, die verder strekt dan alleen mijn project en proefschrift.
    De leden van mijn manuscriptcommissie, Ans van Kemenade, David Birdsong en Marianne Gullberg wil ik bedanken voor de tijd en moeite die ze hebben genomen om mijn proefschrift te beoordelen. Daarnaast wil ik Leah Roberts graag bedanken voor het proeflezen van mijn proefschrift. Ik

waardeer het erg dat ze dit spontaan heeft aangeboden, terwijl ze bij mijn project verder niet betrokken is geweest.

Naast mijn begeleiders en promotores, hebben ook veel andere collega's bijgedragen aan mijn project. Ik denk dan aan de hulp op statistisch gebied van Frans van der Slik en Roeland van Hout en aan de hulp die ik heb gekregen bij mijn analyse van het Nederlands, Duits, Frans en Turks van verschillende moedertaalsprekers en deskundigen. In dit verband noem ik Hans Broekhuis, Simon van Dreumel, Irene Krämer, Gertjan Postma, Helen de Hoop, Ineke van der Craats, Marco Haverkort, Ad Foolen, WUS van Lesse Kloeke, Dominique Nouveau, Mireille Piot, Claire Moyse, Evelyne Vos-Fruit, Marianne Starren, Marianne Gullberg, Aslı Özyürek, Hugo Strötbaum en Ayşe Gürel en hoop ik van harte dat ik niemand vergeet. Ook wil ik Margriet Jagtman bedanken voor haar rol bij de aanvraag van dit project, Brigitte Planken voor het inspreken van de zinnen van de imitatietaak en David Birdsong, Niclas Abrahamson, Laura Sabourin en het publiek van verschillende conferenties en bijeenkomsten voor hun feedback met betrekking tot de opzet en uitvoering van mijn onderzoek.

De laatste - maar zeker niet de minste - mensen die ik heel hartelijk wil bedanken voor hun directe betrokkenheid bij mijn project zijn alle proefpersonen die hebben meegedaan aan mijn pilot-onderzoek en mijn uiteindelijke onderzoek en alle contactpersonen die hebben geholpen bij het vinden van proefpersonen. Zonder hen zou dit proefschrift er niet zijn. Helaas zijn het er teveel om allemaal persoonlijk te bedanken. Het contact met mijn proefpersonen vond ik één van de plezierigste onderdelen van mijn project. Door mijn project heb ik veel mensen ontmoet die ik anders misschien nooit zou zijn tegengekomen en heb ik veel leuke gesprekken gevoerd, waar ik geregeld aan terug denk. Een leuke bijkomstigheid was dat ik voor het afnemen van mijn onderzoek bij de tweedetaalleerders op allerlei plaatsen ben geweest waar ik normaal niet snel zou komen.

Naast alle mensen die ik hierboven heb genoemd, zijn er ook een hoop mensen die niet heel direct hebben bijgedragen aan mijn proefschrift, maar wel een belangrijke rol hebben gespeeld in mijn project of bij de totstandkoming daarvan.

Allereerst wil ik in dit verband Ad Foolen, Eric Kellerman en Lydia White bedanken. Zonder Ads advies als studieadviseur van Algemene Taalwetenschap, Erics tentamen voor de cursus "rol van de moedertaal" en Lydia Whites artikel over Universele Grammatica in tweedetaalverwerving was ik misschien wel nooit in de Toegepaste Taalwetenschap terecht gekomen. Daarnaast wil ik Theo Bongaerts, Ad Foolen en Margriet Jagtman bedanken voor het vertrouwen dat ze in mij hadden en de kans die ze mij hebben geboden om zelf een onderwerp voor een project voor te stellen. Ook

de deelnemers aan de CLS-cursus van Bonnie Schwartz in 1996, waar ik als student aan mee heb gedaan, hebben een belangrijke rol gespeeld bij mijn beslissing om junior onderzoeker te worden. In dit verband wil ik met name Ineke van de Craats en Marianne Starren noemen. Zij hebben mij na de cursus nog geholpen bij mijn scriptie en ik vind het erg leuk dat ze later allebei ook aan de KUN zijn komen werken.

Tijdens mijn project zijn er heel veel collega's geweest met wie ik leuke en nuttige discussies en/of gesprekken heb gevoerd: mijn kamergenoten Ineke Huibregtse, Tan Thein, Ylva Falk, Griet Coupé, Kaoru Yoshioka en Else Havik, (overige) collega's van Taalwetenschap en Bedrijfscommunicatie (teveel om allemaal persoonlijk te noemen), mensen van de acquisitiegroep in Utrecht en van het MPI, docenten van LOT-cursussen, collega's uit binnen- en buitenland die ik ken van congressen e.d. en medepromovendi uit Nijmegen en andere steden. Van deze laatste groepen wil ik met name Aoju Chen, Pytsje van der Veen, Barbara Schmiedtóva, Christiane Dittrich, Laura Sabourin, Wouter Kusters, Sharon Unsworth, Holger Hopp, Marianne Gullberg, Leah Roberts, Christine Dimroth, Aafke Hulk, David Birdsong, Clive Perdue, David Singleton, Mike Long, Cathy Doughty en Bonnie Schwartz noemen. Al deze mensen hebben er in verschillende mate in bijgedragen dat ik met veel plezier terugkijk op mijn werk als junior onderzoeker.

Tot slot wil ik graag Aoju, Irene en Ineke bedanken voor de rol die ze hebben gespeeld en nog zullen spelen bij de voorbereiding van mijn verdediging. Mijn moeder wil ik bedanken dat ze mij de gelegenheid heeft gegeven om mijn eigen weg te vinden, zonder enige druk om een studie te kiezen met veel geld en/of status in het vooruitzicht. Erik bedank ik voor zijn praktische en morele steun tijdens mijn project. Tenslotte wil ik Yana en Erik allebei bedanken voor het creëren van een plezierige thuissituatie, wat het werken aan mijn project en proefschrift zeker ten goede is gekomen.

Contents

List of abbreviations

Types and categories in the sentence preference test

DP-type (Active) sentences with a noun phrase / DP as subject
        DPg   general category of the DP-type
        DPnst non-specific (DP) subjects with a transitive predicate
        DPnsi non-specific (DP) subjects with an intransitive predicate
AS-type Active sentences with a sentential subject (embedded clause)
        ASg   general category of the AS-type
        AScos category of the AS-type containing a change of state
PS-type Passive sentences with a sentential subject (embedded clause)
        PSg   general category of the PS-type
        PSdo category of the PS-type of sentences with a dummy object
                in their active equivalent

List of other abbreviations

1s	first person singular
2s	second person singular
3s	third person singular
AgrSP	Agreement Subject Phrase (see Radford, 1997)
AoA	Age of arrival (in the L2 environment)
ACC	Accusative case
CP	Complementizer Phrase (see Radford, 1997)
CPH	Critical Period Hypothesis
D	Dutch
DAT	Dative case
DP	Determiner Phrase (see Radford, 1997)
DS	Dummy subject
dsp	Dummy subject pair of category excluded from the main analyses
EPP	Extended Projection Principle (see Radford, 1997)
F	French
G	German
GEN	Genitive case
ID	Identification code of the participants
IP	Inflection Phrase (see Radford, 1997)
L1	First language
L2	Second language

LF	Logical Form (see Radford, 1997)
LoR	Length of residence (in the L2 environment)
N	Number (of participants in a group)
NEG	Negation
NOM	morpheme that transforms a verb into a nominal element
NS	Native speaker
p	Predicate excluded from main analyses
PAST	Past tense
POSS	Possessive morpheme
PP	Prepositional Phrase
PROGR	Progressive
SC	Small Clause (see Radford, 1997)
SD	Standard Deviation
SIT	Sentence imitation task
SLA	Second language acquisition
SPT	Sentence preference task
T	Turkish
TP	Tense Phrase (see Radford, 1997)
V2	Verb second (see Radford, 1997)
vP	light verb Phrase (see Radford, 1997)
VP	Verb Phrase (see Radford, 1997)
z	z-score (score - mean, divided by the standard deviation)
*	Ungrammatical sentence
?*	Native speakers have doubts about the grammaticality of the sentence

Chapter

1

Introduction

Few people would disagree with the statement that the younger people start learning a new language, the better their chances of success (i.e. reaching a very high level of proficiency) will be. In this respect, learning a language is not very different from many other kinds of learning, like sports and music.

There is much less agreement, however, on the question of whether there are exceptions to this rule. When people are asked whether they think it is possible for late learners to reach a native level of proficiency in a second language (L2), different answers will come up, depending among other things on people's own learning experiences and those of people around them. When I was collecting my data, for example, some participants who were very interested in my research and keen to participate and see the results, told me that they could already give me the answer to my main research question, considering the possibility of native-like attainment. They were sure it would not be possible for late learners to reach a native level in the second language.

There is also disagreement in the scientific literature on ultimate attainment (see chapter 2). When my research project started in 1999, it had already been shown that there were late L2 learners who had reached a native level of proficiency in pronunciation (e.g. in Bongaerts, Van Summeren, Planken & Schils, 1997). Whether this could also be the case for morphosyntax was not clear. Moreover, the role of the typological distance between the L1 and the L2 had received scant attention. This is why I started to work with great enthusiasm on the research project on ultimate attainment of syntax in second language acquisition (SLA), described in this dissertation.

During my project, the critical period hypothesis for syntax became a popular topic in the field of SLA research and many studies on this topic were carried out. Some recent studies also address the relation between ultimate attainment and the typological distance between source and target languages. However, this issue has not been fully settled yet.

We decided to study native-like attainment for dummy subject constructions (for a definition and examples see chapter 3), because these constructions are known to be very difficult to acquire for second language learners and because learners can only acquire them through processing input. There are no simple rules available to second language learners that they can learn by heart and apply without really having acquired them.

Before I started this research project I believed there might be late learners who are able to reach a native level in L2 syntax. Had I continued to do so, I would probably have formulated the hypothesis that native-like attainment in L2 syntax is still possible for late learners. However, when looking into dummy subject constructions in Dutch, it became very difficult for me to imagine how any learner could acquire the appropriate rules for these constructions in Dutch, let alone late L2 learners. For this reason, my main research question, concerning the possibility of native-like attainment, remained an open question until I analysed the L2 data. To reflect this, I have formulated research questions rather than hypotheses in this dissertation.

In this dissertation, the question of whether there are late second language learners with native-like attainment for syntax will be linked to the Critical Period Hypothesis for syntax in SLA. According to this hypothesis, which will be elaborately discussed in chapter 2, native-like attainment should not be possible for late learners due to a biologically determined decrease in sensitivity to language input after puberty.

Another important issue in the ultimate attainment discussion is the role of the L1. In spite of the importance of this issue, many studies on ultimate attainment with very proficient late learners either use closely related language pairs or do not systematically control for the differences between the first languages and the target language. In our study, we did systematically control for the typological distance between the source languages and the target language, by testing groups of participants from three different L1 backgrounds (German, French and Turkish) which differ in their typological distance to Dutch. We made sure that the languages selected were also different from each other and from Dutch in their use of the constructions we investigated, namely dummy subject constructions (see chapter 3).

If it turns out that native-like attainment of L2 syntax is possible for late learners, it becomes relevant to find out what the input and background characteristics are of those late learners who perform within the native speaker range. To address this question, we used a questionnaire in which we asked the participants about several background characteristics. The results of this questionnaire will be discussed in chapter 5.

We will start, however, with a discussion of the critical period hypothesis and relevant literature on ultimate attainment and the role of the

L1 (in chapter 2). In chapter 3, we will discuss dummy subjects and factors that turned out to be relevant in the choice for dummy subjects, both in general and in the languages involved in our study (Dutch, German, French and Turkish). In this chapter, the results of the native speakers of Dutch in our study will also be discussed. Although our data were analysed in a theory- neutral way, some attention will be paid in this chapter to generative theories that address dummy subject constructions. In chapter 4, we will discuss the methodology of our study and the way in which we analysed our data. Subsequently, the results of the non-native speakers on the two tasks in our study will be presented and discussed (in chapters 5 and 6). In the final chapter of this dissertation, we will summarise the main findings and discuss the outcomes in relation to the research questions presented above and the implications that they have both for research and for second language learners.

Chapter

2

The Critical Period

Hypothesis

In this chapter, we will discuss the Critical Period Hypothesis (hence CPH), which plays an important role in our study. In paragraph 2.1 the biological concept of a critical period and its features will be introduced. Subsequently, we will very briefly discuss the notion of a critical period for language acquisition (paragraph 2.2) and introduce the assumptions associated with it. Paragraphs 2.3 to 2.5 contain a review of the literature related to these assumptions. Neuroimaging studies that contribute to the CPH discussion will be reviewed separately in paragraph 2.6. In paragraph 2.7 other factors that may cause observed age effects in second language acquisition will be discussed. In the last paragraph (2.8), we will try to explain how our study might contribute to the research that has been carried out so far with respect to the CPH in second language acquisition.

2.1   Critical periods

In biology, there are many instances of learning for which there is a critical period, i.e. "a time during the life span of an organism in which the organism may be affected by some exogenous influence to an extent beyond that observed at other times" (Colombo, 1982: 261). During this period, typically early in life, there is a (heightened) sensitivity to stimuli that are necessary for the development of the ability concerned. After this period, there is a non-linear decline in sensitivity. If the relevant stimuli are not present during this critical period, the ability concerned will no longer (fully) develop under normal circumstances.

A well-studied example of a critical period in biology is the development of orientation specificity in the visual cortex for cats (e.g. Baxter, 1966): when cats are only exposed to certain patterns (e.g. only

horizontal stripes) during the first weeks of their lives, they will never be able to perceive other patterns (e.g. vertical stripes). The end of this critical period is rather abrupt and absolute and it has long been thought that these characteristics were typical of all critical periods. There are many critical periods, however, after which the decline in sensitivity and learning ability is much more gradual.[1] Another example of a critical period is imprinting in ducklings (e.g. Ratner & Hoffman, 1974). When ducklings are born, they follow the first moving object they see, but only during the first few days. This critical period is claimed to be related to the maturation of fear for novel objects (Hoffman, 1974). This means that the shape of the underlying maturational mechanism is in this case opposite to that of the critical period itself (rise of fear versus decline in imprinting).

The kind of critical period that is assumed to be most closely related to language learning is the critical period for song learning in certain songbirds, e.g. zebra finches or white-crowned sparrows (e.g. Marler, 1970). When these birds do not hear the specific song of their own species during a certain period after birth, they will usually start singing, but they will not develop the fully-fledged song that is specific for their species during the mating season (contrary to "open learners", like canaries, who can continue to learn new songs throughout their lives). These critical periods are probably affected by hormone levels, since they can be manipulated by castration and administration of hormones.

Although there are differences between different types of critical periods in many aspects (see e.g. Bornstein, 1989), they share certain geometric features, which distinguish them from other types of development. In general, critical periods have an onset, a peak of heightened sensitivity, an offset and a terminus with a flattening after the terminus, as in figure 2.1.

---

[1]   To make a distinction between periods with an abrupt decline versus periods with a more gradual decline, some researchers use the term sensitive period for the latter. The distinction between the terms sensitive (or optimal) period versus critical period is also used, however, to distinguish periods after which no plasticity remains and normal development is no longer possible (critical periods) from periods after which some plasticity in the critical system remains and under special circumstances recovery (to a certain degree) is still possible (optimal or sensitive periods). In spite of these distinctions, most researchers use the term critical period for all such periods, as we will do here.
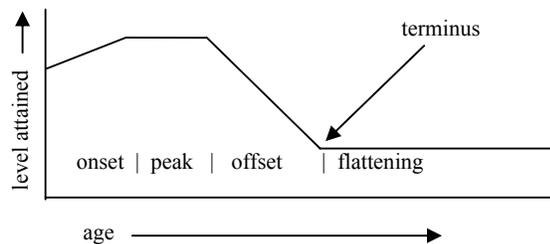
Figure 2.1:   Geometric features of a critical period

During the onset, there is a gradual rise in sensitivity to the critical stimuli. The onset is maturationally determined, but can in some cases be postponed (thus postponing the whole period) when no stimuli are presented, for example when cats are reared in the dark for a number of weeks. During the peak, exposure to the critical stimulus is most effective. This peak can be very short (for example for certain kinds of visual development in cats). The terminus indicates the end of the critical period of heightened sensitivity. After this point, the relevant behaviour should no longer be correlated with age and should stay at approximately the same level. Note that this level of ability (for example the ability to discriminate different visual patterns) should be lower than levels attained before the terminus. For some critical periods, the onset and/or peak can start so early (e.g. before birth for auditory stimuli) or be so short that they do not (seem to) play a role. In order to prove that there is a critical period for a certain ability (e.g. language acquisition), however, one must prove that an offset and flattening are present, and that after the terminus age is no longer correlated with this ability. This means that, under normal circumstances, late learners with the same amount (and quality) of input should behave significantly differently from early learners and not significantly differently from other late learners.

2.2   The critical period hypothesis for language acquisition

In the field of language acquisition, the hypothesis that there is a critical or sensitive period for language acquisition plays an important role. In the late fifties and sixties, Penfield and Roberts (1959) and Lenneberg (1967) were the first to propose that there was a critical period for language acquisition. They based this proposal on different types of evidence:
-    evidence from feral and abused children who grew up without being exposed to human language in childhood and who did not acquire language normally after they were found

7

- evidence from deaf children whose development in spoken language stopped after puberty
- evidence that children with aphasia recovered much better than adults with aphasia

All this evidence comes from first language acquisition. Since we only investigate second language acquisition in this study, we will not discuss any of the evidence for a critical period in first language acquisition here. Lenneberg (1967), however, also makes a claim about second language acquisition. He writes on this topic:

> Most individuals of average intelligence are able to learn a second language after the beginning of their second decade, although the incidence of "language-learning-blocks" rapidly increases after puberty. Also automatic acquisition from mere exposure to a given language seems to disappear after this age, and foreign languages have to be taught and learned through a conscious and labored effort. Foreign accents cannot be overcome easily after puberty. However, a person *can* learn to communicate in a foreign language at the age of forty. This does not trouble our basic hypothesis on age limitations because we may assume that the cerebral organization for language learning as such has taken place during childhood, and since natural languages tend to resemble one another in many fundamental aspects … , the matrix of language skills is present. (p. 176)

It is this notion of a critical period for second language acquisition that most SLA researchers refer to in studies that are performed to test the CPH for SLA.

This notion was later further refined and different causes were proposed (see e.g. Birdsong, 2005). Pinker (1994), for example states that:

> (…) acquisition of a normal language is guaranteed for children up to the age of six, is steadily compromised from then until short after puberty, and is rare thereafter. Maturational changes in the brain, such as the decline in metabolic rate and the number of neurons during early school-age years, and the bottoming out of the number of synapses and metabolic rate around puberty, are plausible causes. (p. 294; cited in Birdsong, in press).
> Language-acquisition circuitry is not needed once it has been used; it should be dismantled if keeping it around incurs any [metabolic] costs. And it probably does incur costs. (p. 294; cited in Birdsong, in press)

In paragraph 2.7 we will discuss causes of age effects in L2 acquisition more elaborately.

In both first and second language acquisition people often discuss "the critical period for language acquisition" as if there were one critical period for

all aspects of language. It is generally assumed, however, that there are different critical periods for different parts of language, e.g. phonology and syntax (see e.g. Seliger, 1978 and Long, 1990) and that not for all parts of language there is a critical period. Scovel (1988), for example, proposed that there is only a critical period for pronunciation.

Second language acquisition
Since the 1960s much research has been done on the factor of age in second and foreign language acquisition. It has been shown that, though older language learners have an initial advantage over younger learners, in the long run young language learners tend to achieve higher levels of success than older learners (e.g. Krashen, Long & Scarcella, 1982; Jia, 1998). This last result is generally interpreted as evidence in favour of the CPH. However, in order to test the assumption that the critical period effect is caused by neurophysiological factors, as Lenneberg (1967) proposes, one needs additional types of evidence (see e.g. Bongaerts, 2003).

More recent research with respect to the CPH in second language acquisition has tended to focus on finding evidence bearing on one of three assumptions connected to the hypothesis that there is a biologically based critical period. First, under the CPH one would expect that "L2-proficiency[2] should show a discontinuous function across ages of acquisition" (Bialystok & Miller, 1999). According to this assumption there should be a significant negative correlation between age of arrival and L2-proficiency until the terminus and no significant correlation between these variables after that. The second assumption is that no late second or foreign language learner will be able to attain a level of success in the target language that is comparable to that of native speakers. The third assumption (which is usually not the focus in studies testing the CPH) is that language learning limitations related to maturation should result in the same relation between age and proficiency for different L1/L2 combinations (see e.g. Birdsong & Molis, 2001). In paragraphs 2.3-2.5, we will discuss literature related to these assumptions.

It should be noted that several researchers have tried to provide decisive evidence for the existence of a critical period for SLA by testing only the assumption that there should be a discontinuous correlation between age of onset of second language acquisition and proficiency. However, in order to prove the existence of a critical period for second language acquisition, all geometric features that are essential for a critical period should be found and the effects found must have a biological (neurological) cause.

---

[2]   The term "proficiency" is not used to refer specifically to "performance" rather than "competence". It is merely used to denote the level reached by second language learners, irrespective of whether this reflects their "competence" or "performance".

2.3 The shape of the age function for SLA

Arguably, the most influential study to date relating to the assumption that there should be a discontinuous relation between L2-proficiency and age of arrival is the one reported on by Johnson and Newport (1989). In this study, 46 adult Chinese and Korean second language learners of English were studied, who had arrived in the United States at different ages, varying from three to 39 and who had had contact with English (in the US) for at least five years. They were divided into four groups, based on their age of arrival (AoA) in the United States: AoA 3-7, AoA 8-10, AoA 11-15 and AoA 17-39. Participants were not selected on the basis of their level of proficiency. There was also a control group consisting of 23 native speakers of English. All participants were given a grammaticality judgement test designed to test twelve morphosyntactic rules, such as plural marking, particle movement, word order, and past tense. Analyses of the results showed that there was a strong significant relationship between age of arrival and test scores (r=-.77, p<.01). The AoA 3-7 group was the only group that did not differ significantly from the native controls. There was a linear decline in proficiency for the AoA 8-10 and 11-15 groups. The correlation between AoA and test scores was high for the AoA < 15 group (r=.87). In the AoA 17-39 group, on the other hand, there was no longer a significant correlation between test scores and AoA (r=-.16, p >.05), and there was a great deal of interlearner variability, which was absent in the other groups. Johnson and Newport conclude from these results that there is a maturationally determined critical period for SLA, which closes after puberty.

Many researchers have criticised this study and its conclusion. Bialystok and Hakuta (1994) reanalysed the data from this study. They show that when the subjects are classified differently according to their AoA, i.e. a group with AoA<20 and a group with AoA>20, the correlation between AoA and proficiency is the same for the younger group (r = -.87) and almost significant for the older group (r = -.49). Bialystok (1997) argues that the difference between these groups, especially in interlearner variablility, could very well be caused by a difference in educational background, since the younger learners all received English language instruction in American schools, whereas (at least most of) the older learners did not. With respect to the discontinuity at age seven, she suggests that this might be due to the fact that for the AoA<7 group, English may have been their dominant language, since there is no information about their proficiency in Chinese or Korean.

Many researchers have replicated Johnson and Newport's 1989 study. Among them are Jia (1998), Bialystok and Miller (1999), Flege, Yeni-Komshian and Liu (1999), DeKeyser (2000) and Birdsong and Molis (2001).

Jia (1998) studied 105 non-native speakers of English with twelve different first languages (most non-natives had an L1 which is unrelated to English), who came to the United States between the ages of 3 and 34 and had lived there for a period ranging from 5 to 32 years. The participants' grammatical proficiency was measured through two grammaticality judgement tasks, based on the task used by Johnson and Newport (1989), one presented orally and the other visually. Jia found that younger age of arrival predicted better performance on L2 grammaticality judgements (r = -.68, $p < .001$). It should be noted that Jia did not distinguish any age groups and did not address the question of whether there was a discontinuity in the age function. Since younger age of arrival and higher L2 proficiency was also strongly related to (lower) L1 proficiency and a richer L2 environment in her study, she concluded that "age differences in L2 acquisition are strongly influenced by factors other than biological maturation" (Jia, 1998, p. ii). This does not answer the question, however, of whether there is also a significant influence of biological maturation.

Bialystok and Miller (1999) studied 28 Spanish and 33 Chinese second language learners of English and a control group of 38 native speakers of English. They used a (oral and written) grammaticality judgement test like the one used by Johnson and Newport, but this test was specifically designed to control for differences in structure between Chinese and English. Besides accuracy, Bialystok and Miller also measured reaction times. They also used a first language proficiency task, to decide whether English was really the participants' second, non-dominant, language. For the analyses, the learners were divided into four groups, AoA 1-8, AoA 9-15, AoA 16-22 and AoA 23-33. This division was based on findings from the Johnson and Newport study. Unlike Johnson and Newport, Bialystok and Miller did not find any significant differences for accuracy between the three oldest learner groups. They only found significant differences between the AoA 1-8 group and the other learner groups. For the Chinese learners they also found a significant difference between the youngest learner group and the native speakers (t=2.47, p=.04), whereas for the Spanish learners only the three older learner groups differed significantly from the native speaker controls. For the reaction times, there was only a main effect of age of arrival for the Spanish-English bilinguals (F=7.16, $p<.01$). For the reaction times of the Chinese-English bilinguals, there was an interaction between grammaticality, age of arrival and modality (F=4.40, $p<.04$). Bialystok and Miller conclude from their findings that there is no evidence for a critical period for second language acquisition. According to them, the influence of the first language that was found for the younger learners and the lack of qualitative differences for task type and structure between the younger and older learners support this result.

Flege et al. (1999) used 128 items from the Johnson and Newport 1989 study (constructions that caused very few errors in previous administrations were left out), complemented by 16 new sentences, in an orally presented grammaticality judgement test. They tested 240 Korean learners of English with AoAs ranging from 1 to 23, divided over 10 AoA groups. T-tests showed that all but the two youngest learner groups (mean AoA = 3 and mean AoA = 5) differed significantly from the native speakers ($p < .01$). For the foreign accent ratings test in their study, all learner groups differed significantly from the native speakers. They also found a significant correlation between AoA and scores on the grammaticality judgement test for both younger and older learners (AoA < 12: r= -.52, $p < .01$; AoA < 15: r = -.71, $p < .01$; AoA > 12: r = -.27, $p < .01$; AoA > 15: r = -.23, $p < .05$).

DeKeyser (2000) used an adapted version of the grammaticality judgement test used by Johnson and Newport in a study with 57 Hungarian second language learners of English. These participants had arrived in the United States between ages one and forty and had lived there for at least ten years. The participants were divided in two groups: AoA < 16 and AoA > 16. DeKeyser found a significant correlation between age of arrival and scores on the grammaticality judgement test (r = -.63, $p<.001$) for all participants taken together, but not for the two separate groups (for AoA > 16: r = -.04; for AoA < 16: r = -.26). DeKeyser considers these results to constitute evidence in favour of a critical period. Bialystok (2002), however, argues that the results from DeKeyser do not show the discontinuity in the age function that should be present if there was a critical period, but that they are instead more consistent with a linear decline in language learning ability through the whole lifespan and therefore could be taken as evidence against a critical period for SLA.

Birdsong and Molis (2001) used the original materials from the Johnson and Newport 1989 study, but with 61 Spanish learners of English, divided into early arrivals (AoA ≤ 16; N = 29) and late arrivals (AoA ≥ 17; N = 32). Their learners had a similar educational background as the participants in the Johnson and Newport study (at least a BA degree), but their mean length of residence was higher (more than 10 years). They found a ceiling effect for learners with AoA < 16 (r = -.24, $p = .22$) and a significant correlation between AoA and performance on the grammaticality judgement task for the late arrival group (r = -.69, $p < .0001$). They also found an inflection point in the age function, but not around puberty or before, but rather at age 27.5. Both the post-maturational age effects and the native language effects found in this study can be considered as evidence for falsification of the CPH.

In 1991, Johnson and Newport (1991) did a study on the acquisition of Subjacency by Korean learners of English with different ages of arrival.

Subjacency is considered to be a part of Universal Grammar that regulates movement of wh-words in questions. Although there are differences between languages, certain restrictions are universal and the options that languages can use are considered to be innate. The aim of this study was therefore to see if similar results would be found for "universal properties of language, considered to be innate" (p. 215). In this study, in which Subjacency was compared to no-inversion sentences, which are considered not to be part of Universal Grammar, Johnson and Newport found a significant correlation between AoA and Subjacency scores (r = -.63, p < .001) with a significant linear trend (F = 25.95, p < .001 on a linear trend test with unequal intervals). Only the AoA 4-7 group did not differ significantly from the natives (t = 1.49, p < .1 (one-tailed)). In a comparison between the Subjacency items and the no-inversion items no significant differences were found for any of the age groups. Johnson and Newport conclude from these results again that there is a critical period for second language acquisition, even though these results do not show the discontinuity and plateau that are required for a critical period.

It is clear from these results that there is not enough evidence at the moment for the existence of a discontinuity in the relation between L2 proficiency and age of arrival with a flattening after puberty. Moreover, it is not clear that the correlation between age of arrival and L2 proficiency has a biological, (neurophysiological) cause.

2.4   Late learners with native-like proficiency in L2 syntax

Let us now turn to the second assumption related to the CPH, that no late second language learner should be able to attain a native level of proficiency. If there is a flattening after the terminus of the critical period, levels attained after this point should be low and significantly different from the level attained by native speakers and early learners. Although any level that is comparable to that of learners who started (long) before the presumed terminus would be counterevidence against a flattening and thus against the CPH, most researchers tend to concentrate on individuals who have reached a level of attainment that does not differ significantly from the level reached by native speakers of the target language. Newport, Bavelier and Neville (2001) point out that in all animal studies there is individual variation "particularly during the waning period or after" (p. 491), e.g. variation among zebra finches in the effects of heterospecific tutors (tutors from a slightly different species) and deafening on song maintenance. They also discuss studies with adjustment to visual prisms in barn owls to show that "re-learning or transfer during adulthood of skills experienced during early

life might be expected to show some resilience and success." (p. 491) This could account for a certain amount of success in the case of second language learning, because the learning task in L2 acquisition is based on a previous learning experience in the same domain. However, they do not give any examples from animal studies where some adult learners perform equally well on a learning task as animals that learned the same behaviour during the peak of the critical period.

Results from studies on native-like attainment in second language acquisition vary a great deal. For pronunciation, it has recently been shown that, at least for certain very advanced late learners under optimal learning conditions, it is possible to achieve a native speaker level in a foreign or second language (see Bongaerts, 1999; Bongaerts et al., 1997; Bongaerts, Mennen and Van der Slik, 2000; Moyer, 1999; Birdsong, 2003).

For morphosyntax, the picture is more complex. On the one hand, there are many studies in which all late learners investigated turned out to deviate in proficiency from native speakers (Johnson & Newport, 1989; Coppieters, 1987; Hyltenstam, 1992[3]). On the other hand, there are also studies in which certain late learners performed like native speakers on a grammaticality judgement test (Birdsong, 1992 and 1997; White & Genesee, 1996; McDonald, 2000). What makes the picture even more complex is that in some studies older learners seemed to be able to attain native-like proficiency whereas younger learners did not. Results from McDonald (2000 and personal communication, 23 March, 2004), for example, show that two adult Spanish learners as well as nine early Spanish learners of English performed like native speakers, whereas many young Vietnamese learners did not and had problems that were similar to those of (most of) the adult Spanish learners.

It is important to make a distinction between studies in which participants were only selected on the basis of their age of arrival and length of residence on the one hand and studies in which participants were carefully selected on the basis of their (high) level of proficiency on the other. In many of the studies in which participants were only selected on the basis of their AoA and length of residence and not on their proficiency level, results were only analysed at the group level and the question of whether there are individuals who reached a native level of attainment is not addressed. However, some of these studies are nevertheless informative with respect to the issue of native-like attainment for late learners. We will therefore first discuss some of these studies that are relevant for this issue and then proceed with the discussion of studies that specifically address this issue by looking at late learners with a very high level of proficiency in the target language.

---

[3]   In Hyltenstam (1992), the late learner group consisted of learners who arrived in the L2-environment between the ages of seven and twelve.

Age of arrival studies

In some of the studies in which participants were selected on the basis of their AoA and length of residence, rather than their L2 proficiency, it is clear that native-like attainment among the late learners was not found (e.g. Johnson & Newport, 1989). Some researchers did find late learners who performed within the native speaker range, however, without selecting them on the basis of their (high) proficiency (e.g. McDonald, 2000; Birdsong & Molis, 2001). This difference often coincides with a difference in the typological distance between the first and second language (no native-like attainment for studies with unrelated languages, e.g. Korean learners of English, and native-like attainment in studies with typologically closely related languages, e.g. Spanish learners of English; see also below).

McDonald (2000) used a grammaticality judgement test based on the one used by Johnson and Newport (1989) with early Vietnamese and Spanish learners of English (AoA ≤ 5), Vietnamese child learners of English (AoA 6-10) and late Spanish learners of English (AoA 14-20). Her group results show that the late Spanish learners and all Vietnamese learners differ significantly from the native speaker controls on accuracy. The reaction times show the same tendency, but the differences between the Vietnamese learners and the native speakers are not significant, probably due to the high variability among the Vietnamese early acquirers. The two Vietnamese learner groups did not differ significantly from each other on accuracy and the early Spanish learners did not differ significantly from the native speaker controls (both on accuracy and reaction times). McDonald's plot for the accuracy data of the Spanish speakers (p. 403) shows that there are five late learners who perform within the range of the early learners. Two of them also performed within the native speaker range. Within the Vietnamese group, there were six early learners and one child learner that fell within the native speaker range (McDonald, personal communication, 23 March 2004[4]). Birdsong and Molis (2001) found one late learner (AoA > 17) in their study that fell within the range of natives from the study of Johnson and Newport (1989). Flege et al. (1999) do not report individual results, but their plots for the grammaticality judgement test (p. 88) suggest that there are some late learners that fall within the native speaker range.

---

[4]   McDonald found slightly different results when considering only the ungrammatical sentences. The scores of the ungrammatical sentences were within the native speaker range for 7 early Spanish learners, one late Spanish learner, 6 early Vietnamese learners and no child Vietnamese learners.

Syntactic studies with near-natives

Considering these results, one might expect more similar outcomes from studies on syntax in which only participants with a very high level of proficiency were selected. However, in these studies, too, opposite results are found. In studies by Coppieters (1987) and Hyltenstam (1992), for example, the scores of all of the very advanced late learners differed from those of the native speakers. Note in this respect, that in Hyltenstam's (1992) study all learners started acquiring their second language before puberty. In other studies with very proficient late learners, however, e.g. Birdsong (1992), White and Genesee (1996) and Hyltenstam and Abrahamsson (2003), (quite) some participants fell within the range of the native speakers for syntax. Ioup, Boustagui, El Tigi and Moselle (1994), finally, found mixed results: they found native-like attainment for a translation task and an anaphora interpretation task, but some non-native like responses on a grammaticality judgement task.

Coppieters (1987) selected 21 very proficient, highly educated, second language learners of French from different L1 backgrounds (ranging from Korean to Italian) who arrived in France after the age of 18. During an interview participants were asked for acceptability judgements (sometimes together with questions about meaning differences) on 107 sentences divided over nine different grammatical constructions, ranging from the place of the adjective to the A-over-A constraint. Coppieters' quantitative analysis revealed that all non-native speakers were over three standard deviations away from the native speaker mean ($p < .005$). Moreover, qualitative analyses (explanations about judgements) revealed that the non-native speakers applied different rules than native speakers for the same constructions. Coppieters also concluded that the largest differences between native speakers and non-native speakers are not in areas considered part of Universal Grammar (such as the A-over-A constraint), but in other constructions, for example the use of different tenses depending on the context.

Hyltenstam (1992) compared Spanish/Swedish and Finnish/Swedish bilinguals from two age of arrival groups (AoA ≤ 6 and AoA ≥ 7) to each other and to monolingual speakers of Swedish. He analysed their oral and written data on several aspects, including error frequency. Hyltenstam does not report any quantitative figures about individual participants, but he claims that "All the individual subjects in the late AoA group had an error frequency well above that exhibited by any single subject in the monolingual group." (p. 365).

Birdsong (1992) tested 20 very advanced, highly educated English learners of French on their acceptability judgements (using a think-aloud protocol) for seven different grammatical constructions in French, such as

adjacency, the use of *il/elle* versus *ce* and *de* + modifier. Their age on arrival ranged from 19 to 48 years and their mean length of residence was almost twelve years (3-36). He found a strong relation between performance on his test and AoA and there were 15 participants who fell within the native speaker range.

White and Genesee (1996) tested 89 learners of English from various L1 backgrounds on their command of Subjacency. They compared near-natives and non-natives with different ages of arrival in Canada with a control group of 19 monolingual native speakers of English. Second language learners were assigned to the near-native group if they were judged to be virtually indistinguishable from native speakers with respect to pronunciation, morphology, syntax and choice of vocabulary by two native English-speaking judges. White and Genesee used a timed grammaticality judgement test in which participants had to indicate whether or not a sentence was possible in English, in combination with a question formation task. Group analyses for proficiency groups (AoA groups taken together) revealed that both learner groups performed quite well on Subjacency, although the non-natives had significantly slower reaction times. The only construction for which the near-natives performed significantly differently from the natives was extraction from non-finite Complementizer Phrases (CPs). Comparison of AoA-groups showed that there were no main effects for AoA for either accuracy or reaction time, even when only the highest AoA group (AoA>16) was compared to the native speakers. It should be noted that most participants in this study had an L1 in which Subjacency works in more or less the same way as in English. White and Genesee note in this respect, however, that Johnson (1988) did not find native-like success for Spanish learners of English for Subjacency and that White and Juffs (1998), using the same task with highly proficient Chinese-speaking adult learners of English, found "that there were few differences between Chinese speakers and English native speakers on the ungrammatical sentence types" (White & Genesee, 1996: 261).

Ioup et al. (1994) studied two near-native English learners of Egyptian Arabic on pronunciation, accent identification abilities and grammar. Both participants started learning Arabic in adulthood and had been resident in Cairo for at least ten years. One of them had studied Arabic whereas the other had learned Arabic without any formal instruction and could not read or write the language. For grammar, they were given an anaphora interpretation task, a translation task and a grammaticality judgement task with UG and language specific constructions, such as wh-questions, relative clauses and conjoined NP word order. One of the participants (the one who had studied Arabic) scored fully native-like on the anaphora interpretation task and on the translation task (except for one error

with a preposition), but deviated from the native speaker norm on six items in the grammaticality judgement task. It should be noted in this respect that, for all of these items, there were also some native speakers who deviated from this norm.

Recently, some researchers, e.g. Hyltenstam and Abrahamsson (2003), have addressed the question of whether comprehensive nativelikeness is attainable for late second language learners. Hyltenstam and Abrahamsson (2003) tested near-native speakers from different L1 backgrounds (ranging from German to Arabic) and non-native speakers of Swedish with different AoAs on three tests: a cloze test, a grammaticality judgement test and a white noise test (verbatim repetition of sentences presented with increasing amounts of white noise, i.e. sound waves with a uniform frequency spectre). For the white noise test, only exact repetitions were counted as correct imitations. None of their participants, including very early learners, fell within the native speaker range (of actual scores) on all three tests. They had one Estonian participant, though, who arrived in Sweden at the age of fourteen who fell within the native speaker range on their grammaticality judgement test. They view their results as evidence in favour of a critical period for second language acquisition. However, under the assumption that there is a critical period as outlined above, one should expect performance to be much better related to AoA, i.e. the early learners should all have been better than the late learners, which was not the case. Moreover, if one assumes that there are different critical periods for different areas of language, rather than one "overall" critical period for SLA, the fact that Hyltenstam and Abrahamsson found a participant who performed native-like on the grammaticality judgement test could just as well be considered as evidence against a critical period for syntax in SLA.

Methodological comments
A methodological problem in most studies in which no differences were found between native speakers and (some) very advanced late learners concerns the structures that were investigated. In order to test whether attainment of a complete native-like grammar is possible after a presumed critical period, it is important to include structures that are among the most difficult ones to be acquired. Moreover, it is also important that participants do not have rules for the constructions under consideration that they have learned by heart and which they might apply in a conscious task (e.g. a grammaticality judgement task) without really having acquired the rule. From this point of view, tests like the one used by Johnson and Newport (1989), in which very straightforward well-known and taught rules (e.g. with respect to morphology) are used, are less suitable for testing the CPH.

On the other hand, in studies with opposite results, it is often far from clear whether the participants were really among the best late learners of the target language. A related problem is the selection of native speaker controls. Should they be representative of the total target language population, or be as comparable as possible to the non-native speakers investigated? Coppieters (1989), for example, compared his very highly educated second language learners to a control group with various backgrounds, ranging from caretakers to linguists. He included ten spouses or close colleagues in his control group and people from different regions in order to make sure that his control group would be representative of the input his second language learners might receive. Birdsong (1992), however, criticized this selection procedure (Birdsong, 1992, p. 715-716) and used a control group with the same educational background as his learners instead.

Differences in participant selection procedures (both for late learners and native speaker controls) make it difficult to interpret results suggesting that there are no late learners who attain a native level of proficiency.

2.5   The role of the mother tongue

As we have argued in paragraph 2.2, the effect of the (psycho)typological distance between the L1 and the L2 should not interact with the relation between age of arrival and L2 proficiency. This means that the beginning of the offset and the terminus should be at the same age for all L2 learner groups, independent of the (psycho)typological distance between the L1 and the L2 (see e.g. Birdsong & Molis, 2001). Since after the terminus native-like attainment should no longer be possible, this also means that native-like attainment in L2 learners should not be found for any language pair.

If one looks at the Johnson and Newport (1989) study and some of the (partial) replications of this study (Bialystok & Miller, 1999; McDonald, 2000; Birdsong & Molis, 2001), it is clear that the relation between age of arrival and proficiency does not look the same for different language pairs. McDonald (2000) compared Vietnamese learners of English with Spanish learners of English. The two Vietnamese learner groups (AoA < 5 and AoA 6-10) differed significantly from the native speakers on accuracy (unlike the Korean and Chinese learners in Johnson and Newport's 1989 study), but did not differ significantly from each other, whereas the early Spanish learners (AoA < 5) did not differ significantly from the native speaker controls (both on accuracy and reaction time). This suggests that the age function for the Viernamese learners looks rather flat for AoA < 10 (but not because of a ceiling effect), whereas Johnson and Newport (1989) found a rather strong correlation between AoA and proficiency until puberty.

Bialystok and Miller (1999), who tested both Chinese and Spanish learners of English, found a significant difference in accuracy between the youngest group of Chinese learners of English and the native speakers (t=2.47, *p*=.04), whereas for the Spanish learners of English only the three older learner groups differed significantly from the native speaker controls. Birdsong and Molis (2001), in a replication of Johnson and Newport's (1989) study with Spanish learners of English (instead of Chinese and Korean learners) found a ceiling effect for learners who arrived in the United States before the age of 16 and an inflection point in the age function at age 27.5. Both these studies seem to indicate that the age function looks different depending on the typological distance between the L1 and the L2: for related languages (such as Spanish and English) the offset and terminus (in as far as these terms are appropriate here) occur much later than for totally unrelated languages (such as Chinese/Korean and English). This suggests that the shape of the age function is not (exclusively) determined by biological factors, as one would expect if a critical period exists for this type of learning.

With respect to native-like attainment, there is more and more evidence suggesting that native-like attainment in different areas is possible for learners of a related language, but it is less clear what the upper limits are for learners of languages totally unrelated to the mother tongue.

In some pronunciation studies, L2 speakers were found who performed like native speakers (e.g. Bongaerts, 1999; Bongaerts et al., 2000; Moyer, 1999; Birdsong, 2003). In all these studies, all learners who were indistinguishable from native speakers spoke a first language that was typologically rather closely related to the target language. Ioup et al. (1994), however, studied two very proficient late English-speaking learners of Egyptian Arabic, a language typologically very distant (and phonologically very different) from English. The results of the analysis of production data were mixed. Both learners were judged to be native by most judges, but there were also a few judges who could tell by certain features of their speech that they were non-natives.

In most other studies of morphosyntax with near-natives, the influence of the typological distance between the first language and the target language of the participants is unclear. Birdsong (1992) used only participants with a first language that is relatively closely related to the target language (English learners of French). Coppieters (1987) and White and Genesee (1996) tested participants with different L1-backgrounds, but this variable was not a systematic part of the research design. White and Genesee (1996) do not even provide information on what other language backgrounds are included in their study besides Germanic and Romance, which are typologically related to the target language of the study (English). Ioup et al.

(1994) tested near-natives with a typologically non-related L1. However, since they only had two subjects and the results were mixed, their study does not provide an unequivocal answer to the question of whether native-like attainment is possible for late learners with a non-related L1.

2.6   Neuroimaging studies on age and proficiency in L2 acquisition

To address the issue of whether there are neurophysiological changes in the brain before puberty that can account for the general tendency that early learners attain a higher level of proficiency than late learners, neuroimaging techniques seem very suitable. If late learners are shown to process their second language in different brain areas than early learners, for example, this could support the CPH for L2 acquisition. It should be noted, however, that most studies do not focus on syntax and that the interpretation of results from neuroimaging studies in general is difficult at this stage, because we still know very little about the exact function of brain areas involved in language production and perception and about how exactly the brain works with respect to language. We know even less about the neurological processes involved in language acquisition. Moreover, certain methods may not be fine-grained enough to reveal subtle differences. Some results of recent neuroimaging studies with second language learners nevertheless provide an interesting perspective on relevant issues with respect to the CPH for second language acquisition.

Abutalebi, Cappa and Perani (2001) give an overview of perception and production studies that have been done with bilinguals using PET (positron emission tomography) and fMRI (functional magnetic resonance imaging) scans. Both PET and fMRI are indirect measures of the amount of blood that flows to different areas in the brain. Both techniques compare the amount of blood flow during a test and a control condition (e.g. looking at pictures versus looking at a blank screen). In PET studies this is measured by probing radioactive tracers in the blood. In fMRI studies magnetic changes are detected resulting from changes in local blood oxygenation levels. If there is more blood in a certain area of the brain during the test condition than during the control condition, this is interpreted as activation of this brain area.

   Both techniques are quite accurate at localising brain activation, but they are not very accurate at establishing at which time the activation occurs. The temporal resolution for fMRI is better, though, than for PET. Due to this problem with the temporal resolution, only tasks in which the same function is performed for at least three seconds can be used. This makes it difficult (if not impossible) to look at syntactic processes, for example. Another

advantage of fMRI studies is that no radionuclides are used, which makes it possible to repeat experiments several times with the same participants. A disadvantage of fMRI is that the air in the middle ear and the mastoid bones creates interference with the magnetic field, resulting in loss of visualisation of certain areas (see also e.g. Abutalebi et al., 2001).

Of the six production studies Abutalebi et al. (2001) review, there are two that compare early and late bilinguals (Kim, Relkin, Lee & Hirsch, 1997, and Chee, Tan & Thiel, 1999) and both use fMRI.

Kim et al. (1997) used twelve proficient bilinguals with different language pairs. However, they did not assess the participants' proficiency. Six participants had been exposed to both languages during early infancy, whereas the other six started learning their L2 after puberty. The task that was used involved the description of activities using covert speech. This is, of course, a task in which many different aspects of language are involved (e.g. semantics, lexical access, syntax), as well as cognitive abilities (such as memory) that are not specific to language. The main result of this study was that, in late learners, L1 and L2 were represented in different parts of Broca's area and in the same regions in Wernicke's area, whereas for early learners overlapping parts were activated in both areas for both languages.

Chee et al. (1999) on the other hand, did not find different patterns of brain activation for both languages in late bilinguals. They compared fifteen early Mandarin-English bilinguals (L2 acquisition before age six) to nine late Mandarin-English bilinguals (L2 acquisition after age twelve), all living in Singapore. According to Abutalebi et al. (2001), this is a community "in which bilingual speakers can be expected to be highly proficient in each language" (p. 183). The task they used was a word generation task, in which words were cued by a visually presented word stem. They found a similar pattern of brain activation in the left prefrontal cortex for both languages and both groups, in spite of differences in writing systems between Mandarin and English.

As Abutalebi et al. suggest, this difference in results between Chee et al.'s study and Kim et al.'s study might be due to differences in proficiency between the late learners in both studies. However, the differences in the results might also be due to the differences in the tasks that were used. Yetkin, Yetkin, Haughton and Cox (1996) also did an fMRI study of language production in multilinguals that speaks to this issue, in spite of some methodological problems (see Abutalebi et al., 2001). They tested multilinguals (with different language combinations) that were fluent in a second language but not in a third language on word generation (phonemic verbal fluency) in their L1, L2 and L3. They found more extended activations for the languages in which the participants were less fluent, suggesting that proficiency (or exposure) is an important factor. On the other

hand, they also found activation in the left prefrontal cortex for all the participants' languages, including the non-fluent L3. This suggests, that this area is not very sensitive to proficiency differences.

Fortunately, the comprehension studies that Abutalebi et al. (2001) review are easier to compare. Of the four studies in which late bilinguals are tested, three use a task in which participants have to listen to stories (Perani et al., 1996; Dehaene et al., 1997 and Perani et al., 1998). Dehaene et al. (1997) tested 8 late French-English bilinguals with a low level of proficiency in an fMRI experiment. They found rather different activations for L2 compared to L1, including activation in the right hemisphere and a great deal of individual variation in activation for listening in the L2. Perani et al. (1996), did a similar experiment in a PET study. In this study, they tested 9 Italian-English late bilinguals with a low level of proficiency. In addition to listening in the L1 and L2, they also measured brain activity during listening in an unknown language (Japanese). Whereas the activation pattern for English and the unknown language was the same (activation in the left and right superior and middle temporal areas), the activation for listening to the L1 was more extensive.

In a follow-up PET study, Perani et al. (1998) compared early and late L2-learners with a high level of proficiency. Twelve Spanish-Catalan bilinguals who acquired Catalan before the age of four were compared with nine Italian-English bilinguals who acquired English after the age of ten. In this study, the activation pattern was similar for both languages and for both types of learners. Moreover, a comparison between the late highly proficient Italian-English bilinguals from this study with the late Italian-English bilinguals with a low level of proficiency from Perani et al. (1996), showed that the highly proficient bilinguals had activation for L2 in the temporal poles and in the left anterior and posterior part of the middle temporal gyrus, whereas the late bilinguals with a low level of proficiency did not.

These results clearly provide evidence "of considerable plasticity in the network that mediates language comprehension in the bilingual brain" (Abutalebi et al., 2001: 186). They also strongly suggest that proficiency is more important as a determinant of cortical representation than age of onset of L2 acquisition.

This conclusion can also be drawn from an ERP (event related brain potential) study by Ardal, Donald, Meuter, Muldrew and Luce (1990). In ERP-studies, electrical brain activity is measured by placing a number of electrodes on a participant's head (the number of electrodes used differs across studies). The signals from these electrodes are displayed in an EEG (electroencephalogram). By time-locking the EEG wave pattern to specific events that occur in the test an ERP is obtained. In order to be able to interpret the brain waves, it is necessary to average over many trials with the

same stimulus type. For the interpretation of the data the ERP of the test-condition is always compared to that of a control condition (e.g. a syntactically incorrect sentence versus an equivalent correct sentence), with respect to polarity (positive or negative), length, amplitude and time. Compared to PET and fMRI, ERP has a good temporal resolution, but is much less accurate at locating brain activation. Because of this, ERP is mostly used to detect more subtle syntactic or semantic processes. Three important linguistic ERP components that have been found for native speakers for these domains are the N400, the P600 and the ELAN (early left anterior negativity). The N400 is a negative deflection in the brain wave at around 400 milliseconds (ms) after the presentation of a semantically anomalous word. The P600 is a positive deflection in the brain wave at about 600 ms after the presentation of a syntactic anomaly. The ELAN is an early negative deflection on the left front side of the brain, which is sometimes found in syntactic conditions.

Ardal et al. tested 12 early (age of onset 3-10) and 12 late (age of onset 13-17) proficient learners of French and English from different L1 backgrounds (mainly French and English) and 24 English monolinguals on their response to incongruent and congruent sentence final words. They "took care to obtain bilinguals with roughly equivalent language fluency or competence at each age of acquisition level" in order to prevent "the strong confounding of current fluency with age of acquisition which very often occurs in the bilingual." (p. 203)

The N400-effect occurred earlier in the monolinguals than in the bilinguals and later in the bilinguals' L2 than in their L1. There was also a left-right parietal asymmetry between the monolinguals and the bilinguals and the bilinguals had reduced frontal negativities for their L2 compared to their L1. However, this was not the case for a group of six highly fluent French-English bilinguals. There was no difference for either the L1 or the L2 between early versus late bilinguals. These results strongly suggest that the differences between the monolinguals and the late bilinguals are caused by the difference in having one versus two languages and differences in proficiency rather than by age of onset of acquisition.

It should be noted that all the tasks used in the studies discussed above do not specifically test syntax. Moreover, differences in activation for syntactic processing may be invisible if there is an overlap in regions (or electricity effects) involved in syntax and other areas of language. Syntax does play an important role, however, in several other ERP studies.

Hahne and Friederici (2001) looked at syntactic and semantic processing in late Japanese learners of German (AoA 18-31) in comparison with native speakers of German. The learners had learned German in formal

settings for 4-60 months (mean 30 months). Their self-estimated proficiency on a six-point scale (1 = no knowledge, 6 = equivalent to native speaker) was 3.5 on average. Participants were presented with correct sentences, semantically incorrect sentences, syntactically incorrect sentences and sentences that were both semantically and syntactically incorrect. They had to judge the sentences for linguistic integrity, while their EEG was recorded from 25 scalp sites.

The L2-learners scored above chance on all conditions. They made most errors in the syntactically incorrect sentences. The ERP-data show differences between the L2-learners and the native speakers in all conditions. For the correct sentences, the ERP-pattern of the L2-learners showed similarities with the ERP-pattern for incorrect sentences obtained from native speakers, suggesting that "the processes which L2-learners have to conduct in order to understand a correct sentence resembles those performed by native listeners during the processing of syntactically incorrect sentences" (p. 129). For the L2-learners, the ERP-pattern for the syntactically incorrect sentences was very similar to that of the correct sentences and there was no clear P600-effect. Furthermore, there was a larger group effect in the syntactic condition on the right than on the left hemisphere.

Weber-Fox and Neville (1996) also compared syntactic with semantic anomalies in an ERP-study with 61 Chinese-English bilinguals with at least five years of immersion. They divided their participants into five age of arrival groups (AoA 1-3, 4-6, 7-10, 11-13, and > 16). Participants were not selected on the basis of their proficiency and the older learners' proficiency in English was much lower than that of the early learners.

Whereas for the semantically incorrect sentences all participants who had arrived before the age of eleven performed like native speakers, there was a linear decrease across AoA groups for the syntactically incorrect sentences in the behavioural data. For the semantic condition, the N400 was later for the highest AoA group. All other groups had a normal N400 pattern. For the syntactic condition, the ELAN found for monolinguals was absent in the learners. According to Weber-Fox and Neville, however, this might be due to the small number of participants in the learner groups. The P600 effect was delayed for the AoA 11-13 group and absent for the oldest learners.

Sabourin (2003) looked at processing of gender by advanced late learners of Dutch in an ERP-experiment. She compared a control group of 23 native speakers of Dutch with 14 native speakers of German, 8 native speakers of a Romance language and 9 native speakers of English on grammatical gender and gender agreement in their L2 Dutch. The learners' exposure to Dutch ranged from 3 to 32 years. They scored between 76% and 100% on sentences in an on-line grammaticality judgement task testing

finiteness (the difference between infinitives and participles) and subject-verb agreement. This was used as a proficiency criterion to make sure that participants were very proficient in Dutch. However, the regular morphological markings for both finiteness and agreement are easy to learn by heart, they are discussed in most Dutch courses and finiteness and subject-verb agreement are also morphologically marked in the source languages in this study. Being able to perform well on such sentences, therefore, does not necessarily imply a high level of proficiency.

In the ERP-experiment, a grammaticality judgement test was used in which sentences were presented visually word by word and participants had to press a button to indicate whether a sentence was grammatical or ungrammatical. To test the participants' command of grammatical gender, four conditions were used: an NP definite condition, in which the right or wrong definite article occurred, an NP indefinite condition, in which the adjective was presented with or without inflection for gender agreement, an RP definite condition, in which the relative pronoun did or did not agree with the definite antecedent and an RP indefinite condition, in which the relative pronoun did or did not agree with the indefinite antecedent.

There were significant differences in the behavioural data between the native speakers of Dutch and all L1 groups on all conditions, including both proficiency conditions. For the test conditions the L1 German group performed significantly better that the two other L1 groups, who scored at chance level for the ungrammatical sentences. For the proficiency conditions the differences between groups were not significant. Considering these results, it is not very surprising that there were also differences between the native speakers and the learner groups in the ERP-data.

For the native speakers and the German group, the ERP-data were only based on the sentences that were responded to correctly. However, because of the poor performance of the Romance and English groups on the behavioural task, sentences to which they responded incorrectly were included in their ERP-data. This makes the comparison between these groups and the native speakers possibly unreliable. The German group had a P600 effect for all conditions, except for the NP-indefinite condition, for which they had a positivity with an atypical distribution. Only for the finiteness and NP-definite conditions, the P600 was similar to that of the native speakers. For the other four conditions, the effect was more delayed, more restricted in distribution and/or had a lower amplitude. The Romance group had a P600 on the finiteness condition (although delayed and with decreased amplitude), a possible (late) P600 for the relative pronoun (RP) conditions and no P600 for the NP- and subject-verb conditions. The English group had a P600 effect (although different from that of the NSs) for the finiteness and subject-verb agreement conditions, a possible late and

restricted P600 for the RP-definite condition and no significant positivity for the RP-indefinite condition.

Sabourin concludes from these data that native-like processing is only possible for grammatical features that are very similar in the L1 and the L2. However, because of the problems with the proficiency criterion and the lack of analysis of individual learners, this conclusion seems unfounded on the basis of these data.

Since the late learners in these ERP-studies were not highly proficient, it is hard to tell whether the processing differences for syntax between late learners and monolinguals are specific to syntax or due to proficiency differences. The results do show, however, that the influence of age in SLA is not the same for syntax and other domains, such as semantics.

Brain plasticity
Pallier et al. (2003) address the issue of brain plasticity and the role of the L1 by looking at adults who were adopted as a child and who reported having completely forgotten their first language. They tested eight Korean participants who were adopted by French speaking families between the ages of 3 and 8 and who had probably had normal exposure to Korean before that age. According to Pallier et al., seven of them had no perceptible foreign accent in French. The control group consisted of eight native monolingual French speakers who had had no exposure to any Asian language.

Participants were given three tests. In the first test participants had to listen to sentences in Korean, Japanese, Polish, Swedish and Wolof and indicate on a seven-point confidence scale whether they thought sentences were in Korean or not. In the second test, participants had to decide which of two orally presented Korean words was the correct translation of a visually presented French word. Both of these tasks were presented out of the scanner. The third task was an event-related fMRI experiment in which participants had to listen to French, Korean, Japanese and Polish sentences and fragments and decide whether a fragment had appeared in the sentence or not.

The behavioural data revealed that the Korean participants did not differ significantly from the French controls and that both groups were unable to recognize Korean sentences or words. There were no Korean adoptees who performed markedly differently from the control group. The fMRI-data revealed no significant differences between the Korean and French participants when French was compared to Polish or Japanese. However, there was a difference in activation for Korean versus Polish: for the French participants, the activation for Korean versus Polish was stronger in the right superior temporal sulcus and the left cerebellum. It should be noted, though, that these differences were not found for the Korean-Japanese

and Japanese-Polish contrasts. Moreover, individual analyses showed that the extent of activation for the French stimuli relative to Polish stimuli was greater for the French participants than for the Korean participants. Whether this is due to a greater activation of French in the French participants or a greater activation of foreign languages in the Korean participants cannot be determined. In the individual analyses, none of the Korean participants showed any Korean-specific activation.

According to Pallier et al. these results provide evidence "in favor of the reversibility of plastic changes associated with language acquisition in the first few years of life" (p.158) and against the crystallisation hypothesis, according to which "a window of brain plasticity is open at birth and progressively closes as the brain networks for language become stabilized, under the possible influence of maturational and/or experiential factors" (p. 160). They are, however, compatible with an interference account, according to which "the presence of processes and representations attuned to the first language acts as a filter that distorts the way a second language can be acquired" (p. 160).

If we consider all these results from neuroimaging studies with bilinguals together, it seems that a second language is initially stored in different areas in the brain and processed differently than the L1. However, with increasing proficiency, the storage and processing of the L2 seems to become more similar to that of the L1, especially for domains such as semantics. At the same time, the processing of the L1 seems to be affected by the L2 (Ardal et al., 1990). In most studies, increasing age of onset of acquisition goes together with increasing entrenchment and input of the L1, and the effect of maturation cannot be separated from the interference effect of the L1. However, when L1 input is no longer available for a long period of time, as was the case for the participants in Pallier et al. (2003), the L2 seems to be able to take over the role of the L1. Further research with more subtle tests and perhaps more subtle measuring techniques is necessary to determine whether or not there are differences between the storage and processing of the L2 by highly proficient late learners and the L1 by monolinguals (or bilinguals) in the area of syntax.

2.7   Cause of age effects in second language acquisition

In paragraphs 2.3-2.5, we reviewed the literature on the influence of age of onset on second language acquisition from the point of view of the Critical Period Hypothesis. In this paragraph, we will discuss some alternative explanations for the observed age effects. There are many factors that may

play a role in SLA that are presumably different for adults than for children, for example anxiety, acculturation, motivation and attitudes towards the new country, culture and language (see e.g. Bialystok & Hakuta, 1994). We will limit ourselves, however, to the role of language input, language use and working memory.

Language input and language use
The input that young children receive is quite different from the kind of input for adults. Children (in cultures that are studied in SLA research) receive input that is adjusted to their cognitive and linguistic abilities. They have ample time to pay attention to the form of the input and practice sounds and structures before they need to perform complex communicative tasks. Adults in a second language environment, on the other hand, often receive input that is not specifically meant for second language learners, and they have to perform complex communicative tasks while learning the L2. Although there is a great deal of variation in the amount of input that second language learners (especially adults) receive, younger learners gradually receive more input in the L2 and less in the L1 than older learners (see e.g. Jia, 1998; Jia & Aaronson, 2003).

In order to test whether there is a critical period during which people are more sensitive to language input, the amount of input in the L2 (ideally) should be kept constant. In most CPH studies in which participants are not selected on the basis of their proficiency, participants are selected on the basis of their length of residence in the L2 environment. In most of these studies, years of residence are either kept constant across groups or set at a minimum of, for example, five years. If it is indeed the case that early learners receive more and better target language input in the same number of years than late learners (as is shown by Jia, 1998, for example), a difference in results between late learners and early learners or native speakers may merely reflect differences in input and not be very informative with respect to differences in sensitivity to this input. In order to compare group results from second language learners at different ages of arrival, a finer measure of L2 input should be used to control for the amount of input that learners receive. Unfortunately, it is practically impossible to control for the qualitative differences in input between early and late learners in a natural environment. This can be done to some extent, though, by looking at the relation between certain demographic variables, such as attending school in the L2 environment, level of education, first language of the spouse etcetera, as has been done by Stevens (1999), for example.

Another problem in comparing late learners with early learners is the interaction between the use of the L1 and the use of the L2. When children start acquiring a second language at a very early age, their L1 has not fully

developed yet and is not as entrenched as it is for older learners. Because of this, there is probably less interference from the L1 and as a consequence it is easier for early learners to reach a high level of proficiency in the L2. Flege, Frieda and Nozawa (1997) tried to separate the influence of maturation from the influence of use of the L1 for L2 accent. They compared accent ratings of two groups of early Italian learners of English (mean AoA = 5, LoR > 18): one group who used Italian 3% of the time on average and one group who used Italian 33% of the time on average (percentages were based on self-report). Both groups had a detectable foreign accent. However, the learners who used Italian more had a significantly stronger foreign accent than the low use group. The study by Pallier et al. (2003), described in the previous paragraph, also strongly suggests a high negative correlation between use of the L1 and level of proficiency in the L2.

In studies on syntax, differences in L1 and L2 use between learners might also account for (part of) the variation in proficiency sometimes found in groups of early learners, for example in the group of Vietnamese early learners in McDonald (2000; see also paragraph 2.4), particularly with respect to reaction times.

Working memory
Another important alternative explanation for the CPH can be found in the "less is more" or "starting small" hypothesis (see e.g. Newport, 1990; Newport, 1991; Elman, 1993; Pitts Cochran, McDonald & Parault, 1999). Proponents of this hypothesis claim that age effects in language acquisition result from changes in working memory with maturation. These changes are also biologically determined, but they are not specific to language. The decrease in language learning ability, according to these authors, is due to an increase in working memory capacity. Having a small working memory capacity forces one to process small units at once and this is claimed to help children to focus on details, such as specific morphemes. Adults, on the other hand, have a larger working memory capacity and try to analyse large parts at once. Due to the complexity of these larger units, details such as the phonological content of specific morphemes get lost.

In an attempt to test this hypothesis, Pitts Cochran et al., 1999, conducted two experiments in which they compared two groups of English-speaking adults learning complex ASL (American Sign Language) verbs from a video-tape. The experimental groups had to count tones while learning the verbs, which limited their memory capacity for the language learning task. The control groups learned the verbs in silence. Although the control groups were better at reproducing signs they had learned, the experimental groups were better at combining morphemes from learned signs into new signs. This experiment suggests that having less working

memory capacity available helps in learning morphology (at least in the initial stages of acquisition). It also suggests that under special circumstances the disadvantages of having a larger working memory can be compensated for. It could also be argued, however, that the differences in this study are caused by a difference in conscious attention, rather than by a difference in working memory capacity (Bas Evers, personal communication, July 2001). In experiments such as the above, these causes are difficult to tease apart, because the introduction of an extra task leads both to a reduced availability of working memory resources and to a reduced amount of conscious attention for the task under investigation.

Kersten and Earles (2001) found an advantage in learning word meaning and morphology in a simple artificial language for adults starting with simple input over adults starting with complex "sentences". Elman (1993) found the same advantage together with an advantage for having a reduced memory capacity for simple recurrent networks in a computer learning relative clauses. In this case, it does not seem very plausible that the results are due to differences in conscious attention, rather than to differences in working memory capacity.

Rohde and Plaut (1999) tested the starting small hypothesis by making a simple recurrent network learn English. They found that the advantage of starting small, found by Elman (1993), disappears if a simple recurrent network can make use of semantic information. In fact, the stronger (more English-like) the semantic constraints were, the more advantageous it was for the network to start with input that represented the full complexity of the language. For early memory restrictions, Rohde and Plaut (1999) found neither an advantage nor a disadvantage.

These studies show that the effect of working memory on acquisition and its relation to other factors (e.g. attention) is not clear enough yet and calls for further research. Moreover, the starting small hypothesis cannot easily explain the advantage of early learners for constructions involving large linguistic units, such as Subjacency. Nevertheless, it seems plausible that working memory skills might play a role in the acquisition of the grammar of a second language, especially in the case of phonologically non-salient items that do not contribute much to the meaning of a sentence, such as inflectional morphology and certain function words (e.g. dummy subjects). The fact that the maturation of working memory capacity is biologically determined, but not linguistic in nature and its effect on language learning can be compensated for by other things, makes it an interesting factor to consider when looking at age effects in second language syntax.

2.8   The present study

In the present study, we test the hypothesis that there is a critical period for the acquisition of L2 grammar by investigating the assumption that no late language learner should be able to acquire a native level of proficiency in L2 grammar. We also look at the role of the L1 for near-native learners.

In order to overcome the problems mentioned in paragraphs 2.4 and 2.5, we used a methodology that differs in two important ways from earlier research in this area. First, we systematically varied the typological distance between the target language and the first language background of our participants. In our study we tested late learners of Dutch from three different L1 backgrounds: German (a Germanic language, like Dutch), French (a Romance language) and Turkish (an Altaic language, which is typologically very different from Dutch and shares no areal features with Dutch). The degree of difference between these languages and Dutch in the structures we used reflects the genetic (and areal) differences between those languages and Dutch.

Second, we chose constructions (namely dummy subject constructions) that are known to be very difficult for learners of Dutch and for which hardly any rules can be found in school books or other materials that learners (even those with a linguistic background) may come across, because they are very complex and ill understood (see chapter 3). This means that even learners who have extensive formal training in Dutch can only acquire these structures on the basis of evidence from the input.

In other respects, we have followed the methodological practice of most of the research on ultimate attainment of very advanced late second language learners. We recruited learners who arrived in the Netherlands after the onset of puberty, when they were twelve years or older. This is the oldest age found in almost all studies that provide evidence for a critical period in first language acquisition and recovery from aphasia (see e.g. Lenneberg, 1967; Long, 1990). It is true that Patkowski (1980) and Johnson and Newport (1989) found a discontinuity at age fifteen, but Bialystok and Hakuta (1994) showed that in the Johnson and Newport study this result was due to the method of analysis applied. In a replication study, Bialystok and Miller (1999) only found a discontinuity around age eight. On the basis of the above results, it seems justified to adopt the age of twelve as the latest age at which an alleged critical period could still be in operation.

We selected only the most successful late learners that we could find for the given source-target language pairs with respect to grammar, i.e. learners that were reported to use (only) target structures the way native speakers use them by people who know them well. We gave these participants a sentence repetition task and a sentence preference task. Their results on these

tasks were compared to those of native speakers of Dutch. The native speaker controls were highly educated. Their educational background is similar to that of most of the learners. Moreover, data from a pilot study we did indicated that on the items we used in our tasks highly educated native speakers are more or less representative of the total Dutch native speaker population.

With this design, we will address the following research questions with respect to the CPH:

1)   Are there any late second language learners who fall within the native speaker range in their command of grammatical constructions that are known to be very difficult for second language learners and which can only be acquired on the basis of the input?
2)   How is the level attained in L2 grammar after the age of twelve related to the typological distance between the L1 and the L2?
3)   What are the input and background characteristics of late learners who perform within the native speaker range (if they exist)?

Chapter

3

Dummy subjects

In chapter two we presented criteria which constructions should meet in order to be used in a study on the Critical Period Hypothesis for syntax in second language acquisition. Such constructions should be very difficult to acquire for second language learners and learners should only be able to acquire them through processing input, i.e. there should not be simple rules available to second language learners that they can learn by heart and apply without really having acquired them. Dummy subject constructions in Dutch seem to meet these criteria. That is why we chose these constructions for our tests.

In this chapter, we will discuss previous literature on dummy subject constructions, the results on these constructions from our sentence preference task with native speakers of Dutch, and what German, French and Turkish learners have to acquire in order to behave like native speakers of Dutch with respect to dummy subject constructions.

We will first review the relevant literature on dummy subjects in Dutch from a purely descriptive perspective (§ 3.1). In paragraph 3.2 we will present the results from our sentence preference test with native speakers. Subsequently, in § 3.3, we will look at some generative analyses for dummy subject constructions. In paragraph 3.4, we will explore how dummy subject constructions, and distinctions that in our study turned out to play an important role in these constructions (in Dutch), are expressed in German, French and Turkish. Finally, we will discuss what native speakers of German, French and Turkish have to acquire with respect to dummy subject constructions in order to behave like native speakers of Dutch in our tests.

In the literature, different terms are used for dummy subject constructions and it is not always clear exactly what constructions researchers using the term "dummy subject" have in mind. Some clarification of what we mean when we refer to dummy subject constructions is, therefore, necessary. We use this term to denote constructions with a logical subject that does not occur in the normal syntactic subject position for semantic or pragmatic

reasons (e.g. heaviness). Instead, the syntactic subject position is filled by a dummy element, which can be *het* (it), *er* (there) or *0*. In this study, *0* is used in sentences in which *het* or *er* is absent, in cases in which an equivalent sentence with *het* or *er* is (theoretically) possible.

We distinguished three types of dummy subject constructions in Dutch:
- sentences with *er* or *0* in which the logical subject is a noun phrase/determiner phrase (DP-type)
- active sentences with *er*, *het* or *0* and a sentential subject (AS-type)
- passive sentences with *er*, *het* or *0* and a sentential subject (PS-type)

Because results from a pilot test with native speakers did not reveal clear differences between active and passive sentences of the DP-type we restricted this type to active sentences. An example of each type is presented in (3.1)-(3.3). In these examples (and all other examples from our study), the dummy subjects are represented by the symbol ◇ and logical subjects are underlined. To the right, we present preference judgements that are extracted from the results on the sentence preference test from the native speaker control group in our study. In these preference judgements, a preference for *0* to *er*, for example, is indicated by the symbol >. When sentences with *0*, for example, are judged as either better or equally good/bad as equivalent sentences with *er*, this is indicated by the symbols > and =. The judgements of the native speakers will be discussed in paragraph 3.2 below.

I   DP-type:
(3.1)   Men beseft niet altijd dat ◇ <u>een pinguïn</u> een vogel is.[5]   0 > er
             one realises not always that ◇ a penguin a bird is
             "One does not always realise that a penguin is a bird."
II   AS-type:
(3.2)   Meestal valt ◇ niet mee <u>om kaartjes voor een concert te krijgen</u>.
                                                                          het > 0 >= er
             usually falls ◇ not with to tickets for a concert get
             "Usually it is not easy to get tickets for a concert."

---

[5]   In sentences like these, it is not clear whether the logical subject is in the normal syntactic subject position or not. It can be argued that it is not, however, in similar sentences with a constituent in between the complementizer and the logical subject, as in:

i   Ik heb gehoord dat volgens haar een pinguïn geen vogel is.
         I have heard that according her a penguin not a bird is
         "I heard that according to her a penguin is not a bird."

III   *PS-type*:
(3.3)   In de krant wordt ◇ beweerd <u>dat hij dronken achter het stuur</u>
          <u>gezeten heeft</u>.   0 >= er >= het
          in the newspaper is ◇ claimed that he drunk behind the wheel sat
          has
          "It is claimed in the newspaper that he was drunk while he was
          driving."

For the different types there are different reasons for using a dummy subject. In sentences of the DP-type with a non-specific subject, the subject is generally not in topic position, because non-specific subjects are typically not topics. In sentences of the AS- and PS-type, sentential subjects (i.e. embedded clauses) occur at the end of the sentence, because heavy elements are preferred in sentence final position in most languages for processing reasons. Because of this, the syntactic position of the subject is supposed to be occupied by some other element, i.e. a dummy subject.

3.1   Traditional analyses of dummy subjects in Dutch

In this paragraph, we will focus on the descriptions in Haeseryn, Romijn, Geerts, De Rooij and Van der Toorn (1997) for the different dummy subject construction types, as this grammar provides the most elaborate description of dummy subjects in Dutch presently available. Haeseryn et al. describe in what kind of sentences dummy subjects can occur and in what sentence types and positions they tend to be obligatory or optional. We will discuss other proposals when relevant. In paragraph 3.4, we will discuss traditional analyses of the other languages involved in our study, German, French and Turkish, both with respect to dummy subject constructions and the factors that play an important role in this study.

DP-subjects
Haeseryn et al. (1997) discuss many factors concerning the acceptability of sentences with a DP-subject and "presentative *er*" (see paragraph 8.6.3 of Haeseryn et al.). They give a wide range of sometimes very subtle judgements for the examples they present, which we have divided below into two categories: a preference for *0* to *er* and a preference for *er* to *0*. Since these are preference judgements (as are the judgements we elicited in our study), they do not provide information about the grammaticality of sentences with the non-preferred dummy subject. In general, however, sentences with a preference for *0* cannot contain *er*, whereas in sentences with a preference for *er 0* is often also grammatical.

For each category we present the factors discussed in Haeseryn et al., together with one of their examples. The presentation of these examples is different from that of the other examples in this dissertation. The main reason for this is that in most of the examples from Haeseryn et al. *er* occurs in initial position, whereas in our study we mainly looked at dummy subjects in non-initial position. The reason for this was that in sentence-initial position, the dummy subject *0* is principally excluded. Haeseryn et al. often only present examples with *er* and not the equivalent sentences with *0*. It should be noted that in the examples in which *er* is in initial position the equivalent sentence with *0* has a different word order.

Factors that conspire against er (0 > er):

Perception verbs with a non-finite embedded clause:
(3.4)   Ik zag (*er*) <u>iemand</u> oversteken. (p.468)
         I saw there someone cross
         "I saw someone cross the road."
Categorical and generic subjects:
(3.5)   *Er* is <u>een zebra</u> gestreept. (p.469)
         there is a zebra striped
         "Zebras are striped."
Definite subjects:
(3.6)   *Er* staat <u>de auto</u> bij de schuur. (p. 470)
         there stands the car near the shed
         "The car is near the shed."
Predicates which strongly emphasize an activity:
(3.7)   *Er* besloop <u>een tijger</u> zijn prooi. (p. 472)
         there stalked a tiger his prey
         "A tiger stalked its prey."
Predicates in which the referent of the subject plays an active role:
(3.8)   *Er* sommeerde mij <u>een agent</u> af te stappen. (p. 472)
         there summoned me a policeman off to get
         "A policeman summoned me to get off my bike."
Questions for clarifications and riddles:
(3.9)   <u>Wat</u> is (*er*) rond en toch vierkant? (p. 473)
         what is there round and yet square?
         "What is round and yet square?"
Questions with a definite direct object:
(3.10)   Schrijft (*er*) <u>één van jullie</u> die brief? (p.474)
         Writes there one of you that letter?
         "Is one of you going to write that letter?"

Questions with a subject complement:
(3.11)   Wie van u is (*er*) de dokter? (p.474)
            who of you is there the doctor?
            "Which one of you is the doctor?"

Sentences with a definite object in initial position:
(3.12)   Al die foto's heeft (*er*) iemand bewaard. (p. 476)
            all those pictures has there someone saved
            "Someone has saved all those pictures"

Factors that favour er (er > 0):

Indefinite (non-categorical) subjects:
(3.13)   *Er* staat een auto bij de schuur. (p. 470)
            there stands a car near the shed
            "A car is near the shed."

Definite subjects in enumerations, and with repetition, habit or custom:
(3.14)   *Er* waren op de receptie aanwezig: de burgemeester, de wethouders
            en de meeste leden van de gemeenteraad. (p. 470)
            there were at the reception present: the mayor, the aldermen and the
            most members of the local council
            "At the reception were present: the mayor, the aldermen and most
            members of the local council."

Questions with an intransitive verb:
(3.15)   Wie komt *er* vanavond op dat feest? (p. 472)
            who comes there tonight at that party?
            "Who is coming to that party tonight?"

Questions with an adjectival subject complement:
(3.16)   Wie is *er* ziek? (p. 473)
            who is there ill?
            "Who is ill?"

Questions with an indefinite direct object:
(3.17)   Wie schrijft *er* een brief? (p. 474)
            who writes there a letter?
            "Who is going to write a letter?"

Sentences with *zijn* or *bestaan* with the meaning of *to exist*:
(3.18)   *Er* is/bestaat ook een aap zonder staart. (p.469)
            there is/exists also a monkey without tail
            "There are also monkeys without tails."

Subject with special emphasis in sentence-initial position:
(3.19)   Een studént kwam *er*! (p. 474)
            a stúdent came there!
            "It was a student who showed up."

Sentences with an indefinite object in initial position:
(3.20)   Foto's heeft *er* altijd wel <u>iemand</u> bewaard. (p. 476)
         pictures has there always someone saved
         "There is always someone who has saved pictures."

Many of these factors are related to specificity. In most examples with a preference for *0*, the entity referred to by the subject does not need to be introduced into the discourse, either because it is already present or because it is clear from the context what is meant. For most examples with a preference for *er*, the opposite holds: it is not clear to the listener in advance what subject the speaker is going to talk about. Therefore, introduction of the subject by means of presentative *er* is necessary.

        The examples given by Haeseryn et al. (1997) of predicates which strongly emphasize an activity (as in (3.7)), and predicates in which the entity referred to by the subject plays an active role (as in (3.8)) are often also transitive. Moreover, in transitive sentences subjects are more easily interpreted as specific. In these sentences it is therefore hard to tell exactly which factors determine the preference for *0* to *er*.

        For declarative sentences with the logical subject in initial position, Haeseryn et al. (1997) claim that *er* is optional in most cases. It should be noted here that from the dominant analysis within a generative perspective (presented in paragraph 3.3) such sentences could never contain a dummy subject and that *er* should always have a locative interpretation in these cases (see e.g. Bennis, 1987)

        A factor not discussed by Haeseryn et al. (1997), but mentioned by other researchers, is the specificity of the verb. Grondelaers, Speelman and Carbonez (2001) make a three-way distinction between verbs with respect to specificity: very specific verbs, such as *branden* (burn), *zitten* (sit) and *verschijnen* (appear), very unspecific verbs, such as *zijn* (be) and an in-between category, consisting of verbs such as *ontstaan* (arise), *blijven* (stay) and *heersen* (rule) (Grondelaers et al. 2001:22).

Sentential subjects in active sentences
For dummy subject constructions in active sentences with a sentential logical subject, it is much less clear which factors play a role. According to Haeseryn et al. (1997), *het* can occur in all sentence types: in main clauses without embedded clauses, in main clauses with both non-finite and finite embedded clauses, in cleft sentences, in "balansschikking"-constructions (complex sentences with the complementizer *of*, of which the first part contains a negative item and the second part has the word order of a main clause (van den Hoek, Houtman & Jullens, 1988: 6)), in pseudo-cleft sentences and nominal predicates with embedded antecedents (Haeseryn et al., 1997: 1133-1137).

Examples of each type are presented in (3.21)-(3.27).

Main clause without embedded clause:
(3.21)   *Het* valt niet te ontkennen: <u>zijn houding is de laatste tijd aanmerkelijk
         veranderd</u>. (p. 1136)
         it falls not to deny: his attitude is the last time considerably changed
         "It cannot be denied: his attitude has changed considerably recently."
Main clause with non-finite embedded clause:
(3.22)   *Het* is gezond <u>om veel te wandelen</u>. (p. 1135)
         it is healthy to much to walk
         "It is good for your health to walk a lot."
Main clause with finite embedded clause:
(3.23)   *Het* zou te betreuren zijn <u>als je je werk niet af kunt maken</u>. (p. 1136)
         it would to regret be if you your work not complete can make
         "It would be regrettable if you cannot finish your work."
Cleft sentence:
(3.24)   *Het* is om vier uur <u>dat de trein vertrekt</u>. (p. 1133)
         it is at four hour that the train leaves
         "It is at four o'clock that the train leaves."
"Balansschikking":
(3.25)   *Het* duurde niet lang <u>of het onweer brak los</u>. (p. 1135)
         it took not long or the thunderstorm broke loose
         "It did not take long for the storm to break."
Pseudo-cleft sentence:
(3.26)   *Het* is niet zo dat we nu al maatregelen gaan nemen. (p. 1137)
         it is not so that we now already measures go take
         "It is not the case that we are going to take measures right now."
Nominal predicate with embedded antecedent:
(3.27)   *Het* is namaak <u>wat je daar ziet</u>. (p. 1137)
         it is imitation what you there see
         "It is an imitation that you see there."

In some of these constructions, *het* can be absent (in non-initial position) or replaced by *er*. Examples of this from the Eindhoven corpus (a Dutch corpus with different text types, see Uit den Boogaart, 1975) are presented in (3.28) and (3.29).

(3.28) Maar steeds duidelijker blijkt *0* <u>dat militaire macht niet zonder meer kan worden omgezet in politieke macht</u>.
but increasingly clearer appears 0 that military power not just like that can be transformed into political power
"However, it becomes more and more clear that military power cannot be transformed into political power just like that."

(3.29) ... *er* bleek alras <u>dat zij allerminst gerekend moesten worden tot pseudo-voorstanders van de vliegengod of zelfs tot potentiële tegenstanders</u>.
... there turned out soon that they not in the least counted should be to pseudo-advocates of the fly god or even to potential opponents
"… it soon turned out that they should not at all be counted as pseudo-advocates of the fly god or even as potential opponents."

It should be noted that the role of predicates with a change of state, which turned out to be important in our study, is not discussed in the traditional literature on dummy subjects. Moreover, in most syntactic studies on dummy subjects hardly any attention, if any at all, has been paid to the question when *er* and *0* can occur in sentences with a sentential logical subject.

Passive sentences
Passive sentences are more or less the opposite of active sentences with respect to the distribution of dummy subjects: they generally show a preference for *er* or *0*. According to Haeseryn et al. (1997), *er* is used with non-factive predicates, as in (3.30) and *het* with factive predicates, as in (3.31). It should be noted that the dummy subject is in initial position in these sentences. Hence, the option of *0* is left out of consideration, because Dutch requires an overt element in sentence initial position in affirmative sentences.

(3.30) *Er* wordt beweerd <u>dat hij gefraudeerd heeft</u>. (p. 1138)
there is claimed that he committed fraud has
"It is claimed that he has committed fraud."

(3.31) *Het* wordt betreurd <u>dat hij gefraudeerd heeft</u>. (p. 1138)
it is regretted that he committed fraud has
"It is regretted that he has committed fraud."

Examples of passive sentences with different dummy subjects in non-initial position are presented in (3.32)-(3.36). As is observed by Haeseryn et al. (1997), the use of *er* is either optional or excluded when the subject is a subordinate clause in initial position (p. 475). It is not clear, however, when it is optional and when it is excluded.

42

(3.32) <u>Dat hij gefraudeerd heeft</u> wordt *er/0* beweerd. (Haeseryn et al., 1997: 475)[6]
that he committed fraud is there/0 claimed
"It is claimed that he has committed fraud."

(3.33) <u>Dat hij hard gewerkt heeft</u>, kan **er/0* niet gezegd worden. (Haeseryn et al., 1997: 475)
that he hard worked can there/0 not said be
"It cannot be said that he has worked hard."

(3.34) <u>Of er een dokter in de zaal was</u>, werd *er/0* gevraagd. (Haeseryn et al., 1997: 475)
if there a doctor in the room was was there/0 asked
"It was asked whether there was a doctor in the room."

(3.35) <u>Of ze misschien toch een verblijfsvergunning kan krijgen</u>, wordt **er/0* nog onderzocht. (Haeseryn et al., 1997, p. 476)
if she perhaps still a residence permit can get is there/0 still investigated
"It is still investigated whether she can get a residence permit after all."

(3.36) … dat *er/het* gezegd werd <u>dat jij komen zult</u>. (Vikner, 1995: 243)[7]
… that there/it said was that you come will
"… that it was said that you will come"

3.2   The present study: results from native speakers of Dutch

The native speakers of Dutch in our study serve as the control group for the results of the native speakers of German, French and Turkish who have acquired Dutch as their second language. Since some of the constructions in our study are ill-understood and subject to a great deal of variation, these results do not merely provide evidence for a pre-established native speaker norm. Rather, we also use these results to determine what preferences native speakers actually have with respect to dummy subjects in Dutch. This is why these results are presented here.

---

[6] When no example with *het* is given, the equivalent sentence with *het* is always ungrammatical.

[7] The dummy subject *0* is left out here, because Vikner does not give a judgement for this option. How Vikner arrived at the other judgements in this sentences is unknown. It seems to us that "het" is not acceptable for many native speakers of Dutch in this context.

Preliminary study and pilot study

In order to find out what factors play a role for native speakers in their choice of dummy subjects, we first did a very small-scale preliminary study in which we elicited absolute grammaticality judgements from five native speakers of Dutch, four of whom were linguists. On the basis of the results from this study we designed a pilot test, which we conducted with sixty native speakers of Dutch: thirty with a high level of education (higher professional education and university) and thirty with a somewhat lower level of education (MAVO and HAVO: lower and higher general secondary education). Each participant was presented with half of the test items. Because some participants in our preliminary study had indicated that they had problems giving absolute judgements, we used a sentence preference design (see chapter 4) for the pilot with a three-point scale (a = I prefer sentence A; a/b = I don't have a preference; b = I prefer sentence B). The order of the a- and b-sentence within items was randomised and the same for all participants. For our final test, we used those factors that seemed to play a role in the choice of dummy subjects in our pilot and those items for which the participants in our pilot study had the most consistent results. Since the results for active and passive sentences with a DP-subject were very similar, we restricted our final test to active sentences for this construction type.

In our final study, we used highly educated native speakers of Dutch (most of whom were undergraduate students). We did this for practical reasons and for comparability with the second language learners (most of whom had a very high level of education). Since the results from our pilot study did not reveal large differences between the highly educated group and a group with a lower level of education, this restriction seems justified for this construction type. Because of the differences between our pilot study and our final study, both in the items and in the scale that we used, we cannot be absolutely certain, however, that there are no differences related to educational level. In order to see whether our results from the native speakers are generalisable to the whole native speaker population, more research would be necessary.

General results

The results from the native speakers who participated in our sentence preference task (for the methodology: see chapter 4) reveal that for each construction type there is a general pattern, and that there are one or two factors which disturb this general pattern. Judgements for the (active) DP-type deviate from the general pattern when the subject is non-specific and the predicate is intransitive. Those for the active sentential type (AS-type) deviate from the general pattern when the predicate expresses a change of state. Judgements for the passive sentential type (PS-type), finally, deviate

from the general pattern when the equivalent active sentence contains a dummy object. The preference patterns with examples are given below:

– DP-type

General pattern (DPg):   **0 > er**
(3.37)   Men beseft niet altijd dat ◇ <u>een pinguïn</u> een vogel is.
         one realises not always that ◇ a penguin a bird is
         "One does not always realise that a penguin is a bird."
Transtive predicates with a non-specific subject (DPnst):   0 >= er
(3.38)   Op televisie doen ◇ <u>veel mensen</u> dingen die ze normaal niet durven
         on television do ◇ many people things that they normally not dare
         "On television many people do things they normally would not dare
         to do."
Intransitive predicates with a non-specific subject (DPnsi):   er >= 0
(3.39)   Ik vind het vervelend dat ◇ boven <u>een raam</u> open staat.
         I find it annoying that ◇ upstairs a window open stands
         "It bothers me that there is a window open upstairs."

– AS-type

General pattern (ASg):   **het > 0 >= er**
(3.40)   Meestal valt ◇ niet mee <u>om kaartjes voor een concert te krijgen</u>
         usually falls ◇ not with to tickets for a concert get
         "Usually it is not easy to get tickets for a concert."
Change of state (AScos):   **0 / het > er[8]**
(3.41)   Nu schiet ◇ mij ineens te binnen <u>dat ik nog boodschappen moet
         doen</u>.
         now occurs ◇ me suddenly that I still shopping must do
         "Now it suddenly occurs to me that I still have to go out shopping."

---

[8]   We use the symbol / in the presentation of the judgements from the native speakers in our experiment when there was no clear preference for one dummy subject. On average, however, the dummy subject to the left of this symbol was preferred.

– PS-type

General pattern (PSg):   **0 >= er >= het**

(3.42)   In de krant wordt ◇ beweerd <u>dat hij dronken achter het stuur</u>
<u>gezeten heeft</u>.
in the newspaper is ◇ claimed that he drunk behind the wheel sat
has
"It is claimed in the newspaper that he was drunk while he was
driving."

Sentences with a dummy object in the active equivalent (PSdo):

het >= 0 >= er

(3.43)   Door haar vrienden wordt ◇ bewonderd <u>dat ze ook in moeilijke</u>
<u>tijden vrolijk blijft</u>.
by her friends is ◇ admired that she also in difficult times cheerful
remains
"Her friends admire her for remaining cheerful, even in difficult
times."

Active sentences with a DP-subject
For active sentences with a DP-subject, the sentences that fit the general
pattern can be divided into three subcategories:
-          perception verbs with a non-finite embedded clause
-          categorical subjects
-          specific subjects

The sentences with categorical subjects get the strongest preference for *0*.
Transitivity turned out to be a weak factor. That is the reason why the
preference for *0* to *er* for transitive predicates with a non-specific subject
(DPnst) is weaker than that for the general category. Moreover, although in
general native speakers have a preference for *0* for DPnst and a weak
preference for *er* for intransitive predicates with a non-specific subject
(DPnsi; see table 3.1 below), there are some exceptions, as can be seen in
(3.44) and (3.45):

Transitive predicate, but weak preference for *er*
(3.44)   Op zijn verjaardag drinken ◇ nooit <u>zo veel mensen</u> bier.   er > 0
on his birthday drink ◇ never so many people beer
"At his birthday party there are usually not many people who drink
beer."

Intransitive predicate, but weak preference for *0*
(3.45)   Op dit feest zijn ◇ <u>heel veel mensen</u> behoorlijk dronken.   0 > er
          on this party are ◇ very many people rather drunk
          "Quite a few people are rather drunk at this party."

It may well be that transitivity and specificity interact. Moreover, (non-) specificity turned out to be a hard factor to control. The reason for this might be that the specificity of the subject is determined in part by the previous discourse, which was lacking in our test. The results for all categories are given in table 3.1. For more detailed information on individual items and deviations from the patterns presented below, see appendix C.

Table 3.1:   Results for the DP-type

Category	NS Pattern	Subcategories
DPg	0 > er	perception verbs with a non-finite embedded clause; categorical subjects; specific subjects
DPnst	0 >= er	-
DPnsi	er >= 0	-

Active sentences with a sentential subject
For active sentences with a sentential subject, the aspect of the predicate plays an important role. For predicates with stative aspect, as in (3.40) above, native speakers of Dutch have a strong overall preference for *het*. For predicates with terminative/inchoative aspect and a change of state verb, as in (3.41), this general pattern is disrupted: *0* gets better and *het* seems to get worse. It should be noted that there is a great deal of variation for this category and many different preference patterns occur, even inconsistent ones (e.g. *0 > het*; *het = er*; *er > 0*). However, only four native speakers have the same preference pattern for both predicate types. It should also be noted that other factors, such as factivity of the predicate, occupation of the object position by a DP and conditionality of the embedded clause, did not disturb the general pattern. An overview of the preference patterns of the native speakers is given in table 3.2. The subcategories presented are categories that seemed to play a role in our preliminary or pilot study, but turned out not to elicit distinctive judgements in our final test.

Table 3.2: Results for active sentences with a sentential subject

Category	DS pair	NS pattern	Subcategories
ASg	het - er	het > er	non-finite embedded clause; conditional embedded clause; DP object; other sentences
	het - 0	het > 0	non-finite embedded clause; conditional embedded clause; DP object; other sentences
	er – 0	0 >= er	non-finite embedded clause; conditional embedded clause; DP object; other sentences
AScos	het - er	het / er[*]	-
	het - 0	none (0 / het)	-
	er – 0	0 >= er	-

*   For some predicates, some NSs had a preference for *er* to *het* in initial position.

Passive sentences with a sentential subject

For the passives, we used various predicate types, but on the basis of our results we could only make a distinction between passives with a dummy object in the corresponding active sentence, as in (3.47), and passives that do not have this, as in (3.46).

(3.46) In de krant wordt ◇ beweerd <u>dat hij dronken achter het stuur gezeten heeft</u>.   0 >= er >= het
in the newspaper is ◇ claimed that he drunk behind the wheel sat has
"It is claimed in the newspaper that he was drunk while he was driving."

(3.47) Door haar vrienden wordt ◇ bewonderd <u>dat ze ook in moeilijke tijden vrolijk blijft</u>.   het >= 0 >=er
by her friends is ◇ admired that she also in difficult times cheerful remains
"Her friends admire her for remaining cheerful, even in difficult times."

Equivalent active sentence:

(3.48) Haar vrienden bewonderen *het* <u>dat ze ook in moeilijke tijden vrolijk blijft</u>.[9]
her friends admire it that she also in difficult times cheerful remains
"Her friends admire her for remaining cheerful, even in difficult times."

---

[9]   In this example, the dummy object is in italics and the semantic object is underlined.

Since dummy objects are always *het* in Dutch, native speakers also have a preference for *het* for the corresponding passive sentences. The preference pattern for this category is the same as the general pattern for active sentences, but it is weaker (compare table 3.3 to table 3.2), probably under the influence of the general preference pattern for passives. This general pattern is a clear preference for *er* and *0* to *het* and a very weak preference for *0* to *er*. It should be noted here that, for passives that do not have a dummy object in their equivalent active sentence, there is a great deal of variation, both between predicates and between native speakers. The patterns for the PS-type are presented in table 3.3. This table does not have a subcategories column, since there were no subcategories that could be distinguished in our pilot.

Table 3.3:   Results for the PS-type

Category	DS pair	NS pattern
PSg	het - er	er >= het
	het - 0	0 >= het
	er - 0	0 / er
PSdo	het - er	het > er
	het - 0	het >= 0
	er - 0	0 >= er

Relevant factors
On the basis of our results for the three dummy subject construction types we will focus in the remaining paragraphs on the following factors:
-        specificity and transitivity (for the DP-type)
-        change of state (for the AS-type)
-        dummy objects (for the PS-type)

3.3   Generative analyses

In generative syntactic research, dummy subjects have received considerable attention, but the constructions discussed are often limited to expletive *there* and some passive constructions, and to sentences with dummy subjects in initial position (see e.g. Hoekstra & Mulder, 1990; Chomsky, 1995; Lasnik, 1995; Moro, 1997; Radford, 1997). According to the Minimalist Program (Chomsky, 1995), features of subjects are case-checked in the AgrSP through spec-head agreement (against the features of the verb). In languages like Dutch, this checking can take place either by overt movement of the subject DP to spec AgrSP, or by attraction of case features from the subject

DP by spec AgrSP (covert movement). When case checking takes place through attraction, a dummy subject must fill the specifier position of the AgrSP (Radford, 1997)[10] to satisfy the Extended Projection Principle (EPP), according to which the subject position must be filled (see e.g. Chomsky, 1995:55). An example of such an analysis of item ns8 from the DPnsi category, presented in example (3.49), is given in figure 1.

(3.49)   … dat *er* vandaag nog <u>een bus</u> komt bij deze halte
         … that there today still a bus comes at this stop
         "… that there will be another bus at this bus stop today."

These "true" dummy subjects are assumed to lack a Θ-role and therefore "occur with verbs that fail to assign an external Θ-role such as existential *be*, unaccusative or raising verbs, and they always require the presence of a thematic associate (the logical subject)" (Felser & Rupp, 2001: 290). For the analysis of this construction (expletive *there*), different proposals have been made within generative grammar.

Chomsky (1995) analyses the expletive as an LF-affix and assumes that the formal features of the associate (the logical subject) adjoin to T. The Φ-features of the associate are assumed to consist only of a [person] feature (i.e. [number] features are lacking). According to Lasnik (1995), the associate must have partitive case, which is assigned by *be*, unaccusative verbs and raising verbs. This means that the associate must be indefinite, because definites have objective case. It also means that expletive *there* (or its equivalent in other languages) can never occur with a transitive verb. Several researchers, however, have challenged the link between partitive case and indefinites and have shown that expletives can sometimes combine with transitive verbs (see e.g. Felser & Rupp, 2001).

---

[10]   This proposal deviates from the original Minimalist theory as proposed by Chomsky (1995) in that, according to Chomsky, there is no AgrSP and case checking takes place within TP. The mechanism of checking case features, however, is the same in both proposals.
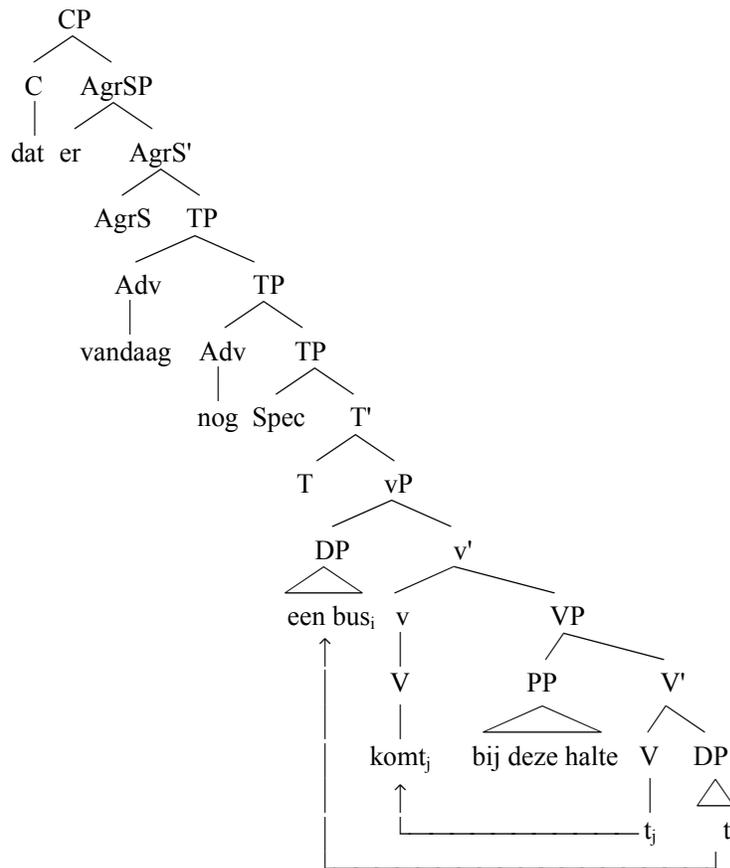
Figure 3.1:   *Minimalist analysis of dummy subject construction*

Felser and Rupp (2001) propose instead that the associate in existential sentences carries a [number], but no [person] feature. Because the [person] feature of T cannot be checked by the associate, the associate cannot be assigned nominative case and receives default case (case used for DPs that check neither structural nor inherent case). The expletive, which is considered a spatio-temporal argument and carries a [person] feature, is subsequently merged into spec TP and is assigned nominative case. In existential sentences (with or without expletive) the associate is interpreted in predicate-internal position, a position that is not available for individual level predicates, such as *to be intelligent* (see Carlson, 1977). This accounts

for why stage-level predicates (such as *to be available*) can be interpreted existentially and combine with expletive *there*, whereas individual level predicates cannot[11]. In order for stage-level predicates to receive a strong (specific) interpretation, the logical subject must appear outside the VP. Felser and Rupp assume that transitive sentences have two predicate-external subject positions. Because of this, transitive sentences can contain expletive *there* as well.

Moro (1997) analyses the expletive as the predicate of a small clause (SC) complement of a copula, which raises to spec-IP to check nominative case. The PP that is often (obligatorily) present in these sentences, as in (3.50), is analysed by Moro as IP-adjunct.

(3.50) $[_{IP} [_{IP} \textit{There}_i \text{ is}_j [_{VP} \text{t}_j [_{SC} [\underline{\text{a man}}] \text{ t}_i]]] [_{PP} \text{ in the room}]]$ (p. 98)

In these frameworks, it is not clear why overt and covert movement to AgrSP can coexist (compare (3.51) and (3.52)), what determines the choice of a particular dummy subject and why dummy subjects in some sentence types can be absent or phonetically unrealised, as in (3.53), in languages which have the Extended Projection Principle (or an EPP-feature).

---

[11]   According to H. de Hoop (personal communication, October 2004) individual level predicates can occur in Dutch with *er*, depending on the predicate and the context, e.g. in sentences like:
(i)   *Er* zijn <u>veel taalkundigen</u> een beetje gek
       there are many linguists a bit mad
       "Many linguists are a bit mad."
(ii)   *Er* zijn <u>weinig vrouwen</u> hoogleraar.
       there are few women professor
       "Few women are professors."
(iii)   *Er* kent in elk geval <u>één persoon</u> Latijn.
        there knows in any case one person Latin
        "There is at least one person who knows Latin."
(iv)   *Er* hebben <u>maar twee van de drie jonge katjes</u> een wit befje.
        there have two of the three young cats a white chest
        "Two of the three kittens have a white chest."
It should be noted that these sentences are somewhat marked, that their grammaticality can be influenced by the presence of adverbs and that *0* may in fact be preferred for several of these examples. However, the fact that *er* can be used in these examples is a problem for the theory outlined here.

(3.51) *Er* komt <u>een auto</u> de hoek om.[12]
        there comes a car the corner around
        "There comes a car around the corner."

(3.52) <u>Een auto</u> komt de hoek om.
        a car comes the corner around
        "A car comes around the corner."

(3.53) Zelden wordt *0* aan ouderen precies uitgelegd <u>hoe de kaartautomaat op het station werkt</u>.
        rarely is 0 to elderly exactly explained how the ticket machine at the station works
        "It is rarely explained to elderly people exactly how the ticket machine at the station works."

In the Minimalist Program, overt movement must always be motivated by checking requirements of strong features, and "arbitrary" scrambling is not possible. What motivates overt movement in sentences like (3.51), but not in what have been termed equivalent sentences, such as (3.52), has not been accounted for within these frameworks. In the case of sentences with a dummy subject and a logical subject in the form of a DP (DP-type), one could imagine that a presentational or existential feature is checked, but such features have not been proposed in the literature.

Another problem is that the subject position does not have to be phonetically filled in all contexts. If the subject position can be phonetically empty in sentences like (3.53), why can it also be filled by *er* in the same sentence and why can there not be a phonetically empty subject in initial position and in sentences like (3.54) to (3.56)?

(3.54) Langzaam drong *het/*0* tot hem door <u>dat hij nooit meer zou kunnen lopen</u>.
        slowly got it/0 to him through that he never anymore would can walk
        "It was slowly getting through to him that he would never be able to walk again."

(3.55) Fatima zegt dat *het/*0* zelden voorkomt <u>dat zij te laat zijn</u>.
        Fatima says that it/0 rarely happens that they too late are
        "Fatima says that it rarely happens that they are late."

(3.56) Zij vond dat *het/*0* een schande zou zijn <u>als dat plan door zou gaan</u>.
        she found that it/0 a disgrace would be if that plan through would go
        "She thought it would be a disgrace if that plan went through."

---

[12] This example is not taken from our sentence preference task, because all pairs were constructed in such a way that they only differed with respect to the choice of dummy subject.

Bennis (1987), in a pre-minimalist framework, proposes that dummy subjects are not semantically empty. Instead, he proposes that in dummy subject constructions[13] either an adverbial pronoun (*er*) is adjoined to the left of the subject (like PPs), or a referential pronoun (*het*) occupies the subject position and the external argument (the logical subject) is extraposed. According to Bennis, restrictions with respect to what logical subjects can combine with a dummy subject are semantically and/or pragmatically determined. He proposes two pragmatic rules that govern the use of expletive *er*. They are given in a and b.

a)   In unmarked sentences, presupposition precedes focus (p. 223)
b)   There is a presuppositional hierarchy (from [+presuppositional] to [-presuppositional]):
         weak pronoun - strong pronoun - name - definite NP - indefinite NP [+spec] - indefinite NP [-spec] (p. 223)

Although this proposal accounts for the fact that *er* can be optional in non-initial position (because it is an adjunct), it does not account for why the same seems to hold for *het* (which cannot be an adjunct in Bennis' proposal) with certain verbs (see e.g. (3.41 and 3.43) above).

Only some of the factors that in our experiment with native speakers turned out to play an important role in the choice of dummy subjects in Dutch (specificity, transitivity, change of state and dummy objects), have been elaborately discussed in the generative literature. Instead of specificity, definiteness is considered to play a crucial role. Unaccusativity (which could not be distinguished from change of state as the crucial factor for the AS-type in our study) is discussed in some of the literature, but with respect to the DP-type, rather than the AS-type. Moreover, unaccusative verbs are claimed to behave in the same way as raising verbs in these accounts (see e.g. Felser & Rupp, 2001), which turned out not to be the case for sentences of the AS-type in our study.

        What all these generative accounts have in common is that they cannot account for (part) of the optionality of *er* and/or *het* in our native speaker data. If Dutch has an EPP-feature, it should always work and an overt subject should always be present. Our data clearly show that both *er* and *het* can sometimes be optional, but not for all predicates. Accounting for this optionality will be a challenge for future generative work.

---

[13]   The term "dummy subject constructions" is used here as proposed in this paragraph. Since Bennis (1987) argues that these elements are not dummy subjects, he would not call these constructions "dummy subject constructions".

3.4   German, French and Turkish

So far, we have discussed some general generative analyses of dummy subject constructions, some traditional analyses of dummy subjects in Dutch and the results from the native speakers of Dutch in our experiment. To gain more insight into the possible role of transfer in the results from our second language learners, we will now discuss relevant constructions and factors in German, French and Turkish that may affect the acquisition of dummy subject constructions by German, French and Turkish speaking learners of Dutch. We will also discuss what German, French and Turkish learners of Dutch have to acquire in order to behave like native speakers of Dutch on items with dummy subject constructions in our tests.

Dummy subjects and relevant factors in German
German sentences in which the logical subject is not in clause initial position for semantic or pragmatic reasons normally have *es* in initial position, as in (3.57).

(3.57)   *Es* wartet <u>jemand</u> auf dich (Drosdowski & Augst, 1984: 720)
         there waits someone on you
         "There is someone waiting for you."

Some German syntacticians claim that *es* can only occur in sentence-initial position (see e.g. Drosdowski & Augst, 1984: 720 or Engel, 1988: 860). According to Müller's generative analysis (Müller, 1998), the reason for this is that German has only one functional projection (F(inite)P) above VP, which is always occupied in embedded or inverted sentences (either by a complementizer or by an object or adjunct), so it can never contain *es*. Others, however, give examples of German sentences with *es* in non-initial position, as in (3.58)-(3.60). As is the case for dummy subjects in Dutch, *es* can occur both in sentences with inversion, as in (3.58) and (3.59) and in embedded sentences, as in (3.60).

(3.58)   Heute macht *es* mir großes Vergnügen, <u>Sie hier zu sehen</u>. (Curme, 1952: 458)
         today makes it me great joy you here to see
         "I am very pleased today to see you here."
(3.59)   Uns freute *es*, <u>den Freund wiederzusehen</u> (Helbig & Buscha, 1991: 110)
         us pleased it the friend again to see
         "It pleased us to see our friend again."

(3.60)   Es scheint, daß *es* ihr unsagbar mühsam ist, <u>den Mund zu öffnen</u>
         (Zifonun et al., 1997: 1284)
         it seems that it her unspeakably difficult is the mouth to open
         "It seems that it is unbelievably hard for her to open her mouth."

When *es* can occur in non-initial position, it can also sometimes be left out. According to Curme (1952), this happens most often after a non-emphatic predicate, as in (3.61).

(3.61)   Richtig ist *0*, <u>daß …</u> (Curme, 1952: 460)
         right is, that …
         "It is right that …"

Other examples in the literature of sentences in which *0* can presumably occur are given below.

(3.62)   Einmal war *0* <u>ein König</u> in großer Not. (Curme, 1952: 462)
         once were a king in great need
         "Once there was a king in great need."
(3.63)   Mir war *es/0* sehr peinlich, <u>ihn um das Buch zu bitten</u>. (Helbig &
         Buscha, 1991: 110)
         to me was (it) very painful him for the book to ask
         "It was very painful for me to ask him for the book."
(3.64)   Natürlich ist *es/0* gut, <u>daß du gekommen bist</u> (Vikner, 1995: 226)
         of course is (it) good, that you come are
         "Of course it is good that you have come."
(3.65)   Gesagt wird *es/0* nicht <u>daß Johan krank ist</u>. (Vikner, 1995: 230-231)
         said is it/0 not that Johan ill is
         "It is not said (emphasis) that Johan is ill."

It should be noted that in the examples in which *es* can occur in non-initial position, the logical subject always has the form of an embedded clause. According to Vikner (1995), in these cases *es* or *0* (which he calls *pro*) is a (quasi-)argument and can occur both in spec IP or in spec CP. If it occurs in spec IP, *es* is optional according to Vikner, as in (3.64) and (3.65), except with raising verbs, as in (3.66).

(3.66)   … daß *es/*0* scheint <u>daß du kommen würdest</u>. (Vikner, 1995: 264)
         … that it/*0 seems that you come will
         "… that it seems that you will come."

Expletive *es* (which has a DP-subject), on the other hand, can only occur in

spec CP, because the verb is in C in verb second (V2) languages. When this position is occupied by a complementizer, a topic or a wh-element, as in (3.67)-(3.69), the expletive (which is generated in spec IP) must be realized (in spec IP) as *pro* in German.

(3.67) Ich weiß, daß **es/0* <u>ein Junge</u> gekommen ist. (Vikner, 1995:70)
        I know, that it/*pro* a boy come is
        "I know that a boy has come."

(3.68) Gestern ist **es/0* <u>ein Junge</u> gekommen. (Vikner, 1995:185)
        yesterday is it/*pro* a boy come
        "Yesterday a boy came."

(3.69) Warum ist **es/0* <u>ein Junge</u> gekommen? (Vikner, 1995:185)
        why is it/*pro* a boy come?
        "Why did a boy come?"

Vikner (1995) does not indicate, however, why this should be the case. Moreover, this hypothesis cannot be carried over to Dutch. Although Dutch behaves like German with respect to V2, expletive *er* (DP-type) can occur in non-initial position in Dutch. It should be noted that in this analysis (contrary to Müller, 1998), there are at least two functional projections above VP, so there is no obvious reason why *es* could not be in spec IP position (like *pro*).

     If we consider in what sentence types *es* can occur in German, there seem to be very few restrictions. According to Vikner, there is no restriction in German (and Dutch) on the predicate types with which *es* can occur: it can occur with unaccusatives (which Vikner calls ergatives), active and passive intransitives, and both active and passive transitives. Some examples are given in (3.70)-(3.72).

(3.70) *Es* ist <u>ein Junge</u> gekommen.   (Vikner, 1995:69: unaccusative predicate)
        there is a boy come
        "A boy has come."

(3.71) *Es* hat <u>jemand</u> Bücher gekauft (Vikner, 1995: 235: active transitive)
        there has someone books bought
        "Someone has bought books."

(3.72) *Es* wurde am Tatort <u>ein Dänischer Linguist</u> gesehen (Vikner, 1995: 175: passive transitive)
        there was at the scene of the crime a Danish linguist seen
        "A Danish linguist was seen at the scene of the crime."

With respect to specificity for constructions with a DP-subject, German, unlike languages like Dutch and English, can have *es* with definite/specific

subjects, as can be seen in (3.73) and even with categorical subjects as in (3.74)

(3.73)   *Es* ist heute <u>der Brief von Maria</u> gekommen (Vikner, 1995: 174)
         there is today the letter from Maria come
         "Today Maria's letter arrived."
(3.74)   *Es* hat <u>der Basilgenpeterlein</u> eine Krafft ...
         (http://www.kraeuter.ch/trachyspermum/Trachyspermum.htm)
         there has the bishop's weed a warm strength …
         "The bishop's weed has a warm strength..."

The only clear restriction on *es* in German seems to be that it does not occur in sentences with individual level predicates, as in (3.75)

(3.75)   *Es war <u>jemand</u> intelligent. (Felser & Rupp, 2001: 302)
         there was someone intelligent
         "Someone was intelligent."

Looking at the differences between dummy subjects in German and Dutch, it becomes clear that German learners of Dutch have to learn several things in order to behave like a native speaker of Dutch with respect to dummy subject constructions. They have to find out that:
-   *es* should be translated as *er* in sentences with DP-subjects and in passive sentences with no dummy object in the active equivalent, but mostly as *het* in active sentences with a sentential subject and in passive sentences with a dummy object in their active equivalents.

-   *er* can occur in non-initial position in sentences with a DP-subject, but not with specific and categorical subjects.
-   The choice between *het* and *0* is not optional in active sentences with a sentential subject, but related to properties of the verb (change of state verbs versus stative verbs)

If we consider the factors that determine which dummy subject should be used in Dutch, we can try to predict which of the aspects mentioned above will be the most difficult to acquire for German learners of Dutch. Specificity, which is the most important factor in dummy subject sentences with a DP-subject in Dutch, is not expressed through grammatical means in German. However, German does distinguish between definites and indefinites by using different articles. Since specificity and definiteness often coincide, in many cases German learners of Dutch might use definiteness to decide whether to use *er* or not. However, categorical subjects are indefinite

but behave like specific subjects in that they do not take *er*. We might expect, therefore, that for dummy subject sentences with a DP-subject German learners of Dutch, if they have discovered that *er* can occur in non-initial position for this construction in Dutch, will have most problems with categorical subjects in our test.

For the AS-type, it is difficult to tell whether German learners of Dutch will have problems acquiring the distinction between change of state verbs and other predicates. Since this distinction seems never to have been looked at before, we cannot tell whether German and Dutch behave similarly or differently in this respect.

For passive sentences with a sentential subject, the relevant factor in Dutch is the presence of a dummy object in equivalent active sentences. Since German also has dummy objects and generally with the same predicates as Dutch, this factor in itself should not cause great problems for German learners of Dutch. However, it might be difficult to connect the right dummy subject preference to the right construction. German learners of Dutch will have to find out that the general pattern for passives is a preference for *er* and *0* to *het*, whereas for passives with a dummy object in the active equivalent, *het* is preferred to *0* and *er*.

All in all, we expect German learners of Dutch to have the fewest problems with active sentences with a sentential subject. For the DP-type, we expect less advanced learners to overgeneralise *0* to the whole construction, and more advanced learners to overgeneralise *er* to categorical and perhaps specific subjects. For the passives, we predict problems in finding out the right preference pattern for the right category. We do, however, expect German learners to be consistent within the two categories of the PS-type (PSg and PSdo).

Dummy subjects and relevant factors in French
In French there are two dummy subjects, *il* and *ce*. Dummy subjects in French can hardly ever be left out. Examples of different constructions with dummy subjects are presented in (3.76)-(3.89) below. The judgements presented in these examples are based in part on the literature and in part on judgements from three native speakers of French.

(3.76)   Hier *il/*c'/*0* est venu <u>quelques hommes</u>. (unaccusative + DP-subject; Hoekstra & Mulder, 1990: 47)
yesterday there is came some men
"Yesterday some men came."

(3.77) Elle disait qu'*il*/?que *c*'est/*que *0* est établi que le ministre a fraudé. (unaccusative + sentential subject) [14]
she said that it is established that the minister has frauded
"She said that it has been established that the minister has committed fraud."

(3.78) En 1930 *il* roulait encore quelques trams dans Paris. (unergative + DP-subject; Hoekstra & Mulder, 1990: 48)[15]
in 1930 there ran still some trams in Paris
"In 1930 there were still some trams running in Paris."

(3.79) Maintenant *il/*ce/*0* parait que le ministre a fraudé. (raising verb + sentential subject)
now it seems that the minister has frauded
"It seems now as if the minister has committed fraud."

(3.80) A mon avis *il/c'/*0* est possible qu'il vienne. (*être* + adjective + sentential subject)
according to me it is possible that he comes
"I believe it is possible that he will come."

(3.81) **Il* a dansé quelqu'un dans le jardin. (active intransitive + DP-subject; Vikner, 1995: 203)
there has danced someone in the garden
"Someone has danced in the garden."

(3.82) **Il* a mangé quelqu'un une pomme. (active transitive + DP-subject; Vikner, 1995: 198)
there has eaten someone an apple
"Someone has eaten an apple."

(3.83) ?**Il* a été dansé. (passive intransitive without logical subject; Vikner, 1995: 209)
there has been danced
"There were people dancing."

(3.84) Elle disait que ici *il/*c'/0* a été organisé un grand concours. (passive transitive + DP-subject)
she said that here there has been organised a big competition
"She said that a big competition has been organised here."

---

[14] When no reference is given, the grammaticality judgements in these examples are my own. The reason for this is that in the literature examples are often incomplete, in the sense that judgements are not given for all dummy subject types or only for one position. For the sake of consistency, quoted examples are sometimes presented in a slightly different form than in the original source.

[15] We used the original example from Hoekstra and Mulder (1990) here. It should be noted, though, that *en Paris* would be more correct in this context.

(3.85)   *Il* a été mangé <u>une pomme</u> (passive transitive + DP-subject; Vikner, 1995: 202)
there has been eaten an apple
"An apple has been eaten."
(3.86)   *Il me vexe, <u>qu'il ait fait ça</u>. (active + sentential subject).
it me annoys that he has done that
"It annoys me that he did that."
(3.87)   *Il me plait, <u>qu'il ait fait ça</u>. (active + sentential subject)
it me pleases that he has done that
"I am glad that he has done that."
(3.88)   *Il me fait plaisir <u>qu'il ait fait ça</u>. (active + sentential subject)
it me does pleasure that he has done that
"I am glad that he has done that."
(3.89)   *Il* lui passa par la tête <u>qu'il avait oublié de faire des courses</u>. (active + sentential subject)
it him passed through the head that he had forgotten to do shopping
"It struck him that he had forgotten to do the shopping."

According to Vikner (1995), French dummy subjects can only occur with unaccusatives, with raising verbs, with *être* + adjective and in passive transitive sentences. However, as can be seen in (3.89), there are also other verbs that can combine with *il*. In the *être* + adjective construction there seems to be a general preference for *ce* and for unaccusatives there seems to be a general preference for *il*. *Il* is generally considered more formal or poetic than *ce*.

For the unaccusative sentences with a DP-subject, the same restrictions on the subject seem to hold as in Dutch: the subject cannot be specific or categorical, as is shown in examples (3.90)-(3.92).

(3.90)   *Il est venu <u>les hommes</u>. (specific subject)
there is come the men
"The men have come."
(3.91)   *On ne réalise souvent pas, qu'*il* est <u>un pingouin</u> un oiseau.
(categorical subject)
one realises often not that there is a penguin a bird
"One does not always realise that a penguin is a bird."
(3.92)   *On ne réalise souvent pas, que *c*'est <u>un pingouin</u> un oiseau.
(categorical subject)
one realises often not that there is a penguin a bird
"One does not always realise that a penguin is a bird."

With respect to sentences with a DP-subject, French learners of Dutch have to learn that in Dutch dummy subjects are not limited to unaccusative and passive transitive predicates. Once they have learned this, however, they can apply the same restrictions on the subject as in French.

For active sentences with a sentential subject, dummy subjects can roughly occur with the same predicate types as in Dutch, but in most cases alternative sentences without a dummy subject are preferred in French. The different predicate types, however, follow a different pattern for Dutch. Whereas in French there is a distinction between *être* + adjective (preference for *ce*) versus all other predicate types, in Dutch the distinction is between (a subclass of) unaccusative verbs (namely change of state verbs) versus all other predicate types. We might therefore expect that, in the case of active sentences with a sentential subject, French learners of Dutch will have most problems with predicates that are neither unaccusative nor contain an adjective (e.g. reflexive verbs).

For passive sentences with a sentential subject, French learners of Dutch have to learn to use dummy subjects, since the equivalents in French in general do not contain a dummy subject. Furthermore, since French does not have dummy objects, they also have to learn the distinction between sentences that have a dummy object in their active equivalent versus sentences that do not. We therefore expect this sentence type to be very difficult for French learners of Dutch. It seems likely that French learners will have a preference for *0* for all passives. For the passives without a dummy object in their active equivalent, this would actually be a reasonably good strategy, since native speakers also have a (slight) preference for *0* for this category. For passives with a dummy object in their active equivalents, however, this strategy will not work. We therefore expect French learners of Dutch to have most problems in this category.

Relevant factors in Turkish
Turkish does not have any semantically empty elements, like dummy subjects, and pronominal subjects can often be left out (from a generative perspective, Turkish is a pro-drop language and does not have an EPP feature). In Turkish, the logical subject always coincides with the syntactic subject. Constructions taking the form of an embedded clause in Western European languages can either have the form of an embedded clause in Turkish (as in (3.95)) or they can be nominalised (as in (3.96)). The former construction is borrowed from Persian and some native speakers of Turkish do not use it. Examples of Turkish sentences with a dummy subject in their Dutch equivalents are given in (3.93)-(3.99)[16].

---

[16]   I would like to thank Hugo Strötbaum for his help with and comments on the Turkish examples in this paragraph.

(3.93) <u>Bir adam</u> geldi.
a man come-PAST(3s)
"There came a man."

(3.94) Dün <u>bir adam</u> geldi.
yesterday a man come-PAST(3s)
"Yesterday a man came."

(3.95) Biliyorum ki <u>bir adam</u> geldi.
know-1sg, that a man come-PAST(3s)
"I know that a man came."

(3.96) Bir adamın geldiǧini bil-iyor-um.
a man-GEN come-NOM-POSS(3s)-ACC know-PROGR-1s
"I know that a man came."

(3.97) Beni kızdır-dı ki gel-di-n.
I-DAT irritate-PAST(3s) that come-PAST-2s
"It annoys/annoyed me that you came."

(3.98) (Senin) geldiǧin beni kızdır-dı.
(you-gen) come-NOM-POSS(2s) I-DAT irritate-PAST(3s)
"It annoys/annoyed me that you came."

(3.99) Dans edildi.
dance do-PASS-PAST(3s)
"People danced."

Since Turkish does not have dummy subjects, Turkish learners of Dutch have to learn many things in order to be able to perform like native speakers of Dutch with respect to dummy subject constructions. Besides finding out which dummy subject to use in which construction in Dutch (see paragraph 3.2), they have to find out:
- that certain complex constituents that are nominal in Turkish have the form of an embedded clause in Dutch (for those native speakers of Turkish who do not have the "ki-construction")
- the form of embedded clauses in Dutch
- that Dutch has dummy subjects
- that Dutch has dummy objects
- that specificity plays a role in the choice of dummy subjects (of the DP-type)
- that change of state plays a role in the choice of dummy subjects (for the AS-type)

Let us now consider the factors that play a role in the choice of dummy subjects in Dutch. Specificity, which is the most important factor in dummy subject sentences with a DP-subject in Dutch, is expressed through grammatical means in Turkish, but it is the object which is marked, rather

63

than the subject. Specific objects get an accusative case morpheme in Turkish, whereas non-specific objects are not marked morphologically, as can be seen in (3.100)-(3.104).

(3.100) Mektub-u yaz-dı-m.
  letter-ACC write-PAST-1sg
  "I wrote the letter."

(3.101) Mektup yaz-dı-m.
  (a) letter(s) write-PAST-1sg
  "I wrote a letter / I wrote letters."

(3.102) Ali bir piyanoyu kiralamak istiyor. (Enç, 1991: 4)
  Ali a piano-ACC hire want-PROGR
  "Ali wants to hire a (specific) piano."

(3.103) Ali bir piyano kiralamak istiyor. (Enç, 1991: 5)
  Ali a piano hire want-PROGR
  "Ali wants to hire a (any) piano."

(3.104) Oda-m-a birkaç çocuk gir-di. (Enç, 1991: 6)
  room-my-DAT a few child enter-PAST
  Some children entered my room.

a) Iki kız tanı-yor-du-m.
  two girl know-PROGR-PAST-1sg
  "I knew two girls."

b) Iki kız-ı tanı-yor-du-m.
  two girl-ACC know-PROGR-PAST-1sg
  "I knew two of the girls."

The fact that specificity is expressed through grammatical means in Turkish might help Turkish learners of Dutch to choose the right dummy subject for sentences with a DP-subject.

    For the active sentences with a sentential subject, we expect that Turkish learners of Dutch may have problems acquiring the right preference pattern for the two categories (change of state and the general category). Change of state verbs do sometimes behave differently from other predicates in Turkish in that they can get a past tense marking for a present tense meaning (Özyürek, personal communication, 2003), as can be seen in (3.105)-(3.106). This possibility does not exist for other verbs, such as *yemek* in example (3.107).

(3.105) Gel-di-m.
  come-PAST-1s
  "I am coming/came/have come."

(3.106) Ibrahim kız-dı.
    Ibrahim get angry-PAST(3s)
    "Ibrahim is/was/got angry."
(3.107) Ibrahim domuz eti ye-me-di.
    Ibrahim pork (meat) eat-NEG-PAST(3s)
    "Ibrahim *does/did not eat / has not eaten pork."

We expect it will be difficult, however, for the Turkish learners to use this distinction in finding out the right pattern in Dutch, especially given the huge amount of variation for the native speakers of Dutch on AScos.

For passive sentences with a sentential subject, the relevant factor in Dutch is the presence of a dummy object in equivalent active sentences. Since Turkish does not have anything like dummy objects at all, we expect this construction type to cause the greatest problems for Turkish learners of Dutch.

All in all, we expect Turkish learners of Dutch to have the fewest problems with active sentences with a DP-subject, and the most problems with passive sentences with a sentential subject. Since there is nothing that Turkish learners of Dutch can transfer directly from their L1, we expect them to have great difficulties in finding out which dummy subject to use in which construction in Dutch.

Problems for German, French and Turkish learners of Dutch
In paragraph 3.2, we discussed which factors are relevant for the choice of dummy subject in our tests. In the previous subparagraphs, we discussed what German, French and Turkish learners of Dutch have to learn with respect to dummy subject constructions and where we expect most problems to occur. Table 3.4 provides an overview of the relevant facts that each learner group has to acquire for each sentence type. The things we expect to be most difficult to acquire are printed in bold.

Table 3.4: Predicted problems for learners of Dutch with respect to dummy subject constructions

	DP-type	AS-type	PS-type
L1G	- er in non-initial position - distinction between specific versus non-specific subjects + right pattern	- distinction between change of state verbs and other predicates + right pattern[*]	- dummy objects as crucial factor + right pattern
L1F	- dummy subjects for all predicate types	- ce → het - il → 0 for change of state verbs - il → het for all other predicates	- use of dummy subjects - distinction between sentences that have dummy objects and those that do not + right pattern
L1T	- specificity of subject as crucial factor + right pattern	- embedded clause at end of sentence - change of state verbs as crucial factor + right pattern	- embedded clause at end of sentence - distinction between sentences that have dummy objects and those that do not + right pattern

*   It is not clear whether this distinction is also made in German or not.

We expect these learning difficulties to translate into the following accuracy order, where the constructions that we expect to be easiest are on the left and the constructions for which we expect most problems are on the right:

L1 German: ASg > PS > DP
L1 French: DPg/DPnst > DPnsi > AScos > ASg > PSg > PSdo
L1 Turkish: DPg/DPnst > DPnsi > AScos > ASg > PSg > PSdo

It should be noted that for the L1 German group AScos is left out, because we do not know how this works in German. For the L1 French and L1 Turkish group we put AScos before ASg and DPg/DPnst before DPnsi, because for AScos, DPg and DPnst *0* is generally preferred. It should be noted that, in spite of the many differences between French and Turkish, our predictions with respect to accuracy order are the same for speakers from both first language backgrounds. This does not mean, however, that we expect the Turkish and French participants to behave similarly on the sentence preference task. We expect the native speakers of Turkish to have more problems in general than the native speakers of French. However, we do expect both groups to have most problems with categories that are on the right in the accuracy hierarchy and least problems with categories that are on the left in this hierarchy. In chapter 5, we will see whether these predictions are borne out in our study.

Chapter

4

Methodology

In this chapter, the methodology of our study will be presented. We will first give some information about the participants (see § 4.1). Subsequently, in §4.2, we will very briefly discuss the constructions used and give some examples. For a discussion of the literature on dummy subjects or more information about dummy subjects in Dutch, we refer the reader to chapter 3. The tasks we used in our study are described in § 4.3. Finally, in §4.4 we will discuss how we analysed our data.

4.1   Participants

There were four groups of participants in our study:

1)   A control group of 44 native speakers of Dutch
2)   15 very advanced German late learners of Dutch
3)   15 very advanced French late learners of Dutch
4)   13 very advanced Turkish late learners of Dutch

The control group of native speakers of Dutch (17 men and 27 women) consisted of 38 undergraduate students at the University of Nijmegen, two PhD-students and four participants with a somewhat lower level of education. This group originally consisted of 47 participants, but three of them were excluded from the analysis: one who turned out to have lived in Norway for several years before puberty and two who performed very deviantly from the other native speakers in previous analyses (see van Boxtel, Bongaerts & Coppen, 2003). One of the latter had a lower level of education and indicated that she had problems with the task and using a PC. We excluded these native speakers with very deviant behaviour from the other native speakers in order to have a very strict and clear criterion for native-likeness.

We recruited mainly highly educated native speakers for our control group both because we had easy access to this group and for comparability with the second language learners (most of whom had a high level of education). Since the results from a pilot study did not reveal large differences between a highly educated group and a group with a lower level of education, this restriction to highly educated native speakers seems justified for the construction we investigated.[17]

The advanced late learners of Dutch were selected through a network procedure. We selected participants who had arrived in the Netherlands when they were at least twelve years old and who were very proficient in Dutch. For each non-native speaker, at least one native speaker of Dutch (usually an expert in the field of second language teaching or second language acquisition) or another L2 participant indicated that he or she spoke Dutch so well that he/she could pass for a native speaker most of the time in all aspects of language, except perhaps for pronunciation or gender errors. We used a background questionnaire to check the age on arrival criterion, and conversations with the experimenter before or after the test to check the proficiency criterion. If participants turned out not to meet one of these criteria during this stage, they were removed from the analyses.

In the L1 German group, one participant was replaced because it turned out he had lived in Flanders from the age of eleven until the age of twenty-three. One participant from the L1 French group and one from the L1 Turkish group were excluded because they did not meet the proficiency criterion. The French participant was replaced. Another participant in the L1 Turkish group was excluded because he had lived in the Netherlands before the age of twelve. There was also a Turkish participant who had lived in Germany before the age of twelve. However, since Turkish was the language

---

[17] Because of the differences between our pilot study and our final study, both in items and in the scale that we used, we cannot be absolutely certain that there are no differences with respect to choice of dummy subject in Dutch related to educational level for native speakers of Dutch. More (sociolinguistic) research is needed to explore this relation. If there were differences between native speakers of Dutch related to educational level, that would mean that second language learners, who are normally exposed to input from native speakers with different educational backgrounds, do not only receive the kind of input represented in the results of our control group, but also other input. If some second language learners were to base their grammar of dummy subjects on the (different) input from native speakers with a lower level of education, our results would then underrepresent the number of second language learners that behave like native speakers of Dutch.

of his environment and since he had had very little contact with German at that time, he was not excluded from the analyses.[18]

The L2 speakers (20 men and 23 women) had arrived in the Netherlands between the ages of twelve and 35 and had lived in the Netherlands for four to 50 years. Relevant background information was obtained through a questionnaire (see appendix A). Thirty-eight participants had finished a higher professional or university education and 21 participants either had a job or had done a study in which language plays an important role (e.g. translating or work in a language department at a university). Some relevant background characteristics are presented in table 4.1. In the last row of this table, the background characteristics of all three non-native speaker groups are taken together. Additional background characteristics can be found in appendix E.

Table 4.1:   Background characteristics of all participants

	Age	Sex m/f	AoA	LoR	Prof.[*] Dutch	L1 prof.[*]
Native speakers (44)	22 (18-46)	17/27	-	-	-	-
L1 German (15)	51 (28-73)	7/8	26 (14-35)	25 (6-50)	4.1 (3-5)	4.6 (3-5)
L1 French (15)	42 (26-59)	3/12	21 (12-26)	19 (4-44)	4.1 (3-5)	4.6 (3-5)
L1 Turkish (13)	38 (23-50)	10/3	17 (12-28)	21 (7-31)	4.0 (3-5)	4.5 (4-5)
Non-native speakers (43)	44 (23-73)	20/23	22 (12-35)	22 (4-50)	4.1 (3-5)	4.6 (3-5)

AoA: Age of arrival in the Netherlands
LoR: Length of residence in the Netherlands
*   Rating for self-estimated overall proficiency on a five-point scale ranging from very poor (1) to very good (5)
         N.B. One German participant did not answer this question for Dutch

4.2   Dummy subjects

For our tests we chose dummy subject constructions, as these are known to be very hard to acquire for learners of Dutch as a second language. These constructions are not only very complex, they are also poorly understood and therefore hardly covered in Dutch grammars (see chapter 3). This means that

---

[18]   This participant scored within the native speaker range on DPg and DPnst and had deviant scores for all other categories (see appendix D). Leaving him out would, therefore, not have had consequences for our conclusions (see chapter 5).

even learners who have had extensive formal training in Dutch do not have access to explicitly formulated rules, and therefore can only acquire these structures on the basis of evidence from the input. An overview of the factors that play a role in the choice of dummy subjects in Dutch and the results from the data from our control group are given in chapter 3. Here we will repeat our definition and some examples of different dummy subject constructions.

As we explained in chapter 3, we use the term "dummy subject constructions" for constructions that contain a semantically empty element that fulfils the syntactic function of a subject. The logical subject (a nominal constituent or a clause) must occur in the same sentence. In Dutch, the role of dummy subject is fulfilled either by *het* (it), *er* (there) or *0* (i.e. the absence of *het* or *er*). We distinguished three types of dummy subject constructions in Dutch:

(I)   sentences with *er* or *0* in which the logical subject is a noun phrase (DP-type)

(II)   active sentences with *er*, *het* or *0* and a sentential logical subject (AS-type)

(III)   passive sentences with *er*, *het* or *0* and a sentential logical subject (PS-type)

In this study, we will restrict ourselves for the DP-type to active sentences, because there seems to be no difference between active and passive sentences of this type (according to the literature and results from a pilot study). We distinguished categories for each type on the basis of the preference patterns of the native speakers. For each type there is a general pattern (g) and one or two factors for which this general pattern is disturbed. We consider all sentences of the same type that have the general preference pattern to belong to the same category, which we call the general category. An example of each preference pattern for the three types is presented in (4.1)-(4.7) below, repeated from (3.37)-(3.43) on pages 45-46. In these examples, the dummy subjects are replaced by ◇ and the logical subject is underlined. To the right of each example the preference pattern of the native speakers in our study is given.

DP-type, general (DPg)

(4.1)   Men beseft niet altijd dat ◇ <u>een pinguïn</u> een vogel is.   0 > er
          one realises not always that ◇ a penguin a bird is
          "One does not always realise that a penguin is a bird."

DP-type, non-specific subject + transitive predicate (DPnst)

(4.2)   Op televisie doen ◇ <u>veel mensen</u> dingen die ze normaal niet
        durven.   0 >= er

        on television do ◇ many people things that they normally not dare
        "On television many people do things they normally would not dare
        to do."

DP-type, non-specific subject + intransitive predicate (DPnsi)

(4.3)   Ik vind het vervelend dat ◇ boven <u>een raam</u> open staat.   er >= 0

        I find it annoying that ◇ upstairs a window open stands
        "It bothers me that there is a window open upstairs."

AS-type, general (ASg)

(4.4)   Meestal valt ◇ niet mee <u>om kaartjes voor een concert te krijgen</u>.

                                                        het > 0 >= er

        usually falls ◇ not with to tickets for a concert get
        "Usually it is not easy to get tickets for a concert."

AS-type, change of state (AScos)

(4.5)   Nu schiet ◇ mij ineens te binnen <u>dat ik nog boodschappen moet</u>
        <u>doen</u>.   0 / het >= er

        now occurs ◇ me suddenly that I still shopping must do
        "Now it suddenly occurs to me that I still have to go out shopping."

PS-type, general (PSg)

(4.6)   In de krant wordt ◇ beweerd <u>dat hij dronken achter het stuur</u>
        <u>gezeten heeft</u>.   0 >= er >= het

        in the newspaper is ◇ claimed that he drunk behind the wheel sat
        has
        "It is claimed in the newspaper that he was drunk while he was
        driving."

PS-type, dummy object in active equivalent (PSdo)

(4.7)   Door haar vrienden wordt ◇ bewonderd <u>dat ze ook in moeilijke</u>
        <u>tijden vrolijk blijft</u>.   het>=0>= er

        by her friends is ◇ admired that she also in difficult times cheerful
        remains
        "Her friends admire her for remaining cheerful, even in difficult
        times."

4.3   Tasks

We gave our participants two tasks: an oral sentence imitation task and a
sentence preference task. We decided to use a sentence imitation task in
addition to a sentence preference task, because an oral sentence imitation

task can tap participants' implicit knowledge of grammar in a more direct way, without participants being aware of what they are doing exactly.

It has been shown that under high working memory demands, processing of sentences will be slower and less accurate (see e.g. Just & Carpenter, 1992). If a sentence is too long or complex for participants' linguistic processing and storage capacity to be retrieved in its entirety from working memory, reconstruction will take place. This may result in a "repetition" that is different from the original stimulus. If a stimulus sentence contains a grammatical feature that a participant has not acquired, this feature will unconsciously be changed in such a way that the sentence fits the participants' own grammar (see e.g. Bley-Vroman & Chaudron, 1994). For native speakers, therefore, ungrammatical targets are most informative, because native speakers are expected to change these into grammatical targets. For non-native speakers, grammatical targets can also be informative, because if non-native speakers change such a target, this could mean that the target is ungrammatical in the non-native speaker's grammar.

In order for this task to yield good results, sentences need to be slightly too long or complex to be retrieved in their entirety from working memory, but not too long or complex to handle. For advanced adult learners, targets (i.e. the grammatical features studied) should preferably be phonologically non-salient and occur in a non-salient position. Dummy subjects in non-initial position are therefore very suitable for an imitation task with very advanced adult learners.

Note that there are individual differences in working memory capacity (see e.g. Just & Carpenter, 1992). Because of this, people with greater working memory resources may be able to repeat sentences literally, while others with the same grammar but more limited resources may make changes in the same sentences. For this reason, only sentences that are not literally repeated can be used for analysis.

Sentence preference task
Grammaticality judgement tests are widely used in studies of ultimate attainment (see e.g. Birdsong, 1989; Johnson & Newport, 1989, 1991), because they have several advantages over other tasks. First, they are very suitable for measuring participants' command of a grammatical feature that is not frequent in the input (like dummy subjects). Second, this feature can be elaborately tested in a controlled experiment and second language learners cannot avoid using the feature under investigation by using a different construction type. Finally, grammaticality judgement tests can provide information both on grammatical and on ungrammatical sentences.

Judgement tests, however, have been criticised for, among other things, the strong influence of previous judgements and the metalinguistic nature of the task of judging grammaticality (see e.g. Birdsong, 1989; Schütze, 1996; Sorace, 1996). These are problems that apply specifically to judgement tests in which participants have to decide whether individual sentences are grammatical (or acceptable) or not. Sorace (1996) therefore advises not to use an absolute judgement test, in which participants have to indicate whether individual sentences are grammatical/acceptable or not. She advises the use of either magnitude estimation or a relative judgement test instead. In a magnitude estimation task, participants are first presented with an anchor item, which is used as a reference point for all other items. Participants can assign any value to this item (e.g. 100). For all subsequent items, participants have to judge how these items relate to the anchor item. If they find an item twice as acceptable as the anchor item, for example, they should assign that item two times the value of the anchor item (e.g. 200). If on the other hand, they find a sentence twice as unacceptable as the anchor item, they should divide the value of the anchor item by two and assign this value to this sentence (e.g. 50). All sentences are thus presented and judged individually, but relative to the anchor item. The results of such a test are assumed to be interpretable as ratio variables rather than as ordinal or interval variables. In the relative judgement test Sorace proposes, participants have to rank order two or more sentences for degree of acceptability.

The problem with the metalinguistic nature of absolute judgements also turned up in a preliminary study on dummy subjects we did with native speakers of Dutch (van Boxtel, 1999, unpublished data). Some native speakers had problems deciding whether a sentence that they would never use (e.g. a sentence with another dummy subject than the one they preferred) was grammatical or not. For these reasons, we used a sentence preference task with minimal sentence pairs both in a pilot study and in the final study. In this task, participants do not have to decide whether sentences are grammatical or not, but which sentence from a minimal pair they prefer (see Birdsong (1997), who also used this task in a study with very advanced late second language learners).

The sentence preference task (henceforth SPT) in our study was presented on a computer and consisted of 190 items: 124 test items and 66 filler items. Each test item contained a minimal sentence pair, in which only the dummy subject was different, as in (4.8).

(4.8)   a   Het is nu gebleken dat de brand door een ongeluk is ontstaan.
            b   Er is nu gebleken dat de brand door een ongeluk is ontstaan.
                it (a) / there (b) is now turned out that the fire by an accident is
                arisen
                "It has now turned out that the fire was caused by an accident."

For each predicate, all possible combinations of dummy subjects were tested. This means, for example, that for passive sentences with a sentential subject the same predicate occurred in three items: one with *het* versus *er*, one with *het* versus *0* and one with *0* versus *er*. In order to include the option of using *0*, we mainly used sentences with dummy subjects in non-initial position. Only for the AS-type did we also use items with *het* or *er* in non-initial position, because results from our pilot test suggested that in this position *er* could more easily be used than in non-initial position.

For the filler items we selected five construction types that we expected to be difficult for second language learners. We chose both construction types for which we expected a great deal of individual variation among native speakers, and construction types for which we expected relatively homogeneous results for the native speakers. Moreover, we selected items for which we expected a strong preference for one sentence to the other and items for which both sentences seemed to be equally (un)acceptable. The reason for this was to make sure that participants would use as many different points of the 7-point-scale (see below) as possible. We used the following five construction types, which were also used in experiments with native speakers of Dutch by van Dreumel (van Dreumel, 1999, unpublished data), designed to investigate the structure of verbal clusters in Dutch:
- Sentences with *zelf* (self) in different positions
- Sentences in which the infinitive marker *te* (to) was either present (sometimes in the wrong position) or absent
- Sentences in which prepositions were either attached to or detached from *er* or *daar* (there)
- Sentences in which the auxiliary was alternated (*be* versus *have*)
- Sentences with more than one verb in which different verb orders were presented

In the experiments by van Dreumel, sentences were taken from written corpora and were often very long. For our study, we adjusted these sentences or constructed new ones, which were shorter and contained lexical items with a higher frequency. Since these items were filler items, they were not as carefully constructed as the test items and some construction types contained very different items.

For each item in the sentence preference test, participants clicked a button on the following seven-point scale:

I strongly prefer   O   O   O   O   O   O   O   I strongly prefer
sentence A to sentence B   sentence B to sentence A

Participants were told that their reaction times would be measured, and that they should read the sentences carefully and then decide which sentence they preferred, without thinking too long about an item. They were also instructed to use the button in the middle when they found both sentences either equally good or equally bad. It was emphasised in the instructions that there were no good or wrong answers, and that participants should indicate which sentence they preferred, rather than what they thought was the norm in standard Dutch. Each item was presented on a separate screen (until participants clicked the 'proceed' button) and participants could not go back to previous items. Both the order of sentences within items and the order of items were randomised (the latter for each participant separately). After 85 items, participants were asked to take a five-minute break. The whole sentence preference test took between 25 and 78 minutes (Mean = 40 minutes), excluding the break. We also measured participants' reaction times. However, since we were unable to measure reliably how quick participants were in using a mouse, the reaction time data were too unreliable to use in any further analyses.

Sentence imitation task
The sentence imitation task in our study consisted of twenty sentences, which were recorded by a female native speaker of Dutch on a DAT recorder and played to the participants. Participants were instructed to repeat the sentences literally as soon as they had listened to the whole sentence. If participants were unable to repeat (part of) the sentence, the sentence was played again, with a maximum of two repetitions. Their reactions were recorded either on cassette tape or minidisk and analysed later. All participants did the sentence imitation task before the sentence preference task.

The imitation task consisted of twelve sentences containing a dummy subject (four sentences for each dummy subject) and eight 'filler' items. Some of the fillers contained a phonologically salient target. With the exception of some of the fillers, the sentences were based on items in a pilot study that we also used for the sentence preference test (see chapter 3). The length of items was adjusted in such a way that all items consisted of ten or eleven words and fifteen syllables. An example of an item in the sentence

imitation task, containing a dummy subject is given in (4.9). The target is printed in italics in this example. For an overview of all test items in the sentence imitation task, see chapter 6.

(4.9)   Tot nu toe komen *er* steeds meer mensen naar dit popconcert.
        until now come there still more people to this pop concert
        "So far, more and more people have come to this pop concert."

In one item of the active DP-type the dummy subject (*er* vs. *0*) was replaced by *het*, as is illustrated in (4.10). In all other items the dummy subject that was used in the sentence imitation task, also occurred in the sentence preference task.

(4.10) Ik vind het vervelend dat *het* boven een raam openstaat.
        I find it annoying that it upstairs a window open stands.
        "It bothers me that there is a window open upstairs."

The results from the native speakers on the sentence preference task showed that of the twelve dummy subject sentences six contained the dummy subject that most native speakers prefer in this context and six another dummy subject than the one preferred by most native speakers of Dutch (see chapter 6).
        One of the disadvantages of a sentence imitation task is that, because of its demanding nature, it cannot contain many items, so that participants cannot be thoroughly tested on their command of a particular grammatical construction. Another disadvantage of imitation tasks is that perfect imitations of the target cannot be interpreted, because they could result from direct retrieval from working memory without reconstruction taking place. This means that only part of the data can be analysed. These problems do not occur in a sentence preference task.

4.4   Analyses

Analyses of dummy subject items in the sentence preference task
To establish a native speaker norm for the analysis of the second language data from our SPT, we used the results from the native speakers on this task. For both the group and the main individual analyses, we established a native speaker pattern for each dummy subject pair in a given category based on the scores of the majority of the native speakers on most of the items in the category concerned, as indicated below. There were basically two native speaker patterns: either a clear preference for one dummy subject over the

other (i.e. either a raw score of 1 or 2 or a raw score of 6 or 7[19]) or a pattern in which one dummy subject was better than or equal to the other dummy subject (i.e. either raw scores between 1 and 4 or raw scores between 4 and 7). The native speaker patterns for each dummy subject pair for each category in the SPT are presented in table 4.2 below. Five items in the test were not considered in the analyses because they contained errors. These items are not included in this table. For information on which items were excluded from the analysis and how many native speakers had a deviant score for each item, see appendix C.

Table 4.2:   Native speaker patterns for each DS pair in the SPT

Category	DS pair	Nr. of items	Items incl. in analysis	Ns pattern	Raw scores
DPg	er – 0	11	10	0 > er	6-7
DPnst	er – 0	8	6	0 >= er	4-7
DPnsi	er – 0	8	5	er >= 0	1-4
ASg	het – er	20	16	het > er	1-2
ASg	het – 0	16	12	het > 0	1-2
ASg	er – 0	16	14	0 >= er	4-7
AScos	het – er	8	0	het / er	1-4
AScos	het – 0	4	0	none (0 / het)	1-7
AScos	er – 0	4	3*	0 >= er	4-7
PSg	het – er	4	3	er >= het	4-7
PSg	het – 0	4	2	0 >= het	4-7
PSg	er – 0	4	0	0 / er	4-7
PSdo	het – er	4	3	het > er	1-3
PSdo	het – 0	4	3	het >= 0	1-4
PSdo	er – 0	4	2	0 >= er	4-7

*   Items included in total, category excluded from the analysis

Items for which less than 90% of the native speakers had the above pattern, were excluded from the analysis. The reason for doing this was to have a clear and strict criterion for native-likeness. Three items with *er* and *0* in the general category of the passive sentential type (PSg) and ten items in the change of state category of the active sentential type (AScos) containing *het* were removed from the analysis on the basis of this criterion. Since after this procedure there were only three items left in the change of state category (AScos), we had to exclude the whole change of state category, as three items is not enough to compute meaningful statistic analyses. These three items were included, however, in the total score. The other categories contained 24 items for which more than 10% (= 4) of the native speakers deviated from the NS pattern for that item. These items were also excluded from the analyses.

---

[19]   There was one exception to this pattern: for the category of the Passive Sentential type with a dummy object in the active equivalent (PSdo) the NS pattern for *het* vs. *er* was 1-3.

Finally, three items were excluded because the predicates in these items behaved differently from the other predicates in the same category. There were thus 79 dummy subject items and 6 categories (DPg, DPnst, DPnsi, ASg, PSg and PSdo) that were included in the analyses.

In order to determine whether participants behaved like most native speakers of Dutch, for each of these 79 items participants were assigned a value of 1 if they had the native speaker pattern and a value of 0 if they had another pattern. The sum of these values for each category and in total was both used as the basis for the group analyses of the SPT and for individual analyses of the SPT.

For the group analyses we computed t-tests and ANOVAs on the basis of these total scores, followed by Games Howell post hoc tests (see § 5.2). For the individual analyses we transformed these scores into z-scores on the basis of the mean and standard deviation of the native speaker group. This enabled us to determine which participants fell within the native speaker range for each category of the SPT (see § 5.3). We considered everyone with a z-score between –1.96 and 1.96 to fall within the native speaker range (α = .05).

In order to test the hypothesis (derived from the CPH for grammar in SLA) that second language learners did not make the same distinctions as native speakers between the categories within each type, we did an additional analysis in which we computed z-scores based on differences in preference for certain dummy subject pairs between the categories within each type (see § 5.4). In this analysis, we used all items and categories in order to use all the information available in the data. For all statistical analyses we used α = .05 as criterion for significance.

Analysis of dummy subject items in the imitation task
For the analysis of the sentence imitation task we looked at deviations from the stimulus, especially in the target, for the 12 sentences containing a dummy subject. Sentences that were literally repeated were not considered in the analysis. The reason for this is that when someone repeats a sentence literally, one cannot tell whether this reflects his/her grammar or good working memory/processing skills. To determine which second language learners fell within the native speaker range, we compared the changes made in the target by individual second language learners to those made by the native speakers. If a second language learner made no change in the targets of the dummy subject sentences that was not also made by at least one of the native speakers, we considered this participant to fall within the native speaker range. All other participants were considered to fall outside the native speaker range. The reason for using this criterion was that changes other than the one expected on the basis of the results of the SPT were never made by a majority of the native

speakers, but were too numerous to discard as non-native-like responses. Because of the qualitative nature of the data and the small number of items, we did not perform a statistical analysis of the imitation task.

Filler items
Both the sentence imitation task and the sentence preference task also contained filler items, which mainly served to divert the participants' attention from dummy subject constructions and to make sure that the participants would use different parts of the scale (see also paragraph 4.3). Although these items were divided into categories, these were not as carefully investigated and the items were not as carefully constructed as the categories and items containing dummy subject constructions. We did not include any of these filler items in the analyses presented in chapters 5 and 6.

Chapter

    5

Results of the sentence

preference task

In this chapter, we will discuss the results of the second language learners on the sentence preference task (SPT) in comparison to the results of the native speakers of Dutch. We used various methods of analysis for our SPT, which will be briefly discussed in § 5.1. At the group level, we used t-tests and ANOVAs followed by post hoc tests. The results of these analyses are presented in § 5.2.

However, since our main research question is whether there are any individuals who have reached a native level of attainment, our focus will be on analyses that evaluate native-like attainment at an individual level. At this level, we computed z-scores based on the deviation from the native speaker pattern for each participant for each category (§ 5.3). Some of the (sub)categories in our sentence preference task could not be used for this analysis, because there was too much variation among the native speakers. For this and other reasons we performed an additional analysis in which we computed the difference between two categories of the same type. This analysis is reported in § 5.4. Finally, in § 5.5, we will look at relevant background characteristics in relation to the combined results of the analyses presented in paragraphs 5.3 and 5.4.

5.1 Analysis procedures

In chapter 4, we explained how we analysed our data. In this paragraph, we will briefly summarise the analysis procedures for the SPT. For an account of which items were excluded on which grounds, see chapter 4 and appendix C.

To establish a native speaker norm for the analysis of the second language data, we used the results from the native speakers in our experiment. For both the group and the main individual analyses, we established a native

speaker pattern for each dummy subject pair in a given category, based on the scores of the majority of the native speakers on most of the items in the category concerned. An overview of the patterns for each category in the analysis is given in table 5.1 below. For more details, see chapter 4, table 4.2. For examples of items in these categories, see chapter 3 or chapter 4, or appendix B for all dummy subject items.

Table 5.1: Native speaker patterns for each category in the SPT

Type	Category	Abbr.	Nr. of items	NS pattern
Active sentences with a noun phrase / DP as logical subject	general	DPg	10	*0 > er*
	non-specific subject + transitive predicate	DPnst	6	*0 >= er*
	non-specific subject + intransitive predicate	DPnsi	5	*er >= 0*
Active sentences with a sentential subject	general	ASg	42	*het > 0 >= er*
	change of state	AScos	3*	*0 / het >= er*
Passive sentences with a sentential subject	general	PSg	5	*0 >= er >= het*
	direct object in active equivalent	PSdo	8	*het >= 0 >= er*

* Items included in total, category excluded from the analysis

We excluded the AScos category because three items is not enough to compute meaningful analyses. These three items were, however, included in the total score. In total, therefore, 79 items were included in the analyses. For each of these 79 items participants were assigned a value of 1 if they had the native speaker pattern (with the right strength) and a value of 0 if they did not. The sum of these values for each category and for all categories together was used to compute t-tests and ANOVAs followed by post hoc tests for the group analyses (see § 5.2) and z-scores (score minus mean, divided by the standard deviation) for the individual analyses (see § 5.3). For the individual results, we additionally computed z-scores based on differences in preference for certain dummy subject pairs between the categories within each type (see § 5.4). All z-scores were based on the mean and standard deviation (SD) of the native speaker control group.

5.2 Group results

In this paragraph we will discuss the group results of the SPT. It should be noted that, if there were a critical period for syntax in SLA, one would expect major group differences between the second language learners and the native speakers in our experiment. On the other hand, we selected L2 participants who seemed not to differ from native speakers of Dutch (with

respect to grammar). In order to see whether or not they would behave in the same way as native speakers with respect to the dummy subject constructions in our test, we will discuss several group results from the SPT. First, we will look at the results from the native speakers versus all non-native speakers together, using t-tests. Subsequently, we will make a further distinction between the different L1 groups and look at results from post hoc tests based on analyses of variance (ANOVA). In these post hoc tests we will both compare the different L1 groups to the control group of native speakers and to each other, to see in what way and to what extent the L1 influences the level participants attained in Dutch.

Differences between the native speakers and all non-native speakers
Let us first look at the differences between the native speaker control group and all non-native speakers together, using t-tests per category and for the total score. Levene's tests showed that there were unequal variances for all the categories in the sentence preference test ($p < .05$). T-tests revealed that the differences between the native speakers of Dutch and all non-native speakers taken together were highly significant for all categories except DPnst (transitive predicates with a non-specific DP-subject). The t-values (equal variances not assumed) and their significance are presented in table 5.2 below. Because of the small number of items, AScos was excluded from this and subsequent analyses.

Table 5.2: T-tests for each category in the SPT

	NSs Mean	SD	NNSs Mean	SD	Mean difference	t	df	Sig. (2-tailed)
DPg (10)	9.6	.7	8.9	1.2	.71	3.296	68.212	.002*
DPnst (6)	5.4	.9	4.9	1.4	.46	1.820	74.996	.073
DPnsi (5)	4.8	.4	4.0	1.1	.77	4.358	55.198	.000*
ASg (42)	40.8	1.6	38.3	5.0	2.54	3.184	50.097	.002*
AScos (3)	2.9	-	2.6	-	.28	-	-	-
PSg (5)	4.7	.7	3.7	1.4	.96	3.983	63.371	.000*
PSdo (8)	7.7	.6	6.8	1.2	.89	4.316	63.285	.000*
Total (79)	75.8	2.6	69.2	7.6	6.59	5.365	51.427	.000*

SD = standard deviation; df = degrees of freedom; sig. = significance; * = sig. ($p < .05$)

Differences between the four groups
For the analysis of differences between the four (L1) groups of participants we used ANOVA followed by Games Howell post hoc tests. We used Games Howell post hoc tests, because Games Howell can handle unequal variances and because Games Howell is a rather liberal post hoc test, which means that differences between groups become significant more easily than for other post hoc tests for unequal variances. This means that, if the results are biased due to the choice of a particular post hoc test, the bias will be in

favour of confirming the CPH (differences between the native and the non-native speakers). Levene's tests again confirmed that there were unequal variances across groups for each category in the analysis ($p < .01$). Outcomes of the ANOVA were also significant for all categories, except DPnst. The F-value of the total sum of all items was also significant. The F-values and their significance are presented in table 5.3.[20]

Table 5.3: ANOVA for each category and the total score in the SPT

	L1 Dutch		L1 German		L1 French		L1 Turkish		Mean Square	F	Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD			
DPg (10)	9.6	.7	8.3	1.4	9.6	.6	8.8	1.1	7.749	8.974	.000*
DPnst (6)	5.4	.9	4.9	1.5	5.3	1.0	4.5	1.6	3.006	2.253	.088
DPnsi (5)	4.8	.4	3.9	1.3	4.1	1.1	3.9	.8	4.397	6.398	.001*
ASg (42)	40.8	1.6	40.0	2.4	39.5	3.4	34.9	7.0	117.400	10.353	.000*
AScos (3)	2.86	-	3.0	-	2.6	-	2.2	-	-	-	-
PSg (5)	4.7	.7	4.6	.7	3.8	1.2	2.6	1.5	15.884	16.769	.000*
PSdo (8)	7.7	.6	7.2	.7	6.9	1.2	6.2	1.5	8.040	9.425	.000*
Total (79)	75.8	2.6	71.9	3.8	71.8	5.5	63.2	9.7	543.944	22.071	.000*

* = significant ($p < .05$)

For the Games Howell post hoc tests we used a two-tailed test for significance for all group comparisons. We did this because of the discrepancy between the low proficiency expected by the CPH and the high proficiency of our learners, and because we had no predictions as to which L1 group would perform best. However, on the basis of previous studies (see chapter 2), we expected that if there were differences, the L1 German group and the L1 French group would perform better than the L1 Turkish group. For the DPnst category we do not report the results of the post hoc test (which showed no significant differences), because the F-value for this category was not significant. The results for all other categories (except AScos) are presented in table 5.4, as well as for the total over all items in the analysis. Results that are only significant in a one-tailed test are indicated by the symbol ^. It should be noted that such values only occur for differences between the native speakers versus the near-native groups. The means for the four groups are plotted in figures 5.1-5.7. Each group is indicated by the first letter of the L1.

---

[20]   Because of the unequal variances across groups for most of the categories, we also computed Welch and Brown-Forsythe. For these analyses, too, all categories except DPnst were significant at the .05-level.

Table 5.4: *Results of the Games-Howell post hoc tests for the SPT*

	DPg		DPnsi		ASg		PSg		PSdo		Total	
	Md	Sig.	Md	Sig.	Md	Sig.	Md	Sig.	Md	Sig.	Md	Sig.
D-G	1.28	.021*	.82	.115	.82	.628	.08	.982	.51	.080^	3.89	.009*
D-F	.01	1.000	.68	.131	1.35	.468	.88	.069^	.77	.129	4.02	.062^
D-T	.84	.080^	.83	.023*	5.90	.045*	2.07	.002*	1.47	.018*	12.66	.002*
G-F	-1.27	.027*	-1.33	.990	.53	.960	.80	.155	.27	.880	.13	1.000
G-T	-.44	.802	.01	1.000	5.08	.105	1.99	.002*	.97	.173	8.78	.035*
F-T	.83	.109	.14	.980	4.54	.184	1.19	.133	.70	.538	8.65	.048*

Md = Mean difference (= difference between the means of the two groups in the comparison)
* = significant ($p < .05$; two-tailed); ^ = only significant if tested one-tailed ($p < .10$)

As can be seen in tables 5.3 and 5.4 and in figure 5.1 below, the native speakers of Dutch have the highest average total score and the native speakers of Turkish the lowest and this difference is highly significant. The native speakers of German and French have a very high average total, which is significantly higher than the total of the L1 Turkish group. However, the average total of the L1 German group is significantly lower than the total of the native speaker control group. The difference between the native speakers of French and the control group is not significant (2-tailed). However, the native speakers of French do not have higher scores than the native speakers of German.
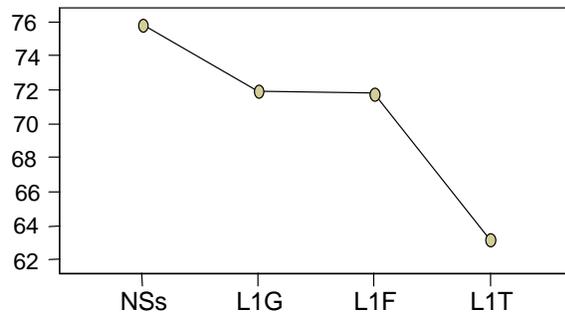
Figure 5.1: *Means plot for total (79 items)*

Let us now look at the scores for the separate categories in the SPT, presented in table 5.4 and figures 5.2-5.7 and compare them to the expectation patterns described in paragraph 3.4 and repeated below (the constructions that we expect to be easiest are on the left and the constructions for which we expect most problems are on the right):

85

L1 German: ASg > PS > DP
L1 French: DPg/DPnst > DPnsi > AScos > ASg > PSg > PSdo
L1 Turkish: DPg/DPnst > DPnsi > AScos > ASg > PSg > PSdo

It should be noted that we do not directly compare categories with each other. Rather, we compare the different L1 groups (including the native speakers of Dutch) for each category and discuss which differences are significant and which differences are largest. We would like to repeat here that the fact that the accuracy order for the L1 French and the L1 Turkish group is the same does not mean that we expect both groups to do equally well.
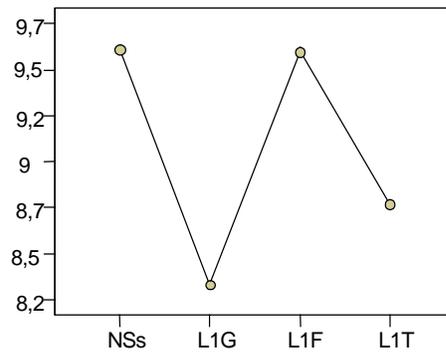
Figure 5.2:   Means plot for DPg (10 items)

Figure 5.2 clearly shows that, of the three learner groups, the German group has most problems with this category, as predicted. DPg is the only category for which the native speakers of German obtain significantly lower scores than both the native speakers of Dutch and the native speakers of French (whose scores are exactly the same). Although the L1 Turkish group also has a lower average score for this category than the native speakers, this is the only category (of the categories with a significant F-value) for which it does not differ significantly from the native speakers of Dutch (see table 5.4).
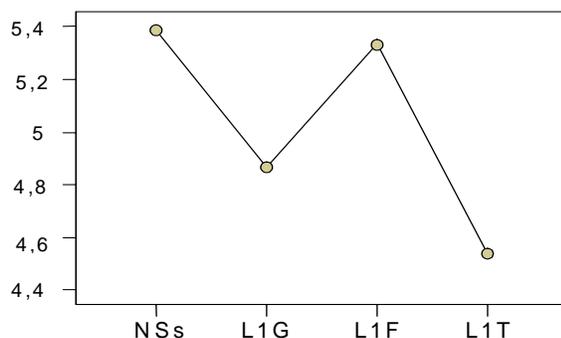
Figure 5.3:   Means plot for DPnst (6 items)

For DPnst the F-value (and post hoc tests) revealed no significant differences between groups. It is clear from the plot that the native speakers of French perform much like the native speaker control group. The native speakers of Turkish have lower scores than the native speakers of German for this category, which cannot be explained on the basis of our predictions, presented above, given the other results from the DP-type. If it were the case that native speakers of Turkish had problems with the combination of specificity and transitivity, one would expect them to be better on the DPnsi category, which they are not.
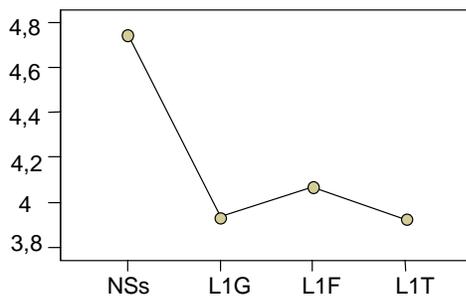
Figure 5.4:   Means plot for DPnsi (5 items)

For DPnsi, as shown in figure 5.4, the three learner groups have very similar scores. They all have lower scores than the native speakers of Dutch. The L1 Turkish group, however, is the only group for which this difference is significant (see table 5.4). All L2 groups obtained lower scores on this category relative to the control group than on the other two categories of the

DP-type. This was predicted for the native speakers of French and Turkish, but not for the native speakers of German. It seems as if using *er* in sentence types (for the L1 French and L1 Turkish group) or in a position (for the L1 German group) in which it does not occur in one's native language, might be the largest problem for L2 learners for the DP-type.
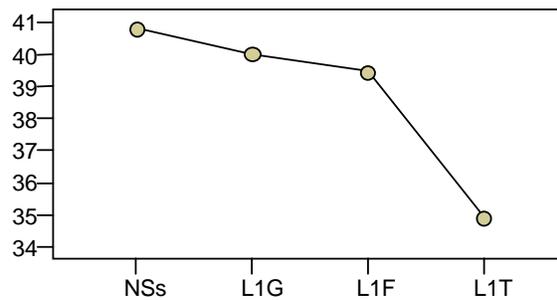
Figure 5.5:   *Means plot for ASg (42 items)*

As can be seen in figure 5.5, the results for ASg look rather different than the results for the DP-type. The native speakers of German have higher scores for this category than the native speakers of French. Both these groups have very high scores relative to the native speaker control group and they do not differ significantly from the control group for this category, unlike the L1 Turkish group. For the L1 German group, this result is as predicted. For the L1 French and L1 Turkish groups, these results for ASg are as predicted in relation to DPg, but not in relation to DPnsi.
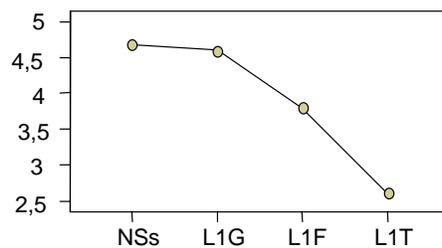
Figure 5.6:   *Means plot for PSg (5 items)*

If one compares figure 5.6 to figure 5.5, one can see that the order of the different L1 groups is the same. However, the native speakers of French and Turkish have lower scores relative to the control group for PSg than for ASg,

as predicted. On the other hand, the native speakers of German did not do worse on PSg than on ASg, contrary to what we predicted. This seems to be due to a ceiling effect. Because of this, not only the difference between the control group and the L1 Turkish group is significant for PSg, but also the difference between the L1 German and the L1 Turkish group. Moreover, the difference between the native speakers of French and the native speakers of Dutch approaches significance.
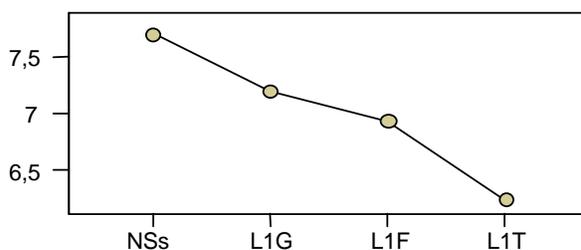
Figure 5.7:   Means plot for PSdo (8 items)

As can be seen in figure 5.7, the order of the four groups is the same as for ASg and PSg. All learner groups seem to perform slightly worse on this category than on ASg, as predicted. However, our prediction that the scores for PSdo would be worse than the scores for PSg for the native speakers of French and Turkish is not borne out. Whereas for the DP-type the greatest problem for all learner groups seemed to be to insert a dummy subject in a sentence or position in which it cannot occur in the L1, this does not seem to be the case for the sentential types. Of the three categories of the sentential types in the analysis, the one where native speakers of Dutch have a preference for *0* seems to be the most difficult one for all learner groups. These results cannot be accounted for by reliance on one preference pattern for all categories, because, if that were the case, the scores for PSg should have been much lower. We also cannot account for the results for PSg by making a distinction between judgements for *het* vs. *er* and judgements for *het* vs. *0*, because the data do not show an asymmetry in accuracy for this distinction.

If we combine the results of the different categories, presented in tables 5.3 and 5.4 and figures 5.2-5.7, we see that the native speakers have higher scores than all three learner groups for all categories, except DPg and AScos. The L1 German and L1 French groups seem to behave more like native speakers of Dutch than the L1 Turkish group, as expected. The results for the general pattern of the Active-sentential type seem to correlate best with the

typological distance between the first languages involved.

For the native speakers versus the L1 Turkish group all differences are significant, except the one for DPg, which is one of the categories we expected to be easiest for the L1 Turkish group. For the native speakers versus the L1 French group, on the other hand, none of the differences are significant, and for the native speakers versus the L1 German group, only the difference for DPg is significant, but it is clear that in general the performance of the L1 German group is worse for the DP-type than for the other types. This confirms our predictions for the DP-type on the basis of German. However, we could not test our prediction for the two sentential types with our participants, because of a ceiling effect at this level.

Whereas the native speakers of German seem to have most problems with the DP-type, the native speakers of French and Turkish, on the other hand, seem to have most problems with the PS-type, as predicted on the basis of relevant properties of French and Turkish. However, on the basis of properties of the L1, we cannot explain why they have relatively higher scores for PSdo than for PSg and higher scores for ASg than for DPnsi (see also paragraph 3.4).

It should be noted that, even though the native speakers have higher scores for all categories than the learner groups, many differences are not significant, which is a problem for the CPH for syntax in SLA. In the next paragraphs, we will take a closer look at individual results in search of crucial evidence with respect to the CPH for syntax in SLA. If the differences found above can also be found in each L2 speaker, this would support the CPH. If, on the other hand, there are L2 speakers that cannot be distinguished from the native speakers of Dutch on all aspects of the SPT, this constitutes counterevidence against the CPH for syntax in SLA.

5.3   Individual results for categories in the sentence preference task

To statistically identify L2 speakers that fell within the native speaker range, we computed z-scores for each category in the sentence preference task (except change of state) on the basis of the pattern scores described in paragraphs 4.4 and 5.1. We considered everyone with a z-score between -1.96 and 1.96 to fall within the native speaker range ($\alpha = .05$). This criterion (rounded off to $-2 < z < 2$) has also been used by Flege, Munro and MacKay (1995) and by Bongaerts, van Summeren, Planken and Schils (1997). By looking at each category separately, we can see whether L2 speakers behave in a native-like way on all aspects of dummy subject constructions in our test and if not, where exactly problems occur. It should be noted that the criterion

for falling within the native speaker range was very strict in our study. Although for each category less than 10% (= 4) of the native speakers had a z-score lower than -1.96, there were in fact 8 native speakers with such a score for at least one category. This means that it is very likely that we underrepresent the number of non-natives who behave like natives with respect to dummy subject constructions in Dutch. Even though such a strict criterion may not be completely fair to the second language learners, it is very suitable for falsification of the hypothesis that no late second language learner should be able to attain a native level in L2 grammar. The results are presented in figures 5.8 and 5.9 below. The individual raw scores and z-scores are presented in appendix D. The first letter under each column in figures 5.8 and 5.9 refers to the L1 of the participant group (Dutch, German, French, Turkish), SL refers to all second language learners together. Z-scores lower than –1.96 are labelled "other pattern" in these figures. Figure 5.8 shows the results for the DP-type. It should be noted that in this figure (as well as in figure 5.9), the y-axis indicates the percentage of participants within each group with a certain pattern (because the number of participants is unequal across groups), whereas the absolute numbers of participants with a certain pattern are given in the bars.
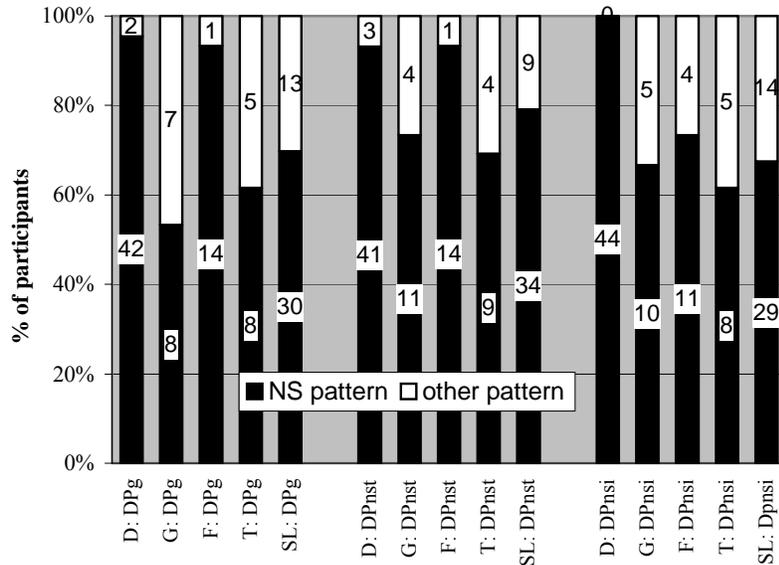
Figure 5.8:   Individual results for the categories of the DP-type

One can see in figure 5.8 that, of the 43 second language speakers, 30 speakers performed within the NS range for DPg, 34 for DPnst and 29 for DPnsi. In the L1 French group, more participants scored within the NS range than in the L1 German and L1 Turkish group. Participants in this group who did make mistakes mainly made them in the DPnsi category, which means they prefer *0* to *er* more often than native speakers of Dutch. This could be attributed to the influence of French. In French fewer predicates with a noun phrase subject can get a dummy subject (*il* or *ce*) than in Dutch (see § 3.4). For this type (unlike for the other two types), the native speakers of German and Turkish perform about equally well, in spite of the huge differences between the two languages and their difference in typological distance to Dutch. Turkish totally lacks dummy subjects, but it has a grammatical notion of specificity, which may help Turkish learners of Dutch in making a distinction between the DPnsi category and the other categories (see §3.4). This might explain why more native speakers of Turkish perform in a native-

like way on this type than on the other types (see also figure 5.9). German, on the other hand, has dummy subjects for constructions with noun phrase subjects, but only in initial position, and these are not limited to specific (or even definite) subjects. Moreover, specificity does not seem to be grammaticalised in German (see § 3.4). This might explain why about half the native speakers of German do not have a native-like preference for *0* for the DPg category.

For the sentential types (AS and PS), the results for the different L1 groups look rather different, as can be seen in figure 5.9.
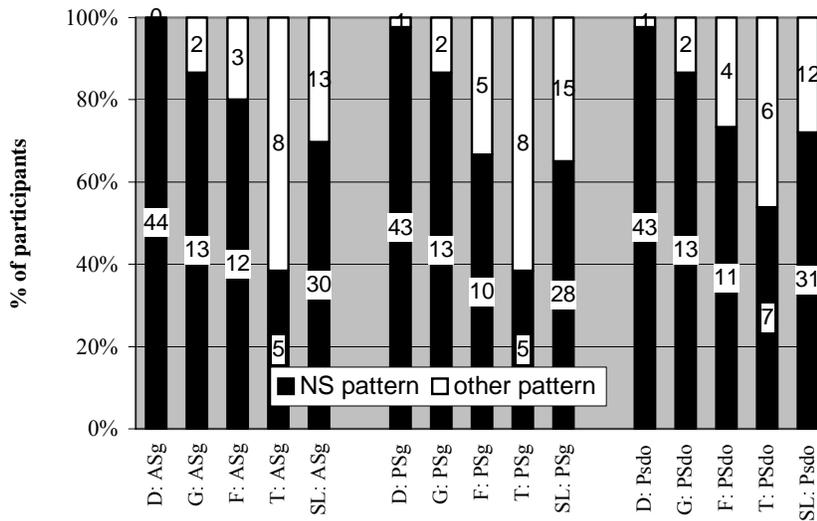
Figure 5.9:   Individual results for the AS- and PS-type.

Again, there are a substantial number of speakers in each non-native speaker group that fall within the native speaker range for all categories. For the ASg category, 30 out of 43 second language speakers fall within the native speaker range. For the two categories of the PS-type, there are as many as 28 (PSg) and 31 (PSdo) second language speakers who fall within the native speaker range.

As we have seen in paragraph 3.4, German has dummy subjects both in active and in passive sentences with a sentential subject and it also has

93

dummy objects. It is not very clear from the literature, though, which predicates can have a dummy subject (in non-initial position) in sentences with a sentential subject in German. French has dummy subjects, but not (normally) in passive sentences, and it does not have dummy objects. In French active sentences the use of dummy subjects is also more limited than in Dutch. Turkish does not have any dummy subjects or objects at all. It should be noted, though, that even though both German and French have dummy subjects and dummy objects, they do not have an equivalent for the *het-er* distinction in Dutch, and native speakers of those languages learning Dutch still have to find out which dummy subject is preferred in which construction for both the AS-type and the PS-type.

These differences between German, French and Turkish could account for the fact that for the sentential types there are slightly more native speakers of German who fall within the NS range than native speakers of French and much more than native speakers of Turkish (contrary to the results for the DP-type). These differences between the source languages cannot account, however, for why so many of the French and Turkish participants fall within the native speaker range. It is also unclear why the Turkish participants have better scores on the PSdo category than on the PSg category (see also the group results in the previous paragraph).

In figure 5.10 we present the percentage and number of participants that fall within the native speaker range on a total score, based on all 79 items in the analysis.
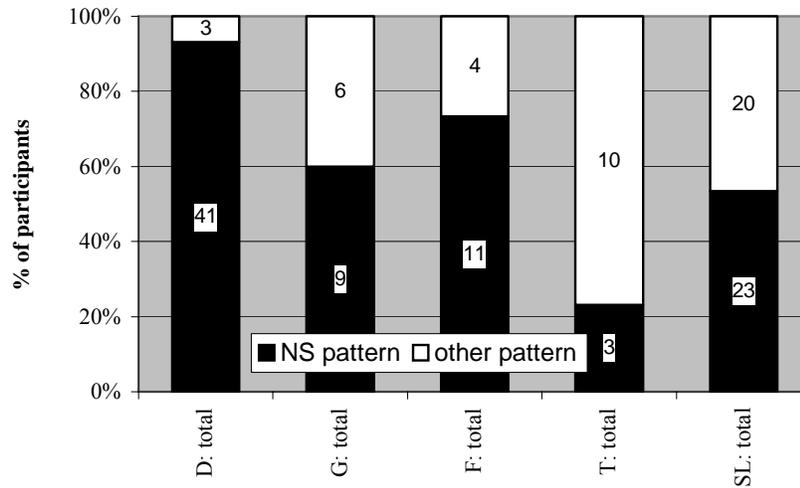
Figure 5.10: *Individual results for the total based on 79 items*

As can be seen in figure 5.10, the number of participants within the native speaker range is somewhat lower than the average for the separate categories, especially for the L1 Turkish group, but still 53% of all second language speakers (23 out of 43) score within the native speaker range.

If we combine the individual results for all categories in the sentence preference task, it turns out that there are 36 native speakers and 8 second language speakers who have the native speaker pattern for each category: three from the L1 German group (ID 59, 62 and 64), four from the L1 French group (ID 68, 71, 72 and 73) and one from the L1 Turkish group (ID 83). All these participants also score within the native speaker range on the total score. These results show that it is possible for late learners (even when they have a typologically very distant L1) to reach a native level in aspects of L2 grammar that are difficult to acquire, for which no explicit rules are available, and for which learners cannot (sufficiently) rely on their L1.

5.4   Additional individual analyses

In the analyses described in the previous paragraph, there was a certain overlap in judgements for different categories of the DP-type and the PS-type: both DPnst and DPnsi included the pattern "*er = 0*" and both categories of the PS-type included the pattern "*het = 0*". Moreover, we excluded the category AScos, because of the (relatively) large amount of variation for items within this category. It is theoretically possible, therefore, that participants who fell within the native speaker range in these analyses had a judgement pattern for certain dummy subject pairs that did not distinguish in a native-like way between categories within each type.

To test whether the participants with a native speaker pattern for all categories also made the same distinction between the different categories of each type as the majority of the native speakers, we computed additional analyses with all items, except the ones that contained a predicate that behaved deviantly. In these analyses we computed z-scores based on difference scores for relevant dummy subject combinations between two types of the same category.

For the DP-type, we looked at the difference between DPnst and DPnsi (*er* vs *0*) to see whether all second language speakers with a native speaker pattern according to the previous analyses made a distinction between transitive and intransitive predicates, rather than having no preference for both categories. This turned out to be the case (all participants with a native speaker pattern for each category had a z-score between −1.96 and 1.96 for this analysis). There was one native speaker, however, with a native speaker pattern for each category who had a z-score lower than −1.96.

For the change of state category of the AS-type (AScos), the native speaker control group had no clear pattern and nine native speakers even had an inconsistent pattern, but all participants in this group had a weaker preference for *het* to *0* for this category than for the general category. We therefore computed z-scores based on the difference between the preference for *het* versus *0* for AScos and ASg. Although all native speakers had a stronger preference for *het* for the ASg category, this difference was so low for two native speakers that it resulted in a z-score lower than −1.96 and so high for two other native speakers that it resulted in a z-score higher than 1.96. The Turkish participant who had a native speaker pattern for each category also made a greater distinction than most (43) of the native speakers, resulting in a z-score of 2.2. Given the huge amount of variation between native speakers and between predicates on the change of state category, it is difficult to interpret what such deviations really mean.

Finally, for the PS-type we looked at the difference between the preference for *het* versus *0* for PSg and PSdo, because the native speakers

had an opposite preference for these dummy subjects for PSg versus PSdo, both including no preference for either dummy subject: *0 >=het* for PSg and *het >= 0* for PSdo. It turned out that there was one native speaker who made this distinction between these two categories, but who made a smaller difference than the majority of the native speakers, resulting in a z-score of -2.0. One of the French participants with a native speaker pattern on all categories in the sentence preference task (ID72) had no preference for either dummy subject for both PSg and PSdo (z = -2.5).

If we combined the results for all native speakers on these three difference scores, there would be 8 native speakers with at least one deviant z-score. Of these eight native speakers, there were only three who also had a deviant pattern for one of the categories of the SPT. There are, therefore, a total of 13 native speakers out of 44 (=30%) who in some way deviate in their behaviour from the majority of the native speakers. To consider all these deviations as non-native like behaviour would not be a meaningful and fair criterion for native-likeness.

Considering this, we will consider all L2 participants with a z-score between -1.96 and 1.96 for each category of the SPT to fall within the native speaker range and we will distinguish between these participants and all other L2 participants when looking at background characteristics.

5.5   Background characteristics

In the previous paragraph, we showed that for each analysis there are participants in all groups that fall within the native speaker range. In this paragraph, we will look at relevant background characteristics and their relation to the results presented in paragraph 5.4.

First, it is clear that the possibility of reaching a native-like command of dummy subject constructions in Dutch is not limited to learners with a typologically similar L1, or to learners with an L1 that is very similar to Dutch with respect to dummy subject constructions. In table 5.5 we compare some other background characteristics of those participants that fall within the native speaker range (within NS range) with those that do not (outside NS range). The questionnaire we used to elicit information about L2 participants' background characteristics can be found in appendix A. For an overview of the results from all (quantifiable) background characteristics, see appendix E.

Because length of residence (LoR) by itself does not say much about how much input individuals receive, we computed a new variable by multiplying the participants' number of years of residence in the Netherlands by the proportion of Dutch use relative to use of their L1 (self-report on a

five-point scale ranging from approximately 10% to 90%)[21]. This new variable, which we will call "input", can be taken as a very rough estimation of the amount of input the second language learners received. For self-reported language skills (speaking, listening, writing, reading, grammar and overall proficiency) we only report the results for those skills for which there was a significant difference between the two groups (see below), namely speaking and writing for Dutch and listening and reading for the participants' best other language. For all skills in all languages, participants had to indicate on a five-point scale, ranging from very poor to very good, how proficient they were. The reason for looking at the participants' best other language was to see whether the participants that were best in Dutch were also better in other languages, for which they usually had not had a very substantial amount of input. For the L1, the differences between the two groups were so small for all skills that we did not include them in the table.

Table 5.5: Comparison of some background characteristics

NS range	AoA	LoR	Input	Age	Sp. D	Wr. D	Li. L3	Re. L3
within (8)	23	20	14	45	4.5	4.0	4.4	4.4
	(12-35)	(10-42)	(3-38)	(33-55)	(4-5)	(2-5)	(3-5)	(4-5)
outside (35)	21	22	13	44	4.1	3.4	3.3	3.5
	(12-35)	(4-50)	(2-39)	(23-73)	(3-5)	(2-5)	(1-5)	(1-5)
total (43)	22	22	13	44	4.2	3.5	3.5	3.7
	(12-35)	(4-50)	(2-39)	(23-73)	(3-5)	(2-5)	(1-5)	(1-5)

AoA	= Age of arrival in the Netherlands
LoR	= Length of residence in the Netherlands (in years)
Input	= LoR (in years) * the proportion of Dutch spoken relative to the L1
Sp. D	= self-estimated level of proficiency in speaking in Dutch
Wr. D	= self-estimated level of proficiency in writing in Dutch
Li. L3	= self-estimated level of proficiency in listening in best other language
Re. L3	= self-estimated level of proficiency in reading in best other language

To see whether the differences reported in table 5.5 were significant or not, we computed t-tests in which we compared the within NS range group to the outside NS range group. Age at the time of testing is very similar for both groups. Interestingly, the participants who fall within the native speaker range have a slightly higher mean age of arrival (AoA), although this difference is not significant and the ranges for both groups are the same. Also, they have had hardly any more input and the difference in input and LoR is not significant. It should be noted, though, that the real amount of input is hard to determine, because it is hard to estimate as it can change

---

[21]   We also asked participants about their use of Dutch and their L1 relative to other languages. Since almost all participants did not use other languages more than 10% of the time, we did not take this variable into account when computing the new input variable.

over the years and because the variable is based in part on a relatively rough measure (a five-point scale for proportion of Dutch used relative to the L1). There may also have been qualitative differences in the input. However, even under favourable conditions, three years of "full input" (30% use of Dutch during nine years LoR (for one of the French participants)) seems to be a rather small amount to reach native-like attainment in such a difficult area of grammar.

The self-reported skills in Dutch are on average higher for the group that falls within the native speaker range. These differences are significant for speaking and writing ($p < .05$ (one-tailed)), but surprisingly not for self-reported proficiency in grammar. The self-reported skills in the participants' best other language are also on average higher for the group that falls within the native speaker range. These differences are significant for listening and reading ($p < .05$ (one-tailed)), but again not for self-reported proficiency in grammar in these languages. This latter result is difficult to interpret, because we do not have enough information about differences in input for the two groups for these languages. We do know, however, that most L2 participants learned these languages in school and did not receive substantial amounts of input in these languages. Although this information is too limited to draw any conclusions from, it seems to suggest, especially in combination with the results on input, that a large amount of input is insufficient for late learners to reach a native level in L2 grammar, and that something like language aptitude plays a role as well.

None of the t-tests for skills in the native language were significant. Although self-report on a five-point scale is of course a very rough measure of L1 proficiency, these results seem to suggest that one can reach a very high level in L2 grammar without it affecting the normal use of the L1.

Of the five statements we presented in the questionnaire about the L2 participants' attitudes towards Dutch and their native language, only the statement "I like to learn languages" yielded significant differences (t = 1.975; $p = .03$ (one-tailed)).

If we look at variables from the questionnaire that are more difficult to quantify, perhaps the most striking result is that all participants that fell within the native speaker range had either studied a language or worked in an environment where language plays an important role. Of these eight participants, three were translators, one was a teacher of her L1 (French), one was a linguist, one had worked at Maison Descartes (an institute in Amsterdam with the literary and academic mission to promote French culture and stimulate French-Dutch dialogue), one had worked in a German department at a Dutch university and one had studied French. It should be noted, though, that in the other group there were also 23 participants who had some kind of special background with respect to language and two of

them were translators. The eight participants within the native speaker range were also very highly educated. Seven of them had an academic degree and the other one had had pre-university education. Even though most participants in our study had a high level of education (only five participants had a level below higher professional education), the relation between level of education and NS range is significant (t = -2.95 (equal variances not assumed); p = .01 (two-tailed)). Whereas linguistic background seems to play a role, the data with respect to attending Dutch lessons seem to indicate that, for acquiring dummy subject constructions, it does not matter much whether participants receive specific instruction in Dutch or not. More empirical evidence is necessary to shed light on the role of education and metalinguistic awareness in reaching a native level in L2 grammar. It is conceivable that metalinguistic awareness makes people more focussed on form and on phonologically less salient grammatical elements such as dummy subjects, which would enable them to process information that normally goes unheeded in late L2 acquisition.

It is also striking that there should be many more women than men in the within NS range group (1 man and 7 women), whereas there are more men in the outside NS range group (19 men and 16 women). This result is difficult to interpret, however, because the proportion of men and women was very different for the different L1 groups: in the L1 German group there were seven men and eight women, in the L1 French group there were three men and twelve women, and in the L1 Turkish group there were ten men and three women. The fact that women perform better than men may, therefore, be an artefact of the fact that the L1 French group is better than the L1 Turkish group.

Finally, we will look at the situations in which participants use Dutch and their native language. Of the eight participants within the NS range, there are five who speak Dutch at home, there is one who speaks Dutch everywhere except at home and two who speak Dutch at work and with Dutch friends (one of them also speaks Dutch with relatives). The use of the native language in this group ranges from only with relatives (outside home) to at work, at home and with relatives. In the group that falls outside the native speaker range, the use of Dutch ranges from use with relatives (outside home) and friends to "everywhere and with everyone in the Netherlands". The use of the native language in this group varies from "only when there is no other solution" to with relatives (outside home) and friends and either at home or at work. There are no people in either group who speak their native language both at work and at home, except when they talk about emotional issues (at home).

Chapter

6

Results of the sentence

imitation task

In this chapter, we will discuss the results of the sentence imitation task (SIT). We used this task in addition to the metalinguistic sentence preference task (SPT), because it can tap participants' implicit knowledge of grammar (see also paragraph 4.3). If it can be shown that second language learners in an imitation task make the same changes that are made by native speakers, this provides additional evidence against the CPH for syntax in SLA.

After we explain how we analysed the data (§6.1), we will first discuss the results of the native speakers in our experiment on the items with dummy subjects (in §6.2). Finally, we will compare the results of the non-native speakers to the results of the native speakers to determine whether there are any non-native speakers who fall within the native speaker range.

6.1   Analysis

For the analysis of the imitation task we looked at deviations from the stimulus, especially in the target. For 12 of the 20 sentences the target was a dummy subject. For the other 8 sentences the targets involved different grammatical constructions (see chapter 4) and different degrees of phonological salience. An overview of all dummy subject items in the SIT, and the category to which they belong, is given in (1)-(12) below. For a presentation and discussion of the categories, see §3.2.

Item 1, **DP-type, general (DPg)**
(1)   Vandaag zijn alle leerlingen van groep vijf naar het zwembad.
         today are all pupils of group five to the swimming pool
         "Today all pupils of group five went to the swimming pool."

Item 3, **DP-type, non-specific + intransitive predicate (DPnsi)**
(2)   Tot nu toe komen er steeds meer mensen naar dit popconcert.
        until now come there still more people to this pop concert
        "So far, more and more people have come to this pop concert."

Item 4, **AS-type, general (ASg)**
(3)   Het overkomt Tessa zelden dat zij te laat op school is.
        it happens Tessa rarely that she too late at school is.
        "It rarely happens to Tessa that she is late for school."

Item 6, **AS-type, general (ASg)**
(4)   Met dit weer zou gek zijn als de voetbalwedstrijd niet doorging.
        with this weather would strange be if the football match not through
        went
        "With this weather it would be strange if the football match was
        called off."

Item 8, **DP-type, non-specific + intransitive predicate (DPnsi)**
(5)   Ik vind het vervelend dat het boven een raam openstaat.
        I find it annoying that it upstairs a window open stands
        "It bothers me that there is a window open upstairs."

Item 10, **DP-type, non-specific + transitive predicate (DPnst)**
(6)   Op Peters verjaardag drinken er nooit zo veel mensen bier.
        on Peter's birthday drink there never so many people beer.
        "At Peter's birthday party there are usually not many people who
        drink beer."

Item 11, **PS-type, dummy object in active equivalent (PSdo)**
(7)   Door haar vriend wordt het bewonderd dat ze altijd vrolijk blijft.
        by her friend is it admired that she always cheerful remains
        "Her friend admires her for always remaining cheerful."

Item 14, **PS-type, dummy object in active equivalent (PSdo)**
(8)   Hij denkt dat er geheim gehouden wordt dat Max terug is.
        he thinks that there secret kept is that Max back is
        "He thinks it is kept a secret that Max is back."

Item 15, **DP-type, non-specific + intranstive predicate (DPnsi)**
(9)   Normaal staan zeven borden op de hoek van deze tafel.
        normally stand seven plates on the corner of this table
        "There are usually seven plates on the corner of this table."

Item 17, **AS-type, change of state (AScos)**
(10)   Er valt haar ineens op dat Fatima een nieuwe bril heeft.
        there falls her suddenly on that Fatima a new glasses has
        "She suddenly notices that Fatima has new glasses."

Item 18, AS-type, change of state (AScos)

(11)   Volgens Bas is het nu gebleken dat zij onschuldig is.
         according Bas is it now turned out that she innocent is
         "According to Bas it has now turned out that she is innocent."

Item 19, DP-type, general (DPg)

(12)   In de laatste wedstrijd heeft Luc het winnende punt gescoord.
         in the last match has Luc the winning point scored
         "In the last match Luc scored the winning goal."

Sentences that were literally repeated were not considered as evidence for native-like behaviour. The reason for this is that when someone literally repeats a sentence, it cannot be determined whether this reflects his/her grammar or good working memory / processing skills (see also §4.3).

         We manually coded all responses, using the following categories:

a   No change
b   Change, not related to target (e.g. *komen* → *kwamen* in item 3)
c   Predicted change in target (e.g. *0* → *het* in item 6)
d   Change in target, while no change predicted (e.g. *het* → *ze* in
         item 11)
e   Other change in target than predicted (e.g. *het* → *0* in item 8)
f   Large part of the sentence missing or substantially changed (e.g.
         dat ze altijd vrolijk blijft → *dat* in item 11)
g   Part of the sentence with the target missing or changed substantially
         (e.g. *zou gek zijn* → *zou je gek zijn* in item 6)
h   Missing (not on tape or not intelligible)

In order to be able to compare the results of the SIT to the results of the SPT, we only looked at changes made in dummy subjects. Changes not related to the target (b) were, therefore, not considered in the analysis.

         When a second language learner had a response in which important information was missing (f/g), we considered this participant to fall outside the native speaker range. To determine which (other) second language learners fell within the native speaker range, we compared the changes made in the target by individual second language learners to those made by the native speakers in our study. When a second language learner made either no change in the targets of the dummy subject sentences or only changes that were also made by at least one of the native speakers, we considered this participant to fall within the native speaker range.

         There were several reasons why we used this criterion which includes all changes made by native speakers. First, for most items the majority of native speakers made no changes at all (see §6.2). If we had used

a criterion based on the behaviour of the majority of the native speakers in our experiment (as suggested by M. Long, personal communication, September 2003), therefore, we should have considered all changes in at least ten of the twelve dummy subject items to fall outside the native speaker range. Had we done that, we would have had to consider the majority of native speakers to behave in a non-native-like way (see § 6.2).

D. Birdsong (personal communication, 2003) suggested we might use grammaticality of the response as a criterion for native-likeness. We investigated this possibility in a small experiment (van Boxtel, 2003, unpublished data). In this experiment, we elicited judgements from native speakers of Dutch without a linguistic background on a subset of responses made by native speakers and non-native speakers in the SIT. However, there was no strong consensus among the judges as to whether certain responses were grammatical or not. For example the response to item 6, which contained *dat* instead of *als*, was considered grammatical by two native speakers and ungrammatical by three native speakers of Dutch. This response is presented in (13) below.

(13)   Met dit weer zou het gek zijn dat de voetbalwedstrijd niet doorging.
        with this weather it would strange be that the football match not
        through went
        "With this weather it would be strange that the football match was
        called off."

Because of this lack of consensus, we could not use grammaticality of the responses as a criterion for falling within the native speaker range.

Because of the nature of the data and the small number of items, we did not perform a statistical analysis on the imitation task.

6.2   Results of the native speakers of Dutch

Table 6.1 gives an overview of the responses from the native speakers to all dummy subject items in the SIT. There were in total 528 responses (44 participants x 12 items). Two of these 528 responses were not recorded on tape due to technical problems and there was one response in which the target was not intelligible. Columns 4-12 represent the number of native speakers that gave a certain response (see § 6.1). It should be noted that changes not related to the target (b) can occur together with changes in the target. Because of this, the sum of the coding categories exceeds the total number of native speakers (44) for some items and the total number at the bottom of the table (536) exceeds the total number of responses. The last

row represents the percentage of responses in each column relative to the total number of valid responses (525). Because of the difference in the number of coding categories used and the number of valid responses, the percentages in the last row add up to more than 100%.

Table 6.1:   Responses by native speakers of Dutch

Item	Target	Predicted change	no change	changes made in target						other changes		mis- sing
			a	c	d	e	g	total	b	f	
1	0	none	42	-	0	0	0	0	0	0	2
3	er	none	38	-	0	0	0	0	6	0	0
4	het	none	42	-	0	0	0	0	2	0	0
6	0	het	22	13	-	1	1	15	10	0	0
8	het	er	12	27	-	4	0	31	2	0	1
10	er	0	23	21	-	0	0	21	0	0	0
11	het	none	39	-	2	0	0	2	2	1	0
14	er	het	27	13	-	3	0	16	1	0	0
15	0	er	40	1	-	0	0	1	3	0	0
17	er	het	24	18	-	0	0	18	5	0	0
18	het	none/0	40	1	0	0	0	1	3	0	0
19	0	none	41	-	0	0	0	0	3	0	0
Total (536)			390	94	2	8	1	105	37	1	3
% of (525) responses			74%	18%	0%	2%	0%	20%	7%	0%	-

The native speakers repeated 74% of all sentences literally. For four items, the target was never changed. Only for item 8 did a majority of the native speakers change the target. For all other items, the majority of native speakers did not change the target, even when it was not the one preferred by most native speakers. The changes that were made, were mainly made in the targets (105 changes in the target vs. 38 other changes). The items in which the native speakers of Dutch made unpredicted changes in the target were items 6, 8, 11 and 14. Items 11 and 14 belonged to the PSdo category. Why native speakers had problems with this category in the SIT is unclear. In the SPT the native speakers did not seem to have any problems with this category (see chapter 5). In item 8, the target in the stimulus seems to be less grammatical than the targets in the stimuli for the other items. This might also be the case for item 6. Moreover, most unpredicted changes were made in items in which a change in target was predicted. This suggests that the native speakers have some problems in processing dispreferred or less grammatical sentences.

Although most changes were made in phonologically non-salient words, some native speakers also changed words that are phonologically rather salient, like *voetbalwedstrijd* (item 6), *onschuldig* (item 18) and *punt*

105

(item 19). An overview of all the changes made by the native speakers is presented below. The changes in the target are underlined. The numbers in parentheses represent the number of native speakers who made the change. All unpredicted changes in the target (d/e/g) are printed bold.

Changes made by the native speakers in the SIT

1   -
3   *dit* → *het* (it; 5); *komen* → *kwamen* (came; 1)
4   *is* → *komt* (come; 1); *op school* → *0* (1)
6   *0* → *het* (it; 13); ***zou gek zijn* → *zou je* (you) *gek zijn* (1)**;
        als → *dat* (that; 5); *zou gek zijn als* → *zo gek dat* (so strange that; 1);
        niet → *0* (1); *voetbalwedstrijd* → *voetbal* (football; 1);
        weer → *uur* (hour; 2)
8   *het* → *er* (there; 27) / ***0* (2)**;
        ***dat het* → *dat em* (that him; 1) / *0* (1)**;
        vind het → *vind hetbut* (non-existing word; 1) / *vinte* (non-existing word; 1)
10   *er* → *0* (21)
11   ***het* → *ze* (she; 1) / *er* (there; 1)**
        dat - blijft → *dat* (1); *vrolijk* → *zo* (so) *vrolijk* (2)
14   ***er* → *het* (it; 13) / *erg* (very; 1) / *0* (1) / *een* (a; 1)**
        geheim gehouden wordt → *geheim geworden houd* (1)
15   *0* → *er* (there; 1)
        deze → *de* (the; 1); *de hoek van deze tafel* → *deze hoek van de tafel* (this corner of the table; 2)
17   *Er* → *het* (it; 18)
        ineens → *opeens* (suddenly; 4) / *nu* (now) *ineens* (1)
18   *het* → *0* (1);
        onschuldig → *erg* (very) *schuldig* (1) / *nu* (now) *onschuldig* (1);
        nu in other position (1)
19   *heeft* → *heb* (colloquial form; 1); *punt* → *doelpunt* (goal; 2)

Two types of changes are most frequent in the data: changes that make the sentence more grammatical (changes in the targets, e.g. *het* → *er* in item 8) and semantic changes (e.g. *ineens* → *opeens* in item 17). There are, however, also a few slips (e.g. *geheim gehouden wordt* → *geheim geworden houd* in item 14), a few non-existing words (e.g. *hetbut* in item 8) and a few sentences that are clearly ungrammatical (e.g. *Met dit weer zo gek dat de voetbalwedstrijd niet doorging* (item 6)) in the responses. This makes it difficult to determine whether responses from second language learners are native-like or not.

6.3   Results of the non-native speakers

Since we did not compute any statistical analyses for the SIT, we cannot tell whether the non-native speaker groups were significantly different from the native speaker control group. However, in our presentation of the data, we will present the results both in tables that give an overview of the number of changes made by each group for each item (in tables similar to table 6.1), and in tables that show how individual non-native speakers responded to each item. These last tables show how many non-native speakers fall within the native speaker range.

Considering the results of the native speaker control group and their problems discussed in the previous paragraph, we decided to consider a second language learner to fall within the native speaker range:
- if he/she made no change in the targets of the dummy subject items and/or
- if he/she only made changes in the targets that were also made by at least one of the native speakers (see also §6.1).

Results of the native speakers of German
The results of all native speakers of German taken together are presented in table 6.2 below. The native speakers of German had 179 valid responses (15 participants x 12 items minus 1 missing). There were ten responses in which a change in the target co-occurred with some other change, so the total use of all coding categories (including missing) is 190.

Table 6.2:   Responses by native speakers of German

Item	Target	Predicted change	no change a	changes made in target c	d	e	g	total	other changes b	f	mis-sing
1	0	none	15	-	0	0	0	0	0	0	0
3	er	none	6	-	4	0	1	5	6	0	0
4	het	none	9	-	0	0	0	0	5	1	0
6	0	het	4	9	0	0	0	9	3	0	0
8	het	er	1	7	0	4	3	14	3	0	0
10	er	0	3	12	0	0	0	12	1	0	0
11	het	none	8	-	3	0	1	4	3	0	0
14	er	het	4	10	0	0	0	10	1	0	1
15	0	er	10	0	0	0	1	1	4	0	0
17	er	het	1	9	0	0	1	10	6	1	0
18	het	none/0	10	3	0	0	1	4	1	0	0
19	0	none	12	-	0	0	0	0	3	0	0
Total (190)			83	50	7	4	8	69	36	2	1
% of (179) responses			46%	28%	4%	2%	4%	39%	20%	1%	-

107

Comparing this table with table 6.1, one can see that the native speakers of German make many more changes than the native speakers of Dutch, both in the target and in other parts of the sentence. Less than half of the responses (46%) are a literal imitation of the stimulus sentence. There is a much smaller difference for the native speakers of German between changes made in the target (50) and other changes (38), than for the native speakers of Dutch. Of the 69 changes that they make in the target, however, a vast majority (50) is as predicted on the basis of the SPT, and only 19 changes are unpredicted. This indicates that the native speakers of German do not have great problems with the sentence imitation task. The items that cause most problems are items 3, 8 and 11. Items 8 and 11 are also among the most difficult ones for the native speakers of Dutch. Items 3 and 8 both belong to the DPnsi category, which was one of the most difficult categories for the native speakers of German in the SPT (see chapter 5). Item 15 also belongs to this category, but causes fewer problems, probably because it has the dummy subject that most native speakers of German prefer.

The individual results of the native speakers of German are presented in table 6.3. In this table (as well as in tables 6.5 and 6.7) the columns represent the participants, identified by their ID-code. Sentences that were literally repeated are not reported in these tables. The changes made by the non-native speakers that were not made by native speakers are marked with grey. Non-predicted changes in the target (d/e) that are different from the non-predicted changes in the target made by some of the native speakers are printed in italics (e.g. the response to item 11 by participant ID57). The participants that fall within the native speaker range for all items with a dummy subject as target are marked with ♪. ID111 in table 6.3 is the participant that replaced ID61.

Table 6.3:   Results of individual native speakers of German

Item	NSs	ID51	ID52	ID53	ID54	ID55	ID56	ID57	ID58	ID59	ID60	ID111	ID62	ID63	ID64	ID65
			♪	♪	♪	♪	♪			♪			?		♪	
1																
3	b	d				b	b	bd	g	b	b	d				bd
4	b	b		b		b			f			b				b
6	bceg	c	c		b	b	c	bc	c			c		c	c	c
8	bce	e	be	c	e	c	c	g	bc	c	g	c			g	be
10	c	c	c		c	c	c	c	c	c	c	bc	c	c		
11	bdf			b				*d*	g		b	d			b	d
14	bce	bc			c	c	c	c	c		c	c	h	c		c
15	bc		b	b					g					b		b
17	bc	g	bc	b		c	c	c	cf	b	c	c	c	b	b	c
18	bc	b				c	c		g						c	
19	b		b						b						b	

♪:   Participant who falls within the native speaker range
           Response not given by any of the native speakers

As can be seen in table 6.3, there are at least 7 native speakers of German who fall within the native speaker range (ID52, 53, 54, 55, 56, 59 and 64). Of these 7 participants, there are two who also fell within the native speaker range for all categories in the SPT (ID59 and 64). For participant ID62 (who fell within the NS range on the SPT), it cannot be determined whether she falls within the native speaker range or not, because her response to the dummy subject target in sentence 14 was not clear (h). If we compare this result to the results of the SPT, there are more native speakers of German who fall within the native speaker range on the SIT than on the SPT. This again seems to indicate that the sentence imitation task did not cause great problems for the German learners of Dutch.

Results of the native speakers of French
The native speakers of French, who performed very similarly (overall) to the native speakers of German on the SPT, have more problems than the native speakers of German on the SIT, as can be seen in table 6.4.

Table 6.4:   Responses by native speakers of French

Item	Target	Predicted change	no change a	changes in target					other changes	
				c	d	e	g	total	b	f
1	0	none	11	-	0	0	0	0	4	0
3	er	none	3	-	10	0	0	10	7	1
4	het	none	6	-	1	1	0	2	8	0
6	0	het	1	14	0	0	0	14	7	0
8	het	er	1	8	0	4	0	12	3	0
10	er	0	2	10	0	2	1	13	1	0
11	het	none	4	-	3	0	1	4	9	0
14	er	het	2	12	0	0	0	12	1	0
15	0	er	10	0	0	0	0	0	5	0
17	er	het	1	11	0	0	0	11	9	0
18	het	none/0	12	0	0	0	0	0	3	0
19	0	none	8	-	0	0	0	0	7	0
Total (204)			61	55	14	7	2	78	64	1
% of (180) responses			34%	31%	8%	4%	1%	43%	36%	1%

The native speakers of French had 180 valid responses (15 participants x 12 items). There were 24 responses in which a change in the target co-occurred with some other change, so the total use of the coding categories a-g is 204.

If one compares this table to tables 6.1 and 6.2, one can see that the native speakers of French have (relatively) fewer literal imitations, more unpredicted changes in the targets and more other changes than the native speakers of Dutch and German. However, the number of predicted changes in the target is about the same as for the native speakers of German (55 vs. 50). It should be noted that the difference in unpredicted changes between the native speakers of French and German is mainly caused by item 3, repeated in (14) below. This item belongs to the DPnsi category.

(14)   Tot nu toe komen er steeds meer mensen naar dit popconcert.
        until now come there still more people to this pop concert
        "So far more and more people have come to this pop concert."

In the SPT, the DPnsi category turned out to be relatively difficult for the native speakers of French. There are two more items of this category in the SIT (item 8 and item 15), but in item 15 the dummy subject in the stimulus sentence is *0* and in item 8 it is *het*, which can never occur in sentences of the DP-type. Still, there are eight native speakers of French who change *het* into *er* in item 8 (= c) and only four that change it into *0* (= e), as do two of the native speakers of Dutch. In item 3, on the other hand, there are 10 native speakers of French who change *er* into *0* (=d). Why this should be the case is unclear. Other items that cause problems for the native speakers of French

are item 10 (DPnst) and item 11 (PSdo). In item 10, repeated below as (15), three native speakers of French change the non-specific subject *zo veel mensen* into the generic subject *mensen* (applying *zo veel* to the object *bier* = e/g). Both generic and non-specific subjects in a transitive sentence, however, get a preference for *0* to *er*, although this preference is stronger for generic subjects (see also chapter 3).

(15)   Op Peters verjaardag drinken er nooit zo veel mensen bier.
         on Peter's birthday drink there never so many people beer
         "At Peter's birthday party there are usually not many people who
         drink beer."

In item 11, repeated here as (16), three native speakers of French change *het* into *ze*, as did one native speaker of Dutch (= d). Without the embedded clause this is a perfect sentence in Dutch, but not in combination with the embedded clause. In the response of one other participant, the entire main clause is missing (=g).

(16)   Door haar vriend wordt het bewonderd dat ze altijd vrolijk blijft.
         by her friend is it admired that she always cheerful remains
         "Her friend admires her for always remaining cheerful."

Why this sentence is more difficult for the native speakers of French than item 14 (which belongs to the same category) is unclear.

        If we now look at the results of the individual native speakers of French, presented in table 6.5 below, it turns out that there are three native speakers of French who fall within the native speaker range (ID68, 70 and 77). In this table, ID117 is the participant that replaced ID67.

Table 6.5:   Results of individual native speakers of French

Item	NSs	ID66	ID117	ID68	ID69	ID70	ID71	ID72	ID73	ID74	ID75	ID76	ID77	ID78	ID79	ID80
					♪		♪							♪		
1						b					b	b				b
3	b	bd	bd		d			d	d	f	d	bd	b	bd	bd	bd
4	b	b			e	b	bd		b	b	b		b			b
6	bceg	c	bc	bc	bc		c	bc	c	c	c	c	c	bc	bc	bc
8	bce	b	c	c	e		c	b	c	c	c	e	c	be	c	e
10	c	c	c	c	c		c	e		bc	e	c	c	c	c	g
11	bdf	b			bd		g	b	b		b	d	b	b	bd	b
14	bce	c		c	c	c	c	b	c	c	c	c	c			
15	bc		b		b		b		b							b
17	bc	c	b		bc	c	bc	c	c	b	bc	c	bc	bc	bc	b
18	bc										b		b		b	
19	b	b		b							b	b		b		b

♪:   Participant who falls within the native speaker range
            Response not given by any of the native speakers

This number of participants within the NS range is similar to the number of native speakers of French who fell within the native speaker range on the SPT (4). However, there is only one native speaker of French who falls within the native speaker range on both tasks. Again, it should be noted, that most deviations from the target occur in item 3. Of the 12 native speakers of French who do not fall within the native speaker range on the SIT, there are seven who only use a different dummy subject than the native speakers for this item (ID66, 117, 73, 74, 76, 78 and 79).

Results of the native speakers of Turkish
The native speakers of Turkish, who had more problems than the other groups on the SPT, also have more problems with the SIT, as is shown in tables 6.6 and 6.7. In these tables, the results of participants ID81 and ID85 have been left out, because these participants did not meet our selection criteria (see chapter 4). The native speakers of Turkish had 155 valid responses (13 participants x 12 items – 1 missing). There were 37 responses in which a change in the target co-occurred with some other change, so the total at the bottom of the table adds up to 193 (including the missing value).

Table 6.6:   Responses by native speakers of Turkish

Item	Target	Predicted change	no change	changes made in target						other changes		mis-sing
			a	c	d	e	g	total	b	f	
1	0	none	5	-	0	0	0	0	8	0	0
3	er	none	1	-	11	0	0	11	9	0	0
4	het	none	2	-	0	2	2	4	9	0	0
6	0	het	3	7	0	0	1	8	6	0	0
8	het	er	0	6	0	3	2	11	1	0	1
10	er	0	1	8	0	4	0	12	7	0	0
11	het	none	0	-	10	0	2	12	6	0	0
14	er	het	2	10	0	1	0	11	2	0	0
15	0	er	6	3	0	0	0	3	4	0	0
17	er	het	1	6	0	1	2	9	8	0	0
18	het	none/0	2	4	0	2	0	6	8	1	0
19	0	none	3	-	0	0	0	0	10	0	0
Total (193)			26	44	21	13	9	87	78	1	1
% of (155) responses			17%	28%	14%	8%	6%	56%	50%	1%	-

As can be seen in this table, the native speakers of Turkish have (relatively) fewer literal imitations, more unpredicted changes in the target and more changes in other parts of the sentences than the three other groups. However, still 17% of all sentences are imitated literally and about half of the changes in the targets (87) are as predicted (44).

Of the 43 unpredicted changes in the target, most (23) occur in items 3 and 11. Both these items also caused problems for the native speakers of French. In fact, the five items that caused most problems for both groups are the same, although the extent to which some of them cause problems is different. Moreover, both groups show a strong asymmetry between their behaviour on item 11 versus item 14 (see above for discussion). For item 11, the target *het* was changed into *ze/zij* by five native speakers of Turkish (and by three native speakers of French and one native speaker of Dutch) and into *0* by four native speakers of Turkish. One participant in this group changed *bewonderd* into *beroemder* (more famous), which also occurs with *het* in Dutch. One participant changed *wordt het bewonderd* into *is ze bewonderd* (is she admired). For this response, the same applies as to the change of *het* into *ze* (see above). Finally, one participant in this group turned the passive sentence into an active sentence without a dummy object (*Haar vriend bewondert dat zij altijd vrolijk blijft*). This is a very good example of how participants sometimes adapt sentences in such a way that they can process them properly according to their own grammar.

The individual results are shown in table 6.7. This table shows that there is one participant in this group who falls within the native speaker

range (ID83). This is the same participant that fell within the NS range for the SPT.

Table 6. 7:   Results on the SIT for individual native speakers of Turkish

Item	NSs	ID82	ID83	ID84	ID86	ID87	ID88	ID89	ID90	ID91	ID92	ID93	ID94	ID95
		♪												
1		b			b				b	b	b	b	b	b
3	b	bd	b	bd	bd	bd		d	d	bd	bd	bd	bd	bd
4	b		b	b	g	b		be	b	b	be	g	b	b
6	bceg	c	b	c		b	c	bc		bc	bc	bc		g
8	bce	c	c	e	c	e	e	c	e	c	g	g	c	h
10	c	bc		c	bc	bc	c	be	e	be	bc	bc	c	e
11	bdf	d	bd	b*d*	g	d	g	d	b*d*	b*d*	bd	b	d	bd
14	bce	bc	c	c	c	e	c			c	c	c	bc	c
15	bc		c			b			c		b	b	c	b
17	bc	bc		b	bc	be	bc	c	c	g	bc	b	b	g
18	bc	b	b			c	f	bc	be	be	b	c	b	bc
19	b	b	b	b		b	b	b	b		b	b		b

♪:   Participant who falls within the native speaker range
Response not given by any of the native speakers

Comparison of groups
If we compare the percentages of the totals per group, as in table 6.8, we see a strong influence of the typological distance to Dutch. As the typological distance increases:
-   the percentage of literal imitations decreases
-   the percentage of unpredicted changes in the target (total – c) increases
-   the percentage of changes in other parts of the sentence (b+f) increases.

However, the percentage of predicted changes in the target, which is 18% for the native speaker control group, is around 30% for all non-native speaker groups. It should be noted that the totals in this table deviate from the totals in tables 6.1, 6.2 and 6.6, because the missing values have been left out here.

Table 6.8:   Total responses of all groups compared

L1		no change	changes made in target					other changes	
		a	c	d	e	g	total	b	f
D	Total (533)	390	94	2	8	1	105	37	1
	% of (525) responses	74%	18%	0%	2%	0%	20%	7%	0%
G	Total (189)	83	50	7	4	8	69	36	2
	% of (179) responses	46%	28%	4%	2%	4%	39%	20%	1%
F	Total (204)	61	55	14	7	2	78	64	1
	% of (180) responses	34%	31%	8%	3%	2%	43%	36%	1%
T	Total (192)	26	44	21	13	9	87	78	1
	% of (155) responses	17%	28%	14%	8%	6%	56%	50%	1%

A comparison of tables 6.2, 6.4 and 6.6 also shows that, in all non-native speaker groups, there are participants that have problems with items 3, 8 and 11, although not to the same extent. For the L1 Turkish and French group, the five items that present most problems are the same. This was predicted on the basis of relevant aspects of French and Turkish (see §3.4). We cannot explain, however, why all groups have (some) problems with item 11, but not with item 14. This is especially strange considering the fact that, for the native speaker control group, item 14 caused more problems than item 11.

If we take the individual results of the three non-native speaker groups together, there are 11 non-native speakers who fall within the native speaker range on the SIT (ID52, 53, 54, 55, 56, 59, 64, 68, 70, 77 and 83). It should be noted that, if we had considered all participants with an unpredicted change in a dummy subject target to fall outside the native speaker range, there would still be six second language learners who would fall within the native speaker range (ID53, 55, 59, 68, 70 and 77). However, there would also have been nine native speakers of Dutch who would not have fallen within the native speaker range.

Chapter

7

Discussion and

conclusions

In this chapter, we will discuss the results of our experiments, presented in chapters 5 and 6, in relation to the research questions presented in chapter 2 and repeated in 1-3 below.

1)   Are there any late second language learners who fall within the native speaker range in their command of grammatical constructions that are known to be very difficult for second language learners and which can only be acquired on the basis of the input? (§ 7.2)

2)   How is the level attained in L2 grammar after the age of twelve related to the typological distance between the L1 and the L2? (§ 7.3)

3)   What are the input and background characteristics of late learners who perform within the native speaker range (if they exist)? (§ 7.4)

We will start with a summary of the main findings from the sentence preference task and the sentence imitation task, and a comparison between these two tasks. Subsequently, we will address our research questions and discuss the implications of our results for the Critical Period Hypothesis for syntax in SLA, and for late learners who want or need to reach a very high level of proficiency in a second language.

7.1   Summary of the main findings

Results of the sentence preference task
For the sentence preference task (SPT) we analysed the scores of all participants on six categories of dummy subject constructions (see chapter 3 for an overview and examples), and the total score for 79 items (see chapter

4 for a description of the analyses of the data).We analysed these scores both at the group level and at the individual level.

At the group level, t-tests and an analysis of variance (ANOVA) revealed significant differences between the native speaker control group and all non-native speakers together for the total and for all categories, except DPnst (transitive predicates with a non-specific noun phrase/DP as logical subject; see table 5.2). To compare the four L1 groups to each other, we used Games-Howell post hoc tests (see table 5.4). These analyses yielded the following significant differences between groups:

- L1 Dutch vs. L1 German: only DPg (general category of sentences with a noun phrase/DP as logical subject) and total
- L1 Dutch vs. L1 French: none
- L1 Dutch vs. L1 Turkish: total + all categories, except DPg
- L1 German vs. L1 French: only DPg
- L1 German vs. L1 Turkish: only PSg (general category of passive sentences with a sentential logical subject) and total
- L1 French vs. L1 Turkish: only total

We found interesting differences between the three non-native speaker groups with respect to the influence of the L1. On the basis of relevant properties of the first languages involved in our study, we predicted the L1 German group to have most problems with the DP-type, and the L1 French and Turkish groups to have most problems with the PS-type (for more detailed predictions see §3.4). These predictions were borne out (see §5.2 and 5.3). Some predictions could not be tested for the native speakers of German and French because of a ceiling effect. In other respects our predictions on the basis of properties of French and Turkish were not borne out:

- The L1 Turkish group performed better for the general category of the DP-type (DPg) than for the sentences with non-specific logical (DP) subjects and a transitive predicate (DPnst), whereas we predicted that their performance on these categories would be similar.
- The L1 French and L1 Turkish groups performed better for the general category of the active sentences with a sentential logical subject (ASg) than for sentences with non-specific logical (DP) subjects and an intransitive predicate (DPnsi), whereas we predicted the opposite.
- The L1 French and L1 Turkish groups did not perform worse for passive sentences with a sentential logical subject that have a dummy object in their active equivalent (PSdo) than for other passive sentences with a logical subject (PSg), contrary to what we predicted.

In all these cases, the category for which the participants performed better than predicted (DPg, ASg and PSdo) involved a stronger preference pattern than that for the categories on which the participants performed worse than predicted (DPnst, DPnsi and PSg). This might be due to the fact that the input is probably clearer and less ambiguous for constructions for which native speakers have a strong preference pattern than for constructions for which native speakers have a weaker pattern. A combination of influence of the strength of the judgement pattern of the native speakers with the influence of the L1 could explain most of our results. However, this cannot explain why the native speakers of German performed better on the DPnst category than on the DPg category.

At the individual level, we combined the results for all categories to see whether there were any non-native speakers who fell within the native speaker range (NS-range) for all dummy subject categories in our analyses. Because we used a very strict criterion for falling within the NS-range, there were eight (out of 44) native speakers of Dutch that did not fall within the native speaker range according to this analysis. Although our criterion was very strict and probably led to an underrepresentation of the actual number of non-native speakers who performed within the native speaker range, there were still eight non-native speakers of Dutch who had the native speaker pattern for each category: three native speakers of German, four native speakers of French and one native speaker of Turkish.

For these eight second language learners who fell within the native speaker range we did some additional analyses to see whether they made the same distinctions between categories within each type of dummy subject construction as the majority of the native speaker control group. Except for one of the French participants, they all did. However, the Turkish participant within the NS-range made a greater distinction between the two categories of the AS-type than most of the native speakers, and there were eight native speakers of Dutch (five of whom fell within the NS-range for all categories) who had at least one deviant z-score for these analyses. For these reasons, we did not consider the results of these additional analyses when looking at background characteristics (see §7.4).

We computed T-tests for the eight non-native speakers mentioned above versus the other 35 non-native speakers with several background characteristics obtained through a questionnaire. These tests revealed significant differences for:
- speaking and writing skills in Dutch (self-report on a five-point scale)
- listening and reading skills in the participants' best other language (self-report on a five-point scale)
- agreement with the statement: "I like to learn languages"
- level of education

119

In all these cases the eight participants within the NS range for all categories had higher scores (i.e. better skills, higher agreement and a higher level of education). Other skills in Dutch and the participants' best other language (including grammar) and all skills in the native language (all self-report on a five-point scale) did not yield any significant differences between both groups. The same is true for attending Dutch lessons, age of arrival and amount of input / length of residence.

Results of the sentence imitation task
For the sentence imitation task (SIT) we did not compute any statistical analyses (see § 4.4). Instead, we compared the changes made in the sentences with a dummy subject by the non-native speakers with those made by the native speakers of Dutch.

At the group level, we looked at those items which caused most problems for each group, and we compared the total percentage for each type of change (or no change) to that of the other groups. This analysis revealed that, as the typological distance to the target language increases, the percentage of literal imitations decreases and the percentage of unpredicted changes in the target as well as the percentage of changes in other parts of the sentence increases. This analysis also revealed that most changes, especially in the L1 French and L1 Turkish groups, were made in one or two items.

At the individual level, we compared all changes for each non-native speaker to those made by the native speaker control group. If a second language learner either made no change in the targets of the dummy subject sentences or only changes that were also made by at least one of the native speakers, we considered this participant to fall within the native speaker range. According to this analysis, there were (at least) seven native speakers of German, three native speakers of French and one native speaker of Turkish who fell within the native speaker range.

The sentence preference task versus the sentence imitation task
The results of the sentence preference task and the sentence imitation task cannot be compared at the level of categories, because the SIT contains too few items to generalise across categories and because not all categories are represented in the SIT. For various reasons, it is also very difficult to compare the number of participants that fall within the native speaker range on the SPT and the SIT.

First, the number of items in the SIT was rather low, whereas for the SPT it was rather high. Second, although it is evident that there are non-native speakers who behave like the native speakers of Dutch on both tasks,

the exact number of non-native speakers that behave in a native-like manner is impossible to determine, because of the variation within the native speaker control group. Due to the difference in nature between the two tasks, the way we dealt with this variation in establishing a norm for native-like behaviour was very different for both tasks. For the SIT we included all native speaker variation in establishing native-likeness (see § 6.1), whereas for the SPT we based the norm on the behaviour of 90% of the native speaker control group for each category (see §5.1). For the SPT this probably led to an underrepresentation of the number of non-native speakers who fell within the native speaker range for all categories according to our criteria, given the fact that there were eight native speakers who did not meet our criteria for falling within the NS range. This means that one should be cautious in comparing the number of non-native speakers who fell within the native speaker range on the two tasks.

One can, however, compare the relative performance of the different groups in the two tasks. For the SPT, the native speakers of German performed at the same level as or slightly worse than the native speakers of French. At the group level, this can be seen both in the mean total score for both groups (71.9 for the native speakers of German and 71.8 for the native speakers of French, see table 5.3), and in the number of significant differences from the native speakers of Dutch in the post hoc tests (two out of six for the native speakers of German and none for the native speakers of French, see table 5.4). At the individual level, there were three native speakers of German who fell within the native speaker range for all categories according to our criteria and four native speakers of French. For the SIT, however, the native speakers of German performed clearly better than the native speakers of French. At the individual level, there were seven native speakers of German who fell within the native speaker range according to our criteria and three native speakers of French (see tables 6.3 and 6.5). At the group level, the L1 German group had more literal imitations than the L1 French group (83 vs. 61) and fewer changes not related to the target (36 vs. 64). The difference in changes in the target, however, was rather small (69 vs. 78). These results suggest that the influence of the typological distance to Dutch is greater for the SIT than for the SPT, at least for languages that are relatively closely related to the target language.

7.2 Native-like attainment

The results discussed in chapters 5 and 6 and summarised in the previous paragraph clearly show that there are late second language learners who fall

within the native speaker range in their command of dummy subject constructions in Dutch, even if we use very strict criteria for falling within the native speaker range. Since dummy subject constructions are known to be very difficult for second language learners and can only be acquired on the basis of the input, the answer to our first research question is affirmative: there are late second language learners who fall within the native speaker range in their command of grammatical constructions that are known to be very difficult for second language learners and that can only be acquired on the basis of the input.

According to the Critical Period Hypothesis (CPH), discussed in chapter 2, levels attained after the presumed terminus of the critical period (around age 12 for L2 syntax) should always be low and significantly lower than levels attained by learners who started to acquire a second language before the alleged terminus (see e.g. Bialystok & Miller, 1999). Native-like attainment by late learners for constructions that are different in the L1 and for which no explicit knowledge is available, therefore, constitutes counterevidence against the CPH for syntax in SLA. In our study we found such evidence.

There has been some debate in the literature with respect to the age at which the critical period for syntax is supposed to end, ranging from before puberty to the end of puberty (compare for example Lenneberg, 1967; Patkowski, 1980; Johnson & Newport, 1989; Long, 1990; Bialystok & Hakuta, 1994; Bialystok & Miller, 1999). In our study, the ages of arrival in the Netherlands of participants with native-like attainment of dummy subject constructions range from 12 to 35 with a mean AoA of 23 (see table 5.5). In other words, there were participants who started to acquire their L2 far beyond puberty and still managed to obtain a native level, thus providing strong evidence against the CPH for syntax in SLA.

Although our study was not designed to test the shape of the age function, it is interesting to note that we did not find a significant difference in AoA between the participants who performed best (within the NS range) and all other participants. This seems compatible with the CPH prediction of a flattening after the terminus, when age of onset is no longer correlated with the level of proficiency attained (see § 2.1 and 2.3). However, this flattening should represent a rather low level of proficiency. The very high (and in some cases native-like) level of proficiency of our participants is, therefore, incompatible with certain predictions derived from the CPH.

This result also does not seem to be in line with results from studies by Flege et al. (1999) and Birdsong and Molis (2001), who found a significant correlation between AoA and proficiency both for early and for late L2 learners. It should be noted, though, that in these studies learners were not selected on proficiency, as was the case in our study.

Our results support findings from other studies that have reported native-like attainment of syntax in late L2 learners, e.g. Schachter (1990), Birdsong (1992), White and Juffs (1998), McDonald (2000), Birdsong and Molis (2001), Hyltenstam and Abrahamsson (2003; results for the grammaticality judgement test), Montrul and Slabakova (2003), Sabourin (2003) and Birdsong (2005). In many of these studies there were methodological problems with respect to the constructions used and/or the languages involved. In this dissertation we have tried to show that native-like attainment can also be found for syntactic constructions that are known to be very difficult for learners, for which no explicit knowledge is available, and which are rather different in the L1 of (at least some of) the learners.

There are also studies which do not report native-like attainment for certain syntactic constructions for highly proficient L2 learners, e.g. Coppieters (1987), Hyltenstam (1992), Ioup et al. (1994)[22]. Moreover, there are some studies that suggest that proficient late learners have more processing problems in their L2 than monolinguals, e.g. Hyltenstam and Abrahamsson (2003) and Sabourin (2003). In fact, our study also shows that some learners who clearly have a native-like command of dummy subject constructions in Dutch sometimes make errors in the imitation task, which makes heavy demands on participants' working memory / processing capacities. This does not mean, however, that there is also a difference in competence.

Our results with respect to the acquisition of dummy subject constructions, which are known to be very difficult for learners and for which no explicit knowledge is available to L2 learners, seem irreconcilable with a critical period for L2 syntax. In order to prove that there is a critical period for syntax in SLA, an account would be necessary of how late learners with a strongly reduced sensitivity to the input due to having passed the critical period could possibly acquire such grammatical constructions. So far, such an account has not been put forward.

Although we question the existence of a critical period for syntax in SLA, we acknowledge that biological and other factors usually confounded with age play an important role in second language acquisition. An important biological factor usually confounded with age is entrenchment of the L1. Because of this entrenchment, adult learners may be less attentive to certain aspects of the L2 input. An important non-biological factor confounded with age is the nature of the input. Differences in prosodic features, stress, utterance length, etc. of the input directed to young children versus the input directed to adults may result in differences in attention to

---

[22]   Although Ioup et al. (1994) stress the high proficiency of the participants in their study, the results of both participants on the grammaticality judgement task show clear deviations from the native speaker norm.

certain characteristics of the input. However, we do believe that the empirical evidence available shows that the general decline in the ability to acquire L2 syntax is not absolute and deterministic, as is the case for orientation specificity in the visual cortex in cats or imprinting in ducklings, for example (see e.g. Baxter, 1966; Ratner & Hoffman, 1974). It has been shown by Pallier et al. (2003), for example, that, when L2 learners are completely deprived of input in the L1 (in this case Korean children adopted by French speaking parents), the levels attained in the L2 are much more native-like than is normally the case in L2 acquisition.

It could be that other biologically determined factors related to age constrain the acquisition of L2 syntax, for example maturation of working memory (see §2.7). It seems plausible that elements such as dummy subjects, which are phonologically non-salient and do not really contribute to the meaning of the sentence, are overlooked by late learners, who can process large units at once because of their larger working memory capacity. Although this is a biologically determined age factor, it is not specific to language and not absolute. Details in the form that in general are not attended to by late learners might be acquired if special attention is paid to them. It seems plausible that second language learners with a linguistic background or high language aptitude are in fact capable of paying special attention to elements like dummy subjects, which are ignored by most L2 learners. The results with respect to the background characteristics of the learners who score within the NS range seem compatible with this hypothesis. Unfortunately, however, our study cannot test this prediction, because we did not investigate the relation between language aptitude and sensitivity to the input. Education may also play a role. Informal contacts with the participants in this study led to the impression that there were differences between participants with respect to their attitude towards prescriptive language standards in general, due to differences in education. It would be interesting to investigate in future studies whether differences in language aptitude and education can lead to differences in attention to the exact form of the input.

7.3   The role of the first language

In our study, we looked at three different L1 backgrounds with different typological relations to Dutch (German, French and Turkish), to see how the level of attainment in L2 grammar after the age of twelve is related to the typological distance between the L1 and the L2 (research question 2). The results from our study clearly show that it is possible to reach a native level in L2 syntax for learners with language backgrounds with different degrees

of typological distance to the L2, ranging from a very close typological relationship (German and Dutch) to a very distant typological relationship (Turkish and Dutch). Since Turkish does not have any dummy arguments, our results also show that it is possible for late learners to acquire a grammatical construction not present in their native language.

In many studies with late learners, especially less recent ones, native-like attainment for L2 syntax has not been found (e.g. Coppieters, 1987; Johnson & Newport, 1989; Johnson & Newport, 1991, Sorace, 1993; Ioup et al., 1994). In those studies in which native-like attainment was reported for L2 syntax, it was mainly found for learners with a typologically rather closely related L1 (see for example Schachter, 1990; Birdsong, 1992; McDonald, 2000; Birdsong & Molis, 2001; Sabourin, 2003).

However, in some very recent studies native-like attainment on a morphosyntactic test has been found for some late learners with a typologically distant L1, namely for one Estonian learner of Swedish (Hyltenstam & Abrahamsson, 2003) and one Vietnamese learner of English (McDonald, 2004 and personal communication, 23 March 2004).

These results, as well as our results, show that it is possible to reach a native level in L2 syntax for late learners with an L1 which is typologically very distant from the L2. This does not mean, however, that there is no L1 influence at very high proficiency levels. First, the number of participants within the native speaker range was much lower for the L1 Turkish group than for the L1 German and L1 French groups. For the imitation task, there was also a clear difference between the L1 German and the L1 French group: the L1 German group performed better than the L1 French group, both at the group level and in the number of individual participants that fell within the native speaker range. Moreover, there was a clear L1 effect for those learners who did not fall within the native speaker range on the SPT (see §7.1). At the same time, some of the predictions on the basis of a contrastive analysis of the source languages involved in our study were not borne out. This might in part be accounted for by a ceiling effect and by the strength of the preference pattern of the native speakers (see §7.1).

7.4   Background characteristics

Now that it is clear that native-like attainment is possible for late learners, we can address our third research question, repeated below.

-    What are the input and background characteristics of late learners who perform within the native speaker range (if they exist)?

With respect to the input of the L2 participants in our study, we saw that there was no significant difference between the participants who fell within the native speaker range on all categories in the SPT and the other participants, neither for length of residence in the Netherlands, nor for amount of input (length of residence multiplied by self-reported percentage of usage of Dutch). Moreover, the minimum amount of input in the group which performed within the NS range was not more than about 30% (self-reported) usage of Dutch for nine years. Although a certain amount of input and L2 use is obviously necessary to reach a high level of proficiency, this result suggests that input alone is not sufficient. The same amount of input and L2 use may lead to very different levels of proficiency under the influence of other factors.

Of the other factors on which we collected information through our questionnaire, the only factors that yielded a significant difference between the participants within the NS range versus the other L2 participants were level of education, speaking and writing skills in Dutch, listening and reading skills in the participants' best other language (all self-report) and agreement with the statement "I like to learn languages". The fact that level of education yielded significant differences could in principle have been caused by the fact that the task of judging minimal sentence pairs on a scale may have been very difficult for some participants with a lower level of education who are not used to taking such tests. However, the fact that there were only two L2 participants who had a level below HAVO (secondary education preparing for higher professional education), makes this explanation unlikely.

Factors such as attending Dutch classes, skills in the native language and statements such as "I find it important to use Dutch without errors" yielded no significant differences. In some cases this is probably due to the fact that there was little variation between the participants on the variables concerned. Almost all participants, for example, found it important or very important to use Dutch without errors.

On the one hand, the role of factors like input, attending Dutch classes and finding it important to use the L2 without errors seemed rather limited in our study. At the same time, some of the skills in the participants' best other language yielded significant differences between the L2 participants within the NS range and all other L2 participants. These facts in combination suggest that something like language aptitude may play an important role. It should be noted in this connection that our results cannot be explained by "explicit reflection on rules" (DeKeyser, 2000: 508) by adult acquirers with high verbal aptitude, as proposed by DeKeyser for his own data, because our participants did not have explicit knowledge of the rules they applied. In fact, DeKeyser explicitly states that:

> The only way that an adult learner can achieve grammatical competence similar to that of a native is by using analytical, problem-solving abilities, because the implicit learning mechanisms of the child are no longer available or accessible. Therefore, only learners with above-average analytical abilities can eventually reach a near-native level, at least on the rather easy kind of test used here, which does not include rare constructions or convoluted sentences and does not require participants to perform under time pressure. On a more subtle test of the kind used by Coppieters (1987), even they, of course, may clearly fall short of native standards. (DeKeyser, 2000: 515)

Our data, therefore, clearly contradict DeKeyser's statement that:

> Early age confers an absolute, not a statistical, advantage – that is, there may very well be no exceptions to the age effect. Somewehere between the age of 6-7 and 16-17, everybody loses the mental equipment required for the implicit induction of the abstract patterns underlying human language, and the critical period really deserves its name. (DeKeyser, 2000: 518)

Given the limitations of our questionnaire, our selection criteria and procedure, and the number of participants in our study (resulting in a bias for highly educated and highly proficient second language learners who often have a more or less linguistic background), more research is necessary to investigate the role of language aptitude in reaching a high level of proficiency in L2 syntax.

In spite of the limitations of our background questionnaire, it is interesting to see what the typical participant that fell within the native speaker range in our study looks like. It turned out that a typical participant in this group is a German or French woman in her mid forties, who arrived in the Netherlands in her early twenties. She studied in the Netherlands, now works as a translator and is rather proficient in at least one other language. She speaks her native language with relatives abroad, when visiting her country of origin, and part of the time at work, and Dutch with everyone else. She likes to speak Dutch and to learn languages, and she finds it important to use Dutch without making errors. Her proficiency in her native language (self-report) does not seem to have suffered from her high proficiency in Dutch and her extensive use of it.

What do these results imply for adolescent or adult immigrants for whom it may be important to reach a native-like level in L2 syntax (e.g. teachers, speech therapists, journalists and translators)?

On the one hand, our results (as well as results from other studies) strongly suggest that only a small percentage of second language learners

reach a native level of proficiency in L2 grammar. Although we cannot say how many people exactly reach a native level, the fact alone that it was hard to find participants that met our selection criteria suggests that native-like attainment is very rare. Moreover, factors that learners have most influence on, such as amount of input and attending Dutch lessons, did not seem to play a major role in our study, whereas factors that seem to be more difficult to change, such as language aptitude / sensitivity to small elements in the input, did seem to play an important role in reaching native-like proficiency. It would be interesting to investigate in future research whether or not such sensitivity to elements of the input can be enhanced by training.

The outcomes described in this paragraph warn against unrealistic expectations for the majority of late second language learners and are important to keep in mind when comparing second language learners with different levels of proficiency.

Rare though native-like attainment by late learners may be, the results from this study show that it is possible for some late learners to reach a native level in L2 syntax, even for learners with a typologically distant mother tongue. This is of course very good news for late second language learners. It may come as a surprise for those participants who told me before I started testing them that they could tell me in advance that I would not find native-like attainment, because "that is impossible for late learners." Apparently their command of Dutch syntax is better than they realise.

In sum, this study has provided important evidence against the Critical Period Hypothesis for syntax in second language acquisition. We showed that native-like attainment of dummy subject constructions in Dutch is possible for late L2 learners, even for learners with a typologically very distant L1.

The time seems ripe now to move on in a different direction. Rather than providing more evidence for the (non-)existence of a critical period, one of the main challenges for future research in this area is to find out what factor(s) contribute most to language learning success and what are the best ways to influence learners' sensitivity to non-salient aspects of the input.

References

Abutalebi, J., Cappa, S., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition, 4*, 179-190.

Ardal, S., Donald, M., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language, 39,* 187-205.

Baxter, B. (1966). Effect of visual deprivation during postnatal maturation of the electroencephalogram of the cat. *Experimental Neurology, 14*, 224-237.

Bennis, H. (1987). *Gaps and dummies.* Dordrecht: Foris.

Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research, 13*, 116-137.

Bialystok, E. (2002). On the reliability of robustness. A reply to DeKeyser. *Studies in Second Language Acquisition, 24,* 481-488.

Bialystok, E., & Hakuta, K. (1994). *In other words: The science and psychology of second language acquisition.* New York: Basic Books.

Bialystok, E., & Miller, B. (1999). The problem of age in second language acquisition: Influences from language, structure and task. *Bilingualism: Language and Cognition, 2*, 127-145.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language, 68*, 706-755.

Birdsong, D. (1997, November). Intransitivity and SE in French: Aspects of late L2 learnability. Paper presented at the 22nd Boston University Conference on Language Development, Boston, MA.

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence.* Berlin: Springer.

Birdsong, D. (2003). Authenticité de prononciation en français L2 chez des apprenants tardifs anglophones: Analyses segmentales et globales. *Acquisition et Interaction en Langue Étrangère, 18*, 17-36.

Birdsong, D. (in press). Interpreting age effects in second language acquisition. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches.* New York: Oxford University Press.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language, 44,* 235-249.

Bley-Vroman, R., & Chaudron, C. (1994). Elicited imitation as a measure of second-language competence. In E. Tarone, S. Gass & A. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 245-261). Hillsdale, NJ: Lawrence Erlbaum.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 133-159). Mahwah, NJ: Lawrence Erlbaum.

Bongaerts, T. (2003). Effets de l'âge sur l'acquisition de la prononciation d'une seconde langue. *Acquisition et Interaction en Langue Étrangère, 18*: 79-98.

Bongaerts, T., Mennen, S., & van der Slik, F. (2000). Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica, 54*: 298-308.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19*: 447-465.

Bornstein, M. (1989). Sensitive periods in development: Structural characteristics and causal interpretations. *Psychological Bulletin, 105*, 179-197.

Carlson, G. (1977). *Reference to kinds in English*. Unpublished doctoral dissertation. University of Massachusetts at Amherst, Amherst.

Chee, M., Tan, E., & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *Journal of Neuroscience, 19,* 3050-3056.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Colombo, J. (1982). The critical period concept: Research, methodology, and theoretical issues. *Psychological Bulletin, 91,* 260-275.

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

Curme, G. (1952). *A grammar of the German language*. New York: Ungar.

Dehaene, S., Dupoux, D., Mehler, J., Cohen, L., Paulesu, E., Perani, D., van de Moortele, P., Lehéricy, S., & LeBihan, D. (1997). Anatomical variability in the cortical representation of first and second languages. *NeuroReport, 8*, 3809-3815.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22,* 499-534.

Drosdowski, G., & Augst, G. (1984). *Duden: Grammatik der deutschen Gegenwartssprache* (Der große Duden in 10 Bänden: 4). Mannheim: Bibliographisches Institut.

Elman, J. (1993), Learning and development in neural networks: The importance of starting small, *Cognition, 48*, 71-99.

Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry, 22,* 1-25.

Engel, U. (1988). *Deutsche Grammatik.* Heidelberg: Julius Gross.

Felser, C., & Rupp, L. (2001). Expletives as arguments: Germanic existential sentences revisited. *Linguistische Berichte, 187,* 289-324.

Flege, J., Frieda, A., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*, 169-186.

Flege, J., Munro, M., & MacKay, I. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3134.

Flege, J., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41,* 78-104.

Grondelaers, S., Speelman, D., & Carbonez, A. (2001). Regionale variatie in de postverbale distributie van presentatief er. *Neerlandistiek.nl, 1-4*, 1-41. <http://www.neerlandistiek.nl>.

Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J., & van den Toorn, M. (1997). *Algemene Nederlandse spraakkunst.* Groningen: Martinus Nijhoff.

Hahne, A., & Friederici, A. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4,* 123-141.

Helbig, G., & Buscha, J. (1991). *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht* (4th ed.). Berlin: Langenscheidt.

Hoekstra, T., & Mulder, R. (1990). Unergatives as copular verbs; Locational and existential predication. *The Linguistic Review, 7,* 1-79.

Hoffman, H. (1974). Fear-mediated processes in the context of imprinting. In M. Lewis & L. Rosenblum (Eds.), *The origins of fear* (pp. 25-48). New York: John Wiley.

Hyltenstam, K. (1992). Non-native features of near-native speakers. On the ultimate attainment of childhood L2 learners. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 351-368). Amsterdam: Elsevier Science.

Hyltenstam, K., & Abrahamsson, N. (2003), Age of onset and ultimate attainment in near-native speakers of Swedish. In K. Fraurud & K. Hyltenstam (Eds.), *Multilingualism in global and local perspectives. Papers from the 8th Nordic Conference on Bilingualism*, November 1-3, 2001, Stockholm Rinkeby (pp. 319-340). Stockholm: Centre for Research on Bilingualism, Stockholm University, and Rinkeby Institute of Multilingual Research.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition, 16*, 73-98.

REFERENCES

Jia, G. (1998). Beyond brain maturation: The critical period hypothesis in second language acquisition revisited. Unpublished doctoral dissertation. New York University, New York.

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics, 24*, 131-162.

Johnson, J. (1988). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign, Urbana-Champaign.

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,* 60-99.

Johnson, J., & Newport, E. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition, 39,* 215-258.

Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*: 122-149.

Kersten, A., & Earles, J. (2001). Less really is more for adults learning a miniature artificial language. *Journal of Memory and Language, 44*, 250-273.

Kim, K., Relkin, N., Lee, K., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature, 388*, 171-174.

Krashen, S., Long, M., & Scarcella, R. (1982). Age, rate and eventual attainment in second language acquisition. In S. Krashen, R. Scarcella & M. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 161-174). Rowley, MA: Newbury House.

Lasnik, H. (1995). Case and expletives revisited: On greed and other human failings. *Linguistic Inquiry, 26*, 615-633.

Lenneberg, E. (1967). *Biological foundations of language.* New York: John Wiley.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*: 251-285.

Marler, P. (1970). A comparative approach to vocal learning: Song development in white-crowned sparrows. *Journal of Comparative Physiological Psychology, 71*, 1-25.

McDonald, J. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21*, 395-423.

McDonald, J. (2004). Grammaticality judgment in second language learners: Effects of age of acquisition, modality, construction type, and first language. Unpublished manuscript, Department of Psychology, Louisiana State University, Baton Rouge.

Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite/imperfect Contrast in Spanish. *Studies in Second Language Acquisition, 25,* 351-398.

Moro, A. (1997). *The raising of predicates. Predicative noun phrases and the theory of clause structure.* Cambridge: Cambridge University Press.

Moyer, A. (1999). Ultimate attainment in L2 phonology. The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition, 21*, 81-108.

Müller, N. (1998). UG access without parameter setting: A longitudinal study of (L1 Italian) German as a second language. In M. Beck (Ed.), *Morphology and its interfaces in second language knowledge* (pp. 115-163). Amsterdam: John Benjamins.

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.

Newport, E. (1991). Contrasting conceptions of the critical period for language. In: S. Carey & R. Gelman (Eds.), *Epigenesis of mind: Essays of biology and cognition* (pp. 111-130), Hillsdale, NJ: Lawrence Erlbaum.

Newport, E., Bavelier, D., & Neville, H. (2001). Critical thinking about critical periods. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 481-502). Cambridge, MA: MIT Press.

Pallier, C., Dehaene, S., Poline, J.-B., LeBihan, D., Argentini, A.-M.., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: can a second language replace the first? *Cerebral Cortex, 13,* 155-161.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning, 30,* 449-472.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms.* Princeton, NJ: Princeton University Press.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., Fazio, F., & Mehler, J. (1996). Brain processing of native and foreign languages. *NeuroReport, 7*, 2439-2444.

Perani, D., Paulesu, E., Sebastian-Galles, N. Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F., & Mehler, J. (1998). The bilingual brain: proficiency and age of acquisition of the second language. *Brain, 121*, 1841-1852.

Pinker, S. (1994). *The language instinct: How the mind creates language.* New York: Morrow.

REFERENCES

Pitts Cochran, B., McDonald, J., & Parault, S. (1999). Too smart for their own good: The disadvantage of a superior processing capacity for adult language learners. *Journal of Memory and Language, 41*, 30-58.

Radford, A. (1997). *Syntactic theory and the structure of English. A minimalist approach*. Cambridge: Cambridge University Press.

Ratner, A., & Hoffman, H. (1974). Evidence for a critical period in khaki campbell ducklings (anas platyrhynchos domesticus). *Animal Behavior, 22*, 249-255.

Rohde, D., & Plaut, D. (1999). Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition, 72*, 67-109.

Sabourin, L. (2003). *Grammatical gender and second language processing: An ERP study*. Doctoral dissertation. University of Groningen, Groningen.

Schachter, J. (1990). On the issue of completeness in second language acquisition. *Second Language Research, 6,* 93-124.

Schütze, C. (1996). *The empirical base of linguistics. Grammaticality judgments and linguistic methodology*. Chicago: The University of Chicago Press.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House.

Seliger, H. (1978). Implications of a multiple critical periods hypothesis for second language learning. In W. Ritchie (Ed.), *Second language acquisition research: Issues and implications* (pp. 11-19). New York: Academic Press.

Sorace, A. (1993). Incomplete versus divergent representations of unaccusativity in non-native grammars of Italian. *Second Language Research, 9,* 22-47.

Sorace, A. (1996). The use of acceptability judgements in second language acquisition research. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 375-409). San Diego, CA: Academic Press.

Stevens, G. (1999). Age at immigration and second language proficiency among foreign-born adults. *Language in Society, 28,* 555-578.

Uit den Boogaart, P. (Ed.). (1975). *Woordfrequenties in geschreven en gesproken Nederlands*. Utrecht: Oosthoek, Scheltema en Holkema.

Van Boxtel, S. (1999). [Preliminary experiment with 5 native speakers on dummy subjects in Dutch]. Unpublished raw data.

Van Boxtel, S. (2003). [Native speaker judgements on responses on the sentence imitation task]. Unpublished raw data.

Van Boxtel, S., Bongaerts, T., & Coppen, P.-A. (2003). Native-like attainment in L2-syntax. In S. Foster-Cohen & S. Pekarek-Doehler (Eds.), *Eurosla yearbook, 2* (pp. 157-181). Amsterdam: John Benjamins.

Van den Hoek, T., Houtman, J., & Jullens, J. (1988). *Grammaticaal woordenboek. Termen en begrippen van de traditionele grammatica en hun equivalenten in het Latijn, Duits, Engels en Frans*. Leiden: Martinus Nijhoff.

Van Dreumel, S. (1999). [items from experiments on verb clusters in Dutch]. Unpublished data.

Vikner, S. (1995). *Verb movement and expletive subjects in the Germanic languages*. New York: Oxford University Press.

Weber-Fox, C., & Neville, H. (1996). Maturational constraints on functional specialization for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8*, 231-256.

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

White, L., & Juffs, A. (1998). Constraints on Wh-movement in two different contexts of non-native language acquisition: Competence and processing. In S. Flynn, G. Martohardjono & W. O'Neil (Eds.), *The generative study of second language acquisition* (pp. 111-129). Mahwah, NJ: Lawrence Erlbaum.

Yetkin, O., Yetkin, F., Haughton, V., & Cox, R. (1996). Use of functional MRI to map language in multilingual volunteers. *American Journal of Neuroradiology, 17*, 473-477.

Zifonun, G., Hoffmann, L., Strecker, B., Ballweg, J., Brauße, U., Breindl, E., Engel, U., Frosch, H., Hoberg, U., & Vorderwülbecke, K. (1997). *Grammatik der deutschen Sprache*. Berlin: Walter de Gruyter.

Appendix

A

Questionnaire second language learners

1.   Are you a man or a woman?
☐   Man
☐   Woman

2.   How old are you?
…   years old

3.   Are you left-handed or right-handed?
☐   left-handed
☐   right-handed

4.   Where were you born?
Country: _____
City/region: _____

5.   How long did you live there?
...   years

6.   What is your native language?

_____

7.   How old were you when you arrived in the Netherlands?

_____

8.   Since when have you lived in the Netherlands?
19…

9.   Did you also live in any other country? If so, how long and what
        language was most important there for daily life?
        …   years in ——————   Language: ——————

        …   years in ——————   Language: ——————

        …   years in ——————   Language: ——————

10.   Do you speak other languages than your native language and Dutch?
        ☐   No
        ☐   Yes, namely   ————————————

                                ————————————

                                ————————————

11.   If you answered "yes" to question 10, could you please indicate how
        old you were when you started learning those languages and with
        whom you speak/spoke those languages or in which situations (for
        example in shops, with friends, at home, at work)

        Language: ——————   Age: …   Years old

        I speak/spoke this language in the following situations / with the
        following people:

        _____

        Language: ——————   Age: …   Years old

        I speak/spoke this language in the following situations / with the
        following people:

        _____

        Language: ——————   Age: …   Years old

        I speak/spoke this language in the following situations / with the
        following people:

        _____

12.　If you answered "yes" to question 10, could you please indicate approximately how much you speak or have spoken these languages?

I have spoken the languages indicated in question 10

☐　very much　(Language(s): ＿＿＿＿＿＿＿)

☐　much　(Language(s): ＿＿＿＿＿＿＿)

☐　regularly　(Language(s): ＿＿＿＿＿＿＿)

☐　little　(Language(s): ＿＿＿＿＿＿＿)

☐　very little　(Language(s): ＿＿＿＿＿＿＿)

13.　How much do you use your native language in proportion to Dutch?
☐　about 90% native language, 10% Dutch
☐　about 70% native language, 30% Dutch
☐　about 50% native language, 50% Dutch
☐　about 30% native language, 70% Dutch
☐　about 10% native language, 90% Dutch

14.　How much do you use your native language and Dutch in proportion to other languages you speak?

☐　about 90% native language/Dutch, 10% other languages
☐　about 80% native language/Dutch, 20% other languages
☐　about 70% native language/Dutch, 30% other languages
☐　about 60% native language/Dutch, 40% other languages
☐　50% or less native language/Dutch, 50% or more other languages

15.　With whom and in which situations do you speak Dutch?

＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

16.   With whom and in which situations do you speak your native language?

    _____

    _____

    _____

17.   Please indicate to what extent you agree with the following statements:

    a.   I feel more confident in my native language than in Dutch.

I Totally agree		I Don't agree, don't disagree		I Totally disagree
☐	☐	☐	☐	☐

    b.   I like to speak my native language.

I Totally agree		I Don't agree, don't disagree		I Totally disagree
☐	☐	☐	☐	☐

    c.   I like to speak Dutch.

I Totally agree		I Don't agree, don't disagree		I Totally disagree
☐	☐	☐	☐	☐

    d.   I like to learn languages.

I Totally agree		I Don't agree, don't disagree		I Totally disagree
☐	☐	☐	☐	☐

    e.   I find it very important to use Dutch without errors.

I Totally agree		I Don't agree, don't disagree		I Totally disagree
☐	☐	☐	☐	☐

18.   How good is your proficiency in Dutch at the moment?

	Very poor				Very good
Speaking	☐	☐	☐	☐	☐
Listening	☐	☐	☐	☐	☐
Writing	☐	☐	☐	☐	☐
Reading	☐	☐	☐	☐	☐
Grammar	☐	☐	☐	☐	☐
General	☐	☐	☐	☐	☐

19.   How good is your proficiency in your mother tongue at the moment?

	Very poor				Very good
Speaking	☐	☐	☐	☐	☐
Listening	☐	☐	☐	☐	☐
Writing	☐	☐	☐	☐	☐
Reading	☐	☐	☐	☐	☐
Grammar	☐	☐	☐	☐	☐
General	☐	☐	☐	☐	☐

20.   How good is your proficiency at the moment in the languages you mentioned in question 10?

Language 1:   _____

	Very poor				Very good
Speaking	☐	☐	☐	☐	☐
Listening	☐	☐	☐	☐	☐
Writing	☐	☐	☐	☐	☐
Reading	☐	☐	☐	☐	☐
Grammar	☐	☐	☐	☐	☐
General	☐	☐	☐	☐	☐

Language 2:   _____

	Very poor				Very good
Speaking	☐	☐	☐	☐	☐
Listening	☐	☐	☐	☐	☐
Writing	☐	☐	☐	☐	☐
Reading	☐	☐	☐	☐	☐
Grammar	☐	☐	☐	☐	☐
General	☐	☐	☐	☐	☐

Language 3: _____

	Very poor				Very good
Speaking	☐	☐	☐	☐	☐
Listening	☐	☐	☐	☐	☐
Writing	☐	☐	☐	☐	☐
Reading	☐	☐	☐	☐	☐
Grammar	☐	☐	☐	☐	☐
General	☐	☐	☐	☐	☐

21. Have you attended Dutch classes?
☐ Yes
☐ No

22. Where did you attend Dutch classes?

Country: _____

Type of school/institute: _____

23. How long have you attended Dutch classes?

… weeks/months/years

24. If you have taken a Dutch course, what kind of course was it? (for example: preparation for the state exam, speaking, translation, general proficiency)

_____

_____

_____

25. What is your highest level of education?

Name: _____

Level: _____

Country: _____

Thank you very much for completing this questionnaire!

Appendix

B

Items in the sentence

preference task

The a- and b-sentences of the items in this appendix are presented in their original order (before randomisation). Items which contained errors and the filler items are left out. The item codes are based on Dutch labels for categories in the pilot test and have the following meaning:

wa: Perception verbs
ca: Categorical subjects
sp: Specific subjects
ns: Non-specific subjects
fi: Finiteness of the embedded clause
co: Conditionality of the embedded clause
ob: Object position occupied by a DP
ao: Other active sentences with a sentential subject
in: Change of state
po: Other passives
pv: Passives with a dummy object in their active equivalent

DP-type, general (DPg)

Perception verbs with a non-finite embedded clause
wa1a   Ik zag er bij mij in de straat iedere dag een paar kinderen spelen.
wa1b   Ik zag bij mij in de straat iedere dag een paar kinderen spelen.

wa2a   Zij hoorde in de trein er een man zachtjes haar naam noemen.
wa2b   Zij hoorde in de trein een man zachtjes haar naam noemen.

wa3a   Hij voelde er tijdens dat interview voor de radio steeds nieuwe vragen opkomen.

wa3b   Hij voelde tijdens dat interview voor de radio steeds nieuwe vragen opkomen.

wa4a   Ik voelde er afgelopen nacht onder de warme dekens iets engs bewegen.

wa4b   Ik voelde afgelopen nacht onder de warme dekens iets engs bewegen.

Categorical subjects

ca1a   Men beseft niet altijd dat er een pinguïn een vogel is.

ca1b   Men beseft niet altijd dat een pinguïn een vogel is.

ca3a   Je moet niet vergeten dat er topsporters ook maar mensen zijn.

ca3b   Je moet niet vergeten dat topsporters ook maar mensen zijn.

ca5a   Arlette vertelt mij net dat er poezen geen suiker kunnen proeven.

ca5b   Arlette vertelt mij net dat poezen geen suiker kunnen proeven.

Specific subjects

sp1a   Vandaag krijgt er iedereen van de jarige een klein cadeautje.

sp1b   Vandaag krijgt iedereen van de jarige een klein cadeautje.

sp3a   Vandaag zijn er alle leerlingen van groep vijf op school.

sp3b   Vandaag zijn alle leerlingen van groep vijf op school.

sp4a   Morgen gaan er de leerlingen uit de tweede klas naar het museum.

sp4b   Morgen gaan de leerlingen uit de tweede klas naar het museum.

sp6a   Vandaag hebben er de drie gitaristen duidelijk hun dag niet.

sp6b   Vandaag hebben de drie gitaristen duidelijk hun dag niet.

DP-type, non-specific transitive (DPnst)

Perception verbs with a finite embedded clause + transitive predicate

wa5a   Hij zag bij de bakker dat er een verkoopster een doos had laten vallen.

wa5b   Hij zag bij de bakker dat een verkoopster een doos had laten vallen.

wa6a   Wij voelden allemaal dat er sommige mensen iets voor ons geheim
       hielden.
wa6b   Wij voelden allemaal dat sommige mensen iets voor ons geheim
       hielden.

Non-specific subjects + transitive predicate
ns1a   Tijdens het diner krijgen er maar een paar speeches veel applaus.
ns1b   Tijdens het diner krijgen maar een paar speeches veel applaus.

ns2a   Op zijn verjaardag drinken er nooit zo veel mensen bier.
ns2b   Op zijn verjaardag drinken nooit zo veel mensen bier.
ns3a   Bij de wedstrijd Barcelona-Juventus heeft er een penalty de
        wedstrijd beslist.
ns3b   Bij de wedstrijd Barcelona-Juventus heeft een penalty de wedstrijd
        beslist.

ns4a   Op televisie doen er veel mensen dingen die ze normaal niet durven.
ns4b   Op televisie doen veel mensen dingen die ze normaal niet durven.

ns5a   Iemand heeft er een grote zwarte tas in mijn auto laten liggen.
ns5b   Iemand heeft een grote zwarte tas in mijn auto laten liggen.

ns6a   In het centrum hebben er twee jongens samen een telefooncel kapot
        gemaakt.
ns6b   In het centrum hebben twee jongens samen een telefooncel kapot
        gemaakt.

DP-type, non-specific intransitive (DPnsi)

Perception verbs with a finite embedded clause + intransitive predicate
wa7a   Zij voelden duidelijk dat er een harde wind over de weg waaide.
wa7b   Zij voelden duidelijk dat een harde wind over de weg waaide.

wa8a   Wij konden heel goed horen dat er bij de buren twee baby's huilden.
wa8b   Wij konden heel goed horen dat bij de buren twee baby's huilden.

Non-specific subjects + intransitive predicate
ns7a   Ik vind het vervelend dat er boven een raam openstaat.
ns7b   Ik vind het vervelend dat boven een raam openstaat.

ns8a   Ik hoop dat er vandaag nog een bus komt bij deze halte.
ns8b   Ik hoop dat vandaag nog een bus komt bij deze halte.

ns9a   Op dit feest zijn er heel veel mensen behoorlijk dronken.
ns9b   Op dit feest zijn heel veel mensen behoorlijk dronken.

ns10a   Normaal staan er vier borden op de hoek van deze tafel.
ns10b   Normaal staan vier borden op de hoek van deze tafel.

ns11a   In deze hoek zit er vaak een jongetje met autootjes te spelen.
ns11b   In deze hoek zit vaak een jongetje met autootjes te spelen.

ns12a   Tot nu toe komen er ieder jaar meer mensen naar dit popconcert.
ns12b   Tot nu toe komen ieder jaar meer mensen naar dit popconcert.

AS-type, general (ASg)

Non-finite embedded clauses
fi1a   Meestal valt het niet mee om kaartjes voor een concert te krijgen.
fi1b   Meestal valt niet mee om kaartjes voor een concert te krijgen.

fi2a   Meestal valt er niet mee om kaartjes voor een concert te krijgen.
fi2b   Meestal valt niet mee om kaartjes voor een concert te krijgen.

fi3a   Meestal valt het niet mee om kaartjes voor een concert te krijgen.
fi3b   Meestal valt er niet mee om kaartjes voor een concert te krijgen.

fi4a   Uit onderzoek blijkt dat het gezond is om veel aan sport te doen.
fi4b   Uit onderzoek blijkt dat gezond is om veel aan sport te doen.

fi5a   Uit onderzoek blijkt dat er gezond is om veel aan sport te doen.
fi5b   Uit onderzoek blijkt dat gezond is om veel aan sport te doen.

fi6a   Uit onderzoek blijkt dat het gezond is om veel aan sport te doen.
fi6b   Uit onderzoek blijkt dat er gezond is om veel aan sport te doen.

fi7a   Volgens haar is het beter voor het milieu om weinig water te gebruiken.
fi7b   Volgens haar is beter voor het milieu om weinig water te gebruiken.

fi8a   Volgens haar is er beter voor het milieu om weinig water te gebruiken.

fi8b   Volgens haar is beter voor het milieu om weinig water te gebruiken.

fi9a   Volgens haar is het beter voor het milieu om weinig water te gebruiken.

fi9b   Volgens haar is er beter voor het milieu om weinig water te gebruiken.

fi10a   Zij vindt dat het niet hoort om iemand heel lang aan te kijken.

fi10b   Zij vindt dat niet hoort om iemand heel lang aan te kijken.

fi11a   Zij vindt dat er niet hoort om iemand heel lang aan te kijken.

fi11b   Zij vindt dat niet hoort om iemand heel lang aan te kijken.

fi12a   Zij vindt dat het niet hoort om iemand heel lang aan te kijken.

fi12b   Zij vindt dat er niet hoort om iemand heel lang aan te kijken.

Conditional embedded clauses

co1a   Veel mensen zeiden dat het te betreuren zou zijn als alle discussies voor niets waren geweest.

co1b   Veel mensen zeiden dat te betreuren zou zijn als alle discussies voor niets waren geweest.

co2a   Veel mensen zeiden dat er te betreuren zou zijn als alle discussies voor niets waren geweest.

co2b   Veel mensen zeiden dat te betreuren zou zijn als alle discussies voor niets waren geweest.

co3a   Veel mensen zeiden dat het te betreuren zou zijn als alle discussies voor niets waren geweest.

co3b   Veel mensen zeiden dat er te betreuren zou zijn als alle discussies voor niets waren geweest.

co4a   Met dit weer zou het gek zijn als de wedstrijd van vanmiddag niet doorging.

co4b   Met dit weer zou gek zijn als de wedstrijd van vanmiddag niet doorging.

co5a   Met dit weer zou er gek zijn als de wedstrijd van vanmiddag niet doorging.

co5b   Met dit weer zou gek zijn als de wedstrijd van vanmiddag niet doorging.

co6a   Met dit weer zou het gek zijn als de wedstrijd van vanmiddag niet doorging.

co6b   Met dit weer zou er gek zijn als de wedstrijd van vanmiddag niet doorging.

co7a   Hij zei dat het wel heel erg saai zou zijn als iedereen precies hetzelfde was.

co7b   Hij zei dat het wel heel erg saai zou zijn als iedereen precies hetzelfde was.

co8a   Hij zei dat er wel heel erg saai zou zijn als iedereen precies hetzelfde was.

co8b   Hij zei dat wel heel erg saai zou zijn als iedereen precies hetzelfde was.

co9a   Hij zei dat het wel heel erg saai zou zijn als iedereen precies hetzelfde was.

co9b   Hij zei dat er wel heel erg saai zou zijn als iedereen precies hetzelfde was.

co10a   Gezien de omstandigheden zou het schandalig zijn als de aanleg van die snelweg doorging.

co10b   Gezien de omstandigheden zou schandalig zijn als de aanleg van die snelweg doorging.

co11a   Gezien de omstandigheden zou er schandalig zijn als de aanleg van die snelweg doorging.

co11b   Gezien de omstandigheden zou schandalig zijn als de aanleg van die snelweg doorging.

co12a   Gezien de omstandigheden zou het schandalig zijn als de aanleg van die snelweg doorging.

co12b   Gezien de omstandigheden zou er schandalig zijn als de aanleg van die snelweg doorging.

DP objects

ob1a   Hij zegt dat het jou ook helemaal niks aangaat wat daar gebeurd is.
ob1b   Hij zegt dat jou ook helemaal niks aangaat wat daar gebeurd is.

ob2a   Hij zegt dat er jou ook helemaal niks aangaat wat daar gebeurd is.
ob2b   Hij zegt dat jou ook helemaal niks aangaat wat daar gebeurd is.

ob3a   Hij zegt dat het jou ook helemaal niks aangaat wat daar gebeurd is.
ob3b   Hij zegt dat er jou ook helemaal niks aangaat wat daar gebeurd is.

ob4a   Olga zegt dat het haar wel wat doet dat haar boek slechte recensies heeft gekregen.
ob4b   Olga zegt dat haar wel wat doet dat haar boek slechte recensies heeft gekregen.

ob5a   Olga zegt dat er haar wel wat doet dat haar boek slechte recensies heeft gekregen.
ob5b   Olga zegt dat haar wel wat doet dat haar boek slechte recensies heeft gekregen.

ob6a   Olga zegt dat het haar wel wat doet dat haar boek slechte recensies heeft gekregen.
ob6b   Olga zegt dat er haar wel wat doet dat haar boek slechte recensies heeft gekregen.

ob7a   Volgens Sofie kan het Ludo helemaal niks schelen wat anderen van hem vinden.
ob7b   Volgens Sofie kan Ludo helemaal niks schelen wat anderen van hem vinden.

ob8a   Volgens Sofie kan er Ludo niks schelen wat anderen van hem vinden.
ob8b   Volgens Sofie kan Ludo niks schelen wat anderen van hem vinden.

ob9a   Volgens Sofie kan het Ludo niks schelen wat anderen van hem vinden.
ob9b   Volgens Sofie kan er Ludo niks schelen wat anderen van hem vinden.

ob10a Volgens mij is het haar veel geld waard dat het werk op tijd af is.
ob10b Volgens mij is haar veel geld waard dat het werk op tijd af is.

ob11a   Volgens mij is er haar veel geld waard dat het werk op tijd af is.
ob11b   Volgens mij is haar veel geld waard dat het werk op tijd af is.

ob12a   Volgens mij is het haar veel geld waard dat het werk op tijd af is.
ob12b   Volgens mij is er haar veel geld waard dat het werk op tijd af is.

Other active sentences with a sentential subject
ao1a   Het overkomt haar maar zelden dat zij te laat op school is.
ao1b   Er overkomt haar maar zelden dat zij te laat op school is.

ao2a   Zij zegt dat het haar maar zelden overkomt dat zij te laat op school
        is.
ao2b   Zij zegt dat haar maar zelden overkomt dat zij te laat op school is.
ao3a   Zij zegt dat er haar maar zelden overkomt dat zij te laat op school is.
ao3b   Zij zegt dat haar maar zelden overkomt dat zij te laat op school is.

ao4a   Zij zegt dat het haar maar zelden overkomt dat zij te laat op school
        is.
ao4b   Zij zegt dat er haar maar zelden overkomt dat zij te laat op school is.

ao5a   Het komt maar zelden voor dat Luc op kantoor een stropdas draagt.
ao5b   Er komt maar zelden voor dat Luc op kantoor een stropdas draagt.

ao6a   Ik denk dat het maar zelden voorkomt dat Luc op kantoor een
        stropdas draagt.
ao6b   Ik denk dat maar zelden voorkomt dat Luc op kantoor een stropdas
        draagt.

ao7a   Ik denk dat er maar zelden voorkomt dat Luc op kantoor een
        stropdas draagt.
ao7b   Ik denk dat maar zelden voorkomt dat Luc op kantoor een stropdas
        draagt.

ao8a   Ik denk dat het maar zelden voorkomt dat Luc op kantoor een
        stropdas draagt.
ao8b   Ik denk dat er maar zelden voorkomt dat Luc op kantoor een
        stropdas draagt.

ao9a   Het is mij nu duidelijk wie die moord op dat meisje heeft gepleegd.
ao9b   Er is mij nu duidelijk wie die moord op dat meisje heeft gepleegd.

ao10a   Nu is het mij duidelijk wie die moord op dat meisje heeft gepleegd.
ao10b   Nu is mij duidelijk wie die moord op dat meisje heeft gepleegd.

ao11a   Nu is er mij duidelijk wie die moord op dat meisje heeft gepleegd.
ao11b   Nu is mij duidelijk wie die moord op dat meisje heeft gepleegd.

ao12a   Nu is het mij duidelijk wie die moord op dat meisje heeft gepleegd.
ao12b   Nu is er mij duidelijk wie die moord op dat meisje heeft gepleegd.

ao13a   Het schijnt dat iedereen gisteren naar de finale van Wimbledon
        gekeken heeft.
ao13b   Er schijnt dat iedereen gisteren naar de finale van Wimbledon
        gekeken heeft.

ao14a   Volgens Feliz schijnt het dat iedereen gisteren naar de finale van
        Wimbledon gekeken heeft.
ao14b   Volgens Feliz schijnt dat iedereen gisteren naar de finale van
        Wimbledon gekeken heeft.

ao15a   Volgens Feliz schijnt er dat iedereen gisteren naar de finale van
        Wimbledon gekeken heeft.
ao15b   Volgens Feliz schijnt dat iedereen gisteren naar de finale van
        Wimbledon gekeken heeft.

ao16a   Volgens Feliz schijnt het dat iedereen gisteren naar de finale van
        Wimbledon gekeken heeft.
ao16b   Volgens Feliz schijnt er dat iedereen gisteren naar de finale van
        Wimbledon gekeken heeft.

AS-type, Change of state (AScos)
in1a   Het is nu gebleken dat de brand door een ongeluk is ontstaan.
in1b   Er is nu gebleken dat de brand door een ongeluk is ontstaan.

in2a   Dennis vertelt dat het nu is gebleken dat de brand door een ongeluk
        is ontstaan.
in2b   Dennis vertelt dat nu is gebleken dat de brand door een ongeluk is
        ontstaan.

in3a   Dennis vertelt dat er nu is gebleken dat de brand door een ongeluk is ontstaan.

in3b   Dennis vertelt dat nu is gebleken dat de brand door een ongeluk is ontstaan.

in4a   Dennis vertelt dat het nu is gebleken dat de brand door een ongeluk is ontstaan.

in4b   Dennis vertelt dat er nu is gebleken dat de brand door een ongeluk is ontstaan.

in5a   Het schiet mij ineens te binnen dat ik nog een boek moet terugbrengen.

in5b   Er schiet mij ineens te binnen dat ik nog een boek moet terugbrengen.

in6a   Nu schiet het mij ineens te binnen dat ik nog boodschappen moet doen.

in6b   Nu schiet mij ineens te binnen dat ik nog boodschappen moet doen.

in7a   Nu schiet er mij ineens te binnen dat ik nog boodschappen moet doen.

in7b   Nu schiet mij ineens te binnen dat ik nog boodschappen moet doen.

in8a   Nu schiet het mij ineens te binnen dat ik nog boodschappen moet doen.

in8b   Nu schiet er mij ineens te binnen dat ik nog boodschappen moet doen.

in9a   Het is na jaren van onderzoek vast komen te staan dat roken ongezond is.

in9b   Er is na jaren van onderzoek vast komen te staan dat roken ongezond is.

in10a   Na jaren van onderzoek is het vast komen te staan dat roken ongezond is.

in10b   Na jaren van onderzoek is vast komen te staan dat roken ongezond is.

in11a   Na jaren van onderzoek is er vast komen te staan dat roken ongezond is.

in11b   Na jaren van onderzoek is vast komen te staan dat roken ongezond is.

in12a   Na jaren van onderzoek is het vast komen te staan dat roken
         ongezond is.
in12b   Na jaren van onderzoek is er vast komen te staan dat roken
         ongezond is.

in13a   Het valt haar nu ineens op dat Paul een nieuwe bril heeft.
in13b   Er valt haar nu ineens op dat Paul een nieuwe bril heeft.

in14a   Wat gek dat het haar nu ineens opvalt dat Paul een nieuwe bril heeft.
in14b   Wat gek dat haar nu ineens opvalt dat Paul een nieuwe bril heeft.

in15a   Wat gek dat er haar nu ineens opvalt dat Paul een nieuwe bril heeft.
in15b   Wat gek dat haar nu ineens opvalt dat Paul een nieuwe bril heeft.

in16a   Wat gek dat het haar nu ineens opvalt dat Paul een nieuwe bril heeft.
in16b   Wat gek dat er haar nu ineens opvalt dat Paul een nieuwe bril heeft.

PS-type, general (PSg)
po1a   In de krant wordt het beweerd dat hij dronken achter het stuur
        gezeten heeft.
po1b   In de krant wordt beweerd dat hij dronken achter het stuur gezeten
        heeft.

po2a   In de krant wordt er beweerd dat hij dronken achter het stuur
        gezeten heeft.
po2b   In de krant wordt beweerd dat hij dronken achter het stuur gezeten
        heeft.

po3a   In de krant wordt het beweerd dat hij dronken achter het stuur
        gezeten heeft.
po3b   In de krant wordt er beweerd dat hij dronken achter het stuur
        gezeten heeft.

po4a   Hij zegt dat het hier onderzocht wordt of ze alsnog een visum kan
        krijgen.
po4b   Hij zegt dat hier onderzocht wordt of ze alsnog een visum kan
        krijgen.

po5a   Hij zegt dat er hier onderzocht wordt of ze alsnog een visum kan krijgen.
po5b   Hij zegt dat hier onderzocht wordt of ze alsnog een visum kan krijgen.

po6a   Hij zegt dat het hier onderzocht wordt of ze alsnog een visum kan krijgen.
po6b   Hij zegt dat er hier onderzocht wordt of ze alsnog een visum kan krijgen.

po7a   Zelden wordt het aan ouderen precies uitgelegd hoe de kaartautomaat op het station werkt.
po7b   Zelden wordt aan ouderen precies uitgelegd hoe de kaartautomaat op het station werkt.

po8a   Zelden wordt er aan ouderen precies uitgelegd hoe de kaartautomaat op het station werkt.
po8b   Zelden wordt aan ouderen precies uitgelegd hoe de kaartautomaat op het station werkt.

po9a   Zelden wordt het aan ouderen precies uitgelegd hoe de kaartautomaat op het station werkt.
po9b   Zelden wordt er aan ouderen precies uitgelegd hoe de kaartautomaat op het station werkt.

po10a Ria weet dat het mij is meegedeeld dat ik niet meer terug mag komen.
po10b Ria weet dat mij is meegedeeld dat ik niet meer terug mag komen.

po11a Ria weet dat er mij is meegedeeld dat ik niet meer terug mag komen.
po11b Ria weet dat mij is meegedeeld dat ik niet meer terug mag komen.

po12a Ria weet dat het mij is meegedeeld dat ik niet meer terug mag komen.
po12b Ria weet dat er mij is meegedeeld dat ik niet meer terug mag komen.

PS-type, dummy object in active equivalent (PSdo)
pv1a   Zij denkt dat het niet erg gewaardeerd wordt dat hij hard gewerkt heeft.
pv1b   Zij denkt dat niet erg gewaardeerd wordt dat hij hard gewerkt heeft.

pv2a Zij denkt dat er niet erg gewaardeerd wordt dat hij hard gewerkt heeft.

pv2b Zij denkt dat niet erg gewaardeerd wordt dat hij hard gewerkt heeft.

pv3a Zij denkt dat het niet erg gewaardeerd wordt dat hij hard gewerkt heeft.

pv3b Zij denkt dat er niet erg gewaardeerd wordt dat hij hard gewerkt heeft.

pv4a Door haar vrienden wordt het bewonderd dat ze ook in moeilijke tijden vrolijk blijft.

pv4b Door haar vrienden wordt bewonderd dat ze ook in moeilijke tijden vrolijk blijft.

pv5a Door haar vrienden wordt er bewonderd dat ze ook in moeilijke tijden vrolijk blijft.

pv5b Door haar vrienden wordt bewonderd dat ze ook in moeilijke tijden vrolijk blijft.

pv6a Door haar vrienden wordt het bewonderd dat ze ook in moeilijke tijden vrolijk blijft.

pv6b Door haar vrienden wordt er bewonderd dat ze ook in moeilijke tijden vrolijk blijft.

pv7a Hij gelooft dat het geheim gehouden wordt dat de koningin morgen naar Amsterdam komt.

pv7b Hij gelooft dat geheim gehouden wordt dat de koningin morgen naar Amsterdam komt.

pv8a Hij gelooft dat er geheim gehouden wordt dat de koningin morgen naar Amsterdam komt.

pv8b Hij gelooft dat geheim gehouden wordt dat de koningin morgen naar Amsterdam komt.

pv9a Hij gelooft dat het geheim gehouden wordt dat de koningin morgen naar Amsterdam komt.

pv9b Hij gelooft dat er geheim gehouden wordt dat de koningin morgen naar Amsterdam komt.

pv10a   Door bijna iedereen wordt het jammer gevonden dat Pieter niet meer terug komt.

pv10b   Door bijna iedereen wordt jammer gevonden dat Pieter niet meer terug komt.

pv11a   Door bijna iedereen wordt er jammer gevonden dat Pieter niet meer terug komt.

pv11b   Door bijna iedereen wordt jammer gevonden dat Pieter niet meer terug komt.

pv12a   Door bijna iedereen wordt het jammer gevonden dat Pieter niet meer terug komt.

pv12b   Door bijna iedereen wordt er jammer gevonden dat Pieter niet meer terug komt.

Appendix

C

Native speaker results

for items in the SPT

The sentence preference task contained 190 items: 124 test items and 66 filler items. In this appendix, an overview is given of the results of the native speakers for all test items in the SPT (except the five items that contained errors). Of these 119 items, forty items were excluded, on different grounds:

- Items for which less than 90% of the native speakers had the NS pattern, were excluded from the analysis
- The whole change of state category was excluded because there were too few consistent items to compute meaningful statistic analyses
- Items were excluded when the predicates in these items behaved differently from the other predicates in the same category

In the last column of each table, it is indicated whether the item was included or excluded in the main analyses. When an item was excluded, because the whole predicate (p) or dummy subject pair (dsp) of a certain category was excluded, this is indicated between brackets.

For each item the median (value above and below which half of the cases fall) and mode (most frequent value) for the native speaker control group (N = 44) is given, together with the native speaker pattern for that item, values included in the native speaker pattern and the number of native speakers that did not have this pattern (the deviant values are presented in brackets). The values represent the values on the seven-point scale that participants used. These values have the following meaning:

1:      strong preference for *er* to *0* or
        strong preference for *het* to *er* or
        strong preference for *het* to *0*
4:      no preference

7:      strong preference for *0* to *er* or
        strong preference for *er* to *het* or
        strong preference for *0* to *het*

It should be noted that items of the AS-type with numbers 1-4, 5-8, 9-12 and 13-16 belong to the same predicate. For the PS-type this is the case for items with numbers 1-3, 4-6, 7-9 and 10-12. Information about the categories (total number of items, number of items in main analyses and NS pattern) is presented below the tables. Those dummy subjects for which no information is given about their position occurred in non-initial position.

Table C.1   DP-type, general: er - 0

Item	subcategory	median	mode	NS pattern	deviations from NS pattern	excluded?
wa01		7	7	6-7	4 (1,3,4)	no
wa02	perception verbs with a	7	7	6-7	1 (1)	no
wa03	non-finite embedded clause	7	7	6-7	0	no
wa04		7	7	6-7	6 (1,4,5)	yes
ca01		7	7	6-7	0	no
ca03	categorical subjects	7	7	6-7	1 (1)	no
ca05		7	7	6-7	0	no
sp01		7	7	6-7	4 (1,3,4)	no
sp03	specific subjects	7	7	6-7	1 (3)	no
sp04		7	7	6-7	4 (1,4,5)	no
sp06		7	7	6-7	2 (1,4)	no

Number of items: 11
Number of items included in analysis: 10
Native speaker pattern: 0 > er

Table C.2   DP-type, non-specific transitive: er – 0

Item	subcategory	median	mode	NS pattern	deviations from NS pattern	excluded?
wa05	perception verbs with a	6	7	4-7	2 (1,2)	no
wa06	finite embedded clause	7	7	4-7	1 (1)	no
ns01		5	4	4-7	7 (1,2,3)	yes
ns02		4	4	1-7	-	yes
ns03	non-specific subjects +	7	7	4-7	0	no
ns04	transitive predicate	6	6+7	4-7	1 (1)	no
ns05		7	7	4-7	0	no
ns06		7	7	4-7	2 (2,3)	no

Number of items: 8
Number of items included in analysis: 6
Native speaker pattern: 0 >= er

Table C.3   DP-type, non-specific intransitive: er – 0

Item	subcategory	median	mode	NS pattern	deviations from NS pattern	excluded?
wa07	perception verbs with a	2	1,2,4	1-4	3 (5,6)	no
wa08	finite embedded clause	3	4	1-4	7 (5,6,7)	yes
ns07		3	4	1-4	1 (7)	no
ns08		1	1	1-4	3 (5,6)	no
ns09	non-specific subjects +	4	4	1-7	-	yes
ns10	intransitive predicate	2	1	1-4	0	no
ns11		6	7	4-7	7 (2,3)*	yes
ns12		2.5	4	1-4	4 (5,6,7)	no

*   There were 31 native speakers who had a score between 5 and 7 for this item, and only 13 native speakers had the pattern 1-4 for this item.

Number of items: 8
Number of items included in analysis: 5
Native speaker pattern: er >= 0

Table C.4   AS-type, general: het - er

Item	subcategory	median	mode	NS pattern	deviations from NS pattern	excluded?
ao01		1	1	1-2	3 (4,7)	no
ao05	other sentences, DS in	1	1	1-2	0	no
ao09	initial position	1.5	1	1-2	5 (3,4)	yes (p)
ao13		1	1	1-2	6 (3,4,7)	yes (p)
fi03		1	1	1-2	0	no
fi06	non-finite embedded	1	1	1-2	1 (7)	no
fi09	clauses	1	1	1-2	1 (6)	no
fi12		1	1	1-2	1 (7)	no
co03		1	1	1-2	1 (4)	no
co06	conditional embedded	1	1	1-2	0	no
co09	clauses	1	1	1-2	2 (7)	no
co12		1	1	1-2	0	no
ob03		1	1	1-2	0	no
ob06		1	1	1-2	0	no
ob09	DP objects	1	1	1-2	0	no
ob12		1	1	1-2	0	no
ao04		1	1	1-2	1 (4)	no
ao08		1	1	1-2	1 (7)	no
ao12	other sentences	1	1	1-2	6 (3,4)	yes (p)
ao16		1	1	1-2	9 (3,4,6,7)	yes (p)

Number of items: 20
Number of items included in analysis: 16
Native speaker pattern: het > er

159

Table C.5   AS-type, general: het - 0

Item	subcategory	median	mode	NS pattern	deviations from NS pattern	excluded?
fi01		1	1	1-2	0	no
fi04	non-finite embedded	1	1	1-2	2 (3,4)	no
fi07	clauses	1	1	1-2	4 (3,7)	no
fi10		1	1	1-2	0	no
co01		1	1	1-2	7 (3,4)	yes
co04	conditional embedded	1	1	1-2	1 (4)	no
co07	clauses	1	1	1-2	3 (6,7)	no
co10		1	1	1-2	0	no
ob01		1	1	1-2	2 (4)	no
ob04	DP objects	1	1	1-2	0	no
ob07		1	1	1-2	1 (4)	no
ob10		1	1	1-2	0	no
ao02		1	1	1-2	4 (4,6)	no
ao06	other sentences	1	1	1-2	11 (3,4,6,7)	yes
ao10		4	4	1-7	-	yes (p)
ao14		2.5	1	1-4	5 (5,6,7)	yes (p)

Number of items: 16
Number of items included in analysis: 12
Native speaker pattern: het > 0

Table C.6   AS-type, general: er - 0

Item	subcategory	median	mode	NS pattern	deviations from NS pattern	excluded?
fi02		4	4	4-7	3 (1,2,3)	no
fi05	non-finite embedded	4	4	4-7	1 (3)	no
fi08	clauses	4	4	4-7	2 (3)	no
fi11		4	4	4-7	1 (3)	no
co02		5	4	4-7	3 (1,3)	no
co05	conditional embedded	4	4	4-7	1 (3)	no
co08	clauses	4	4	4-7	1 (3)	no
co11		4	4	4-7	1 (3)	no
ob02		5	4+6	4-7	1 (2)	no
ob05	DP objects	4	4	4-7	0	no
ob08		4	4	4-7	3 (1,3)	no
ob11		4	4	4-7	4 (1,3)	no
ao03		6	7	4-7	2 (1,2)	no
ao07	other sentences	5	4	4-7	1 (3)	no
ao11		7	7	4-7	1 (3)	yes (p)
ao15		6	6	4-7	1 (2)	yes (p)

Number of items: 16
Number of items included in analysis: 14
Native speaker pattern: 0 >= er

Table C.7   AS-type, change of state

Item	DS pair	median	mode	NS pattern	deviations from NS pattern	excluded?
in01		4	4	1-4	10 (5,6,7)	yes (p)
in05	het-er initial	4	4	1-4	8 (5,6,7)	yes
in09		4	4	1-4	12 (5,6,7)	yes
in13		1	1	1-4	1 (6)	yes (dsp)
in04		3.5	4	1-4	10 (5,6,7)	yes (p)
in08	het-er	2	1	1-4	6 (5,6,7)	yes
in12		3	4	1-4	10 (5,6,7)	yes
in16		1	1	1-4	1 (7)	yes (dsp)
in02		4	4	1-7	-	yes (p)
in06	het-0	4	4	1-7	-	yes
in10		5	4	4-7	6 (1,2,3)	yes
in14		4	1	1-7	-	yes
in03		4	4	1-7	-	yes (p)
in07	er-0	7	7	4-7	1 (2)	no
in11		6	7	4-7	4 (2,3)	no
in15		7	7	4-7	1 (2)	no

Number of items: 16
Number of items included in analysis: 3
Native speaker pattern: 0/het >= er

Table C.8   PS-type, general

Item	DS pair	median	mode	NS pattern	deviations from NS pattern	excluded?
po03		6	7	4-7	2 (1,2)	no
po06	het-er	6	7	4-7	3 (2)	no
po09		6	7	4-7	4 (1,2)	no
po12		4	4	4-7	10 (1,2,3)	yes (p)
po01		7	7	4-7	1 (2)	no
po04	het-0	6	6	4-7	5 (2,3)	yes
po07		6	7	4-7	4 (2,3)	no
po10		6	7	4-7	5 (2,3)	yes (p)
po02		5	4	4-7	3 (2,3)	yes (dsp)
po05	er-0	4	4	4-7	9 (1,2,3)	yes
po08		4	4	4-7	8 (2,3)	yes
po11		5	4	4-7	8 (2,3)	yes (p)

Number of items: 12
Number of items included in analysis: 5
Native speaker pattern: 0 >= er >= het

Table C.9   PS-type, dummy object in active equivalent

Item	DS pair	median	mode	NS pattern	deviations from NS pattern	excluded?
pv03		1	1	1-3	1 (4)	no
pv06	het-er	1	1	1-3	4 (4,7)	no
pv09		2	1	1-4	4 (5,7)	yes (p)
pv12		1	1	1-3	2 (4)	no
pv01		1	1	1-4	2 (7)	no
pv04	het-0	2	1	1-4	2 (6,7)	no
pv07		3.5	4	1-4	8 (5,6,7)	yes (p)
pv10		1	1	1-4	0	no
pv02		4	4	1-7	-	yes
pv05	er-0	6	7	4-7	2 (2,3)	no
pv08		4	4	1-7	-	yes (p)
pv11		5	4	4-7	0	no

Number of items: 12
Number of items included in analysis: 8
Native speaker pattern: het >= 0 >= er

Appendix

D

Scores on the sentence preference task

In chapter 5, the results on the sentence preference task are presented and discussed. For the individual results, chapter 5 provides an overview of how many participants fell within the native speaker range for each category in the task and for the total score based on all items in the analyses. In this appendix, we present the individual scores for all second language participants for these analyses. For the additional analyses in which we computed difference scores between categories of the same type, we present the difference scores for all L2 participants within the native speaker range. In all the tables containing individual results (tables D.2, D.4, D.6), the scores that do not fall within the native speaker range (-1.96 < z < 1.96) are marked grey.

In order to be able to interpret the scores of the second language participants, we also included some group statistics in tables D.1, D.3 and D.5. It should be noted that the mean z-score (score minus mean, divided by the standard deviation) for the native speakers is always exactly 0 in these tables and the standard deviation (SD) always exactly 1, because all z-scores were computed on the basis of the mean and SD of the native speaker control group.

Table D.1   Group scores for the DP-type

Group		DPg	zDPg	DPnst	zDPnst	DPnsi	zDPnsi
L1D	N	44	44	44	44	44	44
	Mean	9.61	.0000	5.39	.0000	4.75	.0000
	Minimum	7	-3.6187	3	-2.5242	4	-1.7123
	Maximum	10	.5349	6	.6491	5	.5708
	Std. Deviation	.722	1.0000	.945	1.0000	.438	1.0000
L1G	N	15	15	15	15	15	15
	Mean	8.33	-1.7726	4.87	-.5497	3.93	-1.8645
	Minimum	6	-5.0032	2	-3.5820	1	-8.5613
	Maximum	10	.5349	6	.6491	5	.5708
	Std. Deviation	1.447	2.0041	1.457	1.5415	1.280	2.9220
L1F	N	15	15	15	15	15	15
	Mean	9.60	-.0189	5.33	-.0561	4.07	-1.5601
	Minimum	8	-2.2341	3	-2.5242	2	-6.2783
	Maximum	10	.5349	6	.6491	5	.5708
	Std. Deviation	.632	.8757	.976	1.0323	1.100	2.5108
L1T	N	13	13	13	13	13	13
	Mean	8.77	-1.1691	4.54	-.8969	3.92	-1.8879
	Minimum	7	-3.6187	2	-3.5820	3	-3.9953
	Maximum	10	.5349	6	.6491	5	.5708
	Std. Deviation	1.092	1.5118	1.561	1.6509	.862	1.9687
Total	N	87	87	87	87	87	87
	Mean	9.26	-.4836	5.16	-.2385	4.37	-.8725
	Minimum	6	-5.0032	2	-3.5820	1	-8.5613
	Maximum	10	.5349	6	.6491	5	.5708
	Std. Deviation	1.051	1.4546	1.180	1.2482	.904	2.0631

Table D.2   Individual scores for the DP-type

Group	ID	DPg	zDPg	DPnst	zDPnst	DPnsi	zDPnsi
L1G	51	7	-3.6187	3	-2.5242	5	.5708
	52	8	-2.2341	3	-2.5242	4	-1.7123
	53	10	.5349	6	.6491	3	-3.9953
	54	6	-5.0032	5	-.4087	3	-3.9953
	55	6	-5.0032	3	-2.5242	5	.5708
	56	8	-2.2341	6	.6491	4	-1.7123
	57	7	-3.6187	2	-3.5820	5	.5708
	58	9	-.8496	6	.6491	2	-6.2783
	59	9	-.8496	6	.6491	4	-1.7123
	60	10	.5349	6	.6491	1	-8.5613
	111	10	.5349	6	.6491	3	-3.9953
	62	10	.5349	5	-.4087	5	.5708
	63	7	-3.6187	4	-1.4664	5	.5708
	64	9	-.8496	6	.6491	5	.5708
	65	9	-.8496	6	.6491	5	.5708
L1F	66	10	.5349	6	.6491	5	.5708
	117	10	.5349	6	.6491	2	-6.2783
	68	10	.5349	6	.6491	4	-1.7123
	69	10	.5349	6	.6491	5	.5708
	70	8	-2.2341	6	.6491	4	-1.7123
	71	10	.5349	6	.6491	5	.5708
	72	9	-.8496	6	.6491	4	-1.7123
	73	10	.5349	6	.6491	5	.5708
	74	9	-.8496	5	-.4087	5	.5708
	75	10	.5349	5	-.4087	2	-6.2783
	76	9	-.8496	3	-2.5242	3	-3.9953
	77	10	.5349	5	-.4087	5	.5708
	78	9	-.8496	4	-1.4664	3	-3.9953
	79	10	.5349	4	-1.4664	5	.5708
	80	10	.5349	6	.6491	4	-1.7123
L1T	82	10	.5349	6	.6491	5	.5708
	83	10	.5349	6	.6491	5	.5708
	84	10	.5349	6	.6491	3	-3.9953
	86	10	.5349	6	.6491	3	-3.9953
	87	8	-2.2341	5	-.4087	4	-1.7123
	88	7	-3.6187	2	-3.5820	5	.5708
	89	9	-.8496	3	-2.5242	4	-1.7123
	90	8	-2.2341	6	.6491	4	-1.7123
	91	9	-.8496	2	-3.5820	5	.5708
	92	9	-.8496	3	-2.5242	3	-3.9953
	93	7	-3.6187	5	-.4087	3	-3.9953
	94	8	-2.2341	4	-1.4664	3	-3.9953
	95	9	-.8496	5	-.4087	4	-1.7123

Table D.3   Group results for the AS-type, PS-type and total

Group		ASg	zASg	PSg	zPSg	PSdo	zPSdo	Total	zTotal
L1D	N	44	44	44	44	44	44	44	44
	Mean	40.82	.0000	4.68	.0000	7.70	.0000	75.82	.0000
	Min.	34	-4.3315	1	-4.9756	5	-4.2814	68	-3.0077
	Max.	42	.7508	5	.4300	8	.4677	79	1.2241
	SD	1.574	1.0000	.740	1.0000	.632	1.0000	2.599	1.0000
L1G	N	15	15	15	15	15	15	15	15
	Mean	40.00	-.5198	4.60	-.1106	7.20	-.7987	71.93	-1.4945
	Min.	35	-3.6962	3	-2.2728	6	-2.6983	67	-3.3924
	Max.	42	.7508	5	.4300	8	.4677	78	.8394
	SD	2.449	1.5561	.737	.9957	.676	1.0703	3.826	1.4719
L1F	N	15	15	15	15	15	15	15	15
	Mean	39.47	-.8586	3.80	-1.1917	6.93	-1.2208	71.80	-1.5458
	Min.	29	-7.5079	1	-4.9756	4	-5.8644	60	-6.0853
	Max.	42	.7508	5	.4300	8	.4677	78	.8394
	SD	3.399	2.1592	1.207	1.6313	1.223	1.9357	5.467	2.1031
L1T	N	13	13	13	13	13	13	13	13
	Mean	34.92	-3.7451	2.62	-2.7926	6.23	-2.3330	63.15	-4.8720
	Min.	19	-13.8607	0	-6.3270	3	-7.4474	43	-12.6253
	Max.	42	.7508	4	-.9214	8	.4677	76	.0699
	SD	7.005	4.4505	1.502	2.0210	1.481	2.3439	9.711	3.7359
Total	N	87	87	87	87	87	87	87	87
	Mean	39.56	-.7973	4.21	-.6418	7.26	-.6968	72.56	-1.2522
	Min.	19	-13.8607	0	-6.3270	3	-7.4474	43	-12.6253
	Max.	42	.7508	5	.4300	8	.4677	79	1.2241
	SD	3.878	2.4637	1.212	1.6375	1.051	1.6631	6.539	2.5156

Table D.4   Individual scores for the AS-type, PS-type and total

Group	ID	ASg	zASg	PSg	zPSg	PSdo	zPSdo	Total	ztotal
L1G	51	42	.7508	5	.4300	8	.4677	73	-1.0842
	52	39	-1.1551	5	.4300	7	-1.1153	69	-2.6230
	53	41	.1155	3	-2.2728	7	-1.1153	73	-1.0842
	54	38	-1.7903	5	.4300	7	-1.1153	67	-3.3924
	55	39	-1.1551	5	.4300	7	-1.1153	68	-3.0077
	56	35	-3.6962	5	.4300	7	-1.1153	68	-3.0077
	57	39	-1.1551	5	.4300	7	-1.1153	68	-3.0077
	58	41	.1155	5	.4300	7	-1.1153	73	-1.0842
	59	42	.7508	5	.4300	8	.4677	77	.4547
	60	42	.7508	3	-2.2728	8	.4677	73	-1.0842
	111	35	-3.6962	4	-.9214	7	-1.1153	68	-3.0077
	62	42	.7508	5	.4300	8	.4677	78	.8394
	63	41	.1155	5	.4300	6	-2.6983	71	-1.8536
	64	42	.7508	4	-.9214	8	.4677	77	.4547
	65	42	.7508	5	.4300	6	-2.6983	76	.0699
L1F	66	39	-1.1551	4	-.9214	6	-2.6983	73	-1.0842
	117	40	-.5198	4	-.9214	7	-1.1153	71	-1.8536
	68	40	-.5198	5	.4300	8	.4677	76	.0699
	69	42	.7508	3	-2.2728	8	.4677	77	.4547
	70	41	.1155	5	.4300	7	-1.1153	74	-.6995
	71	40	-.5198	5	.4300	8	.4677	77	.4547
	72	42	.7508	5	.4300	7	-1.1153	76	.0699
	73	42	.7508	4	-.9214	8	.4677	78	.8394
	74	29	-7.5079	4	-.9214	5	-4.2814	60	-6.0853
	75	41	.1155	2	-3.6242	8	.4677	70	-2.2383
	76	36	-3.0609	3	-2.2728	6	-2.6983	62	-5.3159
	77	37	-2.4256	5	.4300	7	-1.1153	71	-1.8536
	78	40	-.5198	1	-4.9756	7	-1.1153	66	-3.7771
	79	42	.7508	3	-2.2728	8	.4677	75	-.3148
	80	41	.1155	4	-.9214	4	-5.8644	71	-1.8536
L1T	82	41	.1155	0	-6.3270	5	-4.2814	70	-2.2383
	83	42	.7508	4	-.9214	7	-1.1153	76	.0699
	84	41	.1155	3	-2.2728	8	.4677	74	-.6995
	86	36	-3.0609	2	-3.6242	6	-2.6983	66	-3.7771
	87	37	-2.4256	1	-4.9756	8	.4677	66	-3.7771
	88	42	.7508	3	-2.2728	7	-1.1153	68	-3.0077
	89	26	-9.4137	2	-3.6242	3	-7.4474	49	-10.3171
	90	41	.1155	4	-.9214	7	-1.1153	71	-1.8536
	91	31	-6.2373	4	-.9214	7	-1.1153	60	-6.0853
	92	31	-6.2373	0	-6.3270	7	-1.1153	55	-8.0088
	93	19	-13.8607	4	-.9214	4	-5.8644	43	-12.6253
	94	32	-5.6020	4	-.9214	6	-2.6983	59	-6.4700
	95	35	-3.6962	3	-2.2728	6	-2.6983	64	-4.5465

Table D.5   Group results for the difference score analyses

Group		difDP		difAS		difPS			
		nst-nsi	zDP	het-0	zAS	het-0	zPS het-0	het-er	zPS het-er
L1D	N	44	44	44	44	44	44	44	44
	Mean	3.9121	.0000	2.9535	.0000	3.8030	.0000	4.2348	.0000
	Min.	.93	-2.4378	.26	-1.9789	.67	-2.0291	.00	-2.8033
	Max.	6.00	1.7087	6.00	2.2398	6.00	1.4213	6.00	1.1685
	SD	1.2219	1.0000	1.3602	1.0000	1.5457	1.0000	1.5107	1.0000
L1G	N	15	15	15	15	15	15	15	15
	Mean	2.6044	-1.0702	3.1619	.1532	3.4444	-.2320	3.3778	-.5674
	Min.	.80	-2.5469	1.67	-.9461	.67	-2.0291	1.67	-1.7001
	Max.	5.07	.9449	5.14	1.6097	6.00	1.4213	5.67	.9478
	SD	1.4899	1.2193	.9555	.7025	1.9586	1.2671	1.2464	.8251
L1F	N	15	15	15	15	15	15	15	15
	Mean	3.2933	-.5064	1.8159	-.8364	2.3556	-.9364	2.9556	-.8468
	Min.	.63	-2.6833	-.07	-2.2239	-.67	-2.8917	-.33	-3.0240
	Max.	5.40	1.2177	4.00	.7694	6.00	1.4213	6.00	1.1685
	SD	1.4245	1.1658	1.0446	.7680	2.1324	1.3795	2.0926	1.3852
L1T	N	13	13	13	13	13	13	13	13
	Mean	2.8872	-.8388	1.3407	-1.1857	.6667	-2.0291	1.4872	-1.8189
	Min.	1.23	-2.1923	-.24	-2.3465	-2.00	-3.7543	-3.00	-4.7892
	Max.	6.00	1.7087	6.00	2.2398	4.00	.12743	6.00	1.1685
	SD	1.5964	1.3065	1.6981	1.2484	1.6499	1.0674	2.4176	1.6004
Total	N	87	87	87	87	87	87	87	87
	Mean	3.4268	-.3972	2.5523	-.2950	3.0230	-.5047	3.4559	-.5156
	Min.	.63	-2.6833	-.24	-2.3465	-2.00	-3.7543	-3.00	-4.7892
	Max.	6.00	1.7087	6.00	2.2398	6.00	1.4213	6.00	1.1685
	SD	1.4420	1.1801	1.4551	1.0698	2.0490	1.3256	1.9655	1.3011

Table D.6   Difference scores for participants within NS range

Group	ID	difDP nst-nsi	difDP zDP	difAS het-0	difAS zAS	difPS het-0	difPS zPS het-0	difPS het-er	zPS het-er
L1G	59	4.60	.5630	4.00	.7694	5.00	.7744	3.67	-.3761
	62	4.30	.3174	2.93	-.0183	4.67	.5587	5.33	.7272
	64	5.07	.9449	1.67	-.9461	5.33	.9900	2.33	-1.2587
L1F	68	3.60	-.2554	2.57	-.2809	6.00	1.4213	5.33	.7272
	71	4.73	.6721	2.36	-.4384	4.67	.5587	5.00	.5065
	72	3.67	-.2009	1.67	-.9461	.00	-2.4604	3.67	-.3761
	73	4.43	.4266	2.52	-.3159	3.00	-.5195	5.00	.5065
L1T	83	6.00	1.7087	6.00	2.2398	2.00	-1.1665	5.00	.5065

Appendix

E

Results questionnaire

second language learners

In chapter 5, some of the results from the questionnaire were related to the results from the sentence preference test. In this appendix, we will present most results from the questionnaire. For each variable, we will distinguish between the results from the eight L2 speakers that scored within the NS range on the SPT according to our very strict criteria and those of all other L2 speakers. For the variables that can be quantified, we will also report the results for all L2 speakers together. We will first present the variables that can be quantified in two tables: one for all variables regarding language skills and one for all other variables that can be quantified. The other variables will be presented in lists. Within these tables and lists, we follow the order in the questionnaire (see appendix A) as much as possible. It should be noted that there is not always a direct relation between questions in the questionnaire and the variables derived from them. The numbers in the second column of each table refer to the question number in the questionnaire that the variable is derived from. In the last three columns, the mean (unless otherwise indicated) and range (within brackets) for the three groups are presented, except for categorical variables, such as sex. For these variables, the numbers in each category are given (e.g. men-women). In this appendix we do not report on significance of the differences between the groups. All relevant significance effects are reported in chapter 5.

Table E.1   Language skills

Variables	Q. nr	within NS range	other L2 speakers	total L2 speakers
speaking Dutch	18	4.5 (4-5)	4.1 (3-5)	4.2 (3-5)
listening Dutch	18	4.6 (4-5)	4.3 (3-5)	4.4 (3-5)
writing Dutch	18	4.0 (2-5)	3.4 (2-5)	3.5 (2-5)
reading Dutch	18	4.8 (4-5)	4.6 (3-5)	4.6 (3-5)
grammar Dutch	18	3.8 (2-5)	3.6 (1-5)	3.6 (1-5)
general proficiency Dutch	18	4.3 (3-5)	4.0 (3-5)	4.1 (3-5)
speaking L1	19	4.5 (3-5)	4.5 (3-5)	4.5 (3-5)
listening L1	19	4.9 (4-5)	4.9 (4-5)	4.9 (4-5)
writing L1	19	4.4 (3-5)	4.4 (3-5)	4.4 (3-5)
reading L1	19	4.9 (4-5)	4.8 (3-5)	4.8 (3-5)
grammar L1	19	4.5 (3-5)	4.5 (3-5)	4.5 (3-5)
general proficiency L1	19	4.6 (4-5)	4.6 (3-5)	4.6 (3-5)
speaking other language[*]	20	3.6 (3-5)	2.9 (1-5)	3.1 (1-5)
listening other language[*]	20	4.4 (3-5)	3.3 (1-5)	3.5 (1-5)
writing other language[*]	20	3.6 (3-5)	2.9 (1-5)	3.1 (1-5)
reading other language[*]	20	4.4 (4-5)	3.5 (1-5)	3.7 (1-5)
grammar other language[*]	20	3.5 (2-5)	3.2 (1-5)	3.2 (1-5)
general proficiency other language[*]	20	3.5 (3-5)	3.2 (1-5)	3.2 (1-5)

1 = very poor; 5 = very good

N.B. Because of the small number of participants who reported their skills in more than one
      other language, we only report the results for the participants' best language here.
[*]   Only participants who had answered this question were considered (n=8 for within NS
      range, n=32 for other L2 speakers and n=40 for total L2 speakers)

Table E.2   Other background characteristics

Variables	Q. nr	within NS range	others	total L2 speakers
sex (m/f)	1	1/7	19/16	20/23
age	2	45 (33-55)	44 (23-73)	44 (23-73)
handedness (left/right)	3	1/6[*]	3/32	4/38[*]
age of arrival (AoA)	7	23 (12-35)	21 (12-35)	22 (12-35)
Length of Residence (LoR)	8	20 (10-42)	22 (4-50)	21 (4-50)
LoR other countries[**]	9	2 (1-4)	3 (1-9)	3 (1-9)
age of onset best other language	11	11 (9-15)	12 (0-30)	12 (0-30)
amount of use of other languages[***]	12	little - very much	very little - very much	very little - very much
proportion use L1-Dutch[***]	13	70%/30% - 10%/90%	70%/30% - 10%/90%	70%/30% - 10%/90%
use other languages	14	max. 30%	max. 40%	max. 40%
statement a[***]	17	totally agree - totally disagree	totally agree - totally disagree	totally agree - totally disagree
statement b[***]	17	totally agree - neither (dis)agree	totally agree - totally disagree	totally agree - totally disagree
statement c[***]	17	totally agree - neither (dis)agree	totally agree - disagree	totally agree – disagree
statement d[***]	17	totally agree - agree	totally agree - totally disagree	totally agree - totally disagree
statement e[***]	17	totally agree - agree	totally agree - neither (dis)agree	totally agree - neither (dis)agree
Dutch lessons (yes/no)	21	6/2	25/10	31/12
duration Dutch lessons[***]	23	2 weeks - 3 years	3 lessons - 10 years	3 lessons - 10 years
level of education[***]	25	VWO-PhD	MBO-PhD	MBO-PhD

VWO =   secondary education preparing for university

MBO =   intermediate professional education

[*]   For one participant in this group, handedness could not be determined

[**]   Only participants who had lived in another country were considered

[***]   For these variables only the range is presented

Countries where the L2 speakers were born:

<u>Within NS range</u>
France, Germany, Turkey

<u>Other L2 participants</u>
Belgium, France, Germany, Netherlands[23], Switserland, Turkey

---

[23]   One participant was born in the Netherlands, but had only lived there for a few months before the age of twelve.

Other countries where the L2 speakers had lived:

<u>Within NS range</u>
Belgium, Germany, USA

<u>Other L2 participants</u>
Belgium, Cameroon, Egypt, Germany, Greece, Guadaloupe, Russia, Syria, Turkey, UK, USA

Languages that were most important (for the participants) in the other countries:

<u>Within NS range</u>
Dutch, English, German

<u>Other L2 participants</u>
Arabic, Aramese, Dutch, English, French, German, Russian, Turkish

Other languages (than Dutch and the L1) that the L2 speakers spoke:

<u>Within NS range</u>
English, French, German, Italian, Spanish

<u>Other L2 participants</u>
Arabic, English, French, German, Greek, Italian, Russian, Spanish, Swedish

Situations in which the L2 speakers used other languages:

<u>Within NS range</u>
at home, at work, at school, for study, with a boyfriend, with colleagues abroad, with friends, during holidays or short trips

<u>Other L2 participants</u>
with partner, during residence in other country, at work, at school, for study, with relatives (in law), at conferences, with friends, during holidays or short trips, with foreigners, in a pizzeria

Situations in which the L2 speakers used Dutch:

<u>Within NS range</u>
with everyone in the Netherlands, at home, everywhere except at home,
at work, with relatives (in law), with (Dutch) friends

<u>Other L2 participants</u>
with everyone in the Netherlands, with everyone except native speakers of
one's L1, with everyone except one child, at home, with (one of) the
children, with the partner, everywhere except at home, at work, with
relatives (in law), with (Dutch) friends, with acquaintances, in the street, in
shops, at meetings, in institutions

Situations in which the L2 speakers used their L1:

<u>Within NS range</u>
at home, with partner, at work, with relatives, with friends with the same L1,
in the country of birth

<u>Other L2 participants</u>
at home, with partner, with (one of the) children, at work, at home when
talking about feelings, with relatives, with everyone with the same L1, with
friends with the same L1, with native speakers of the L1 who don't speak
Dutch, when writing poems, (colleagues) in the country of birth (during
holidays), on the telephone, "If there is no other possibility"

Countries in which the L2 speakers attended Dutch lessons:

<u>Within NS range</u>
Belgium, France, Germany, Netherlands

<u>Other L2 participants</u>
Belgium, France, Germany, Netherlands, Switserland

Institutions in which the L2 speakers attended Dutch lessons:

<u>Within NS range</u>
university, private language school, "Internationale Schakelklas" (secondary
school for beginning learners from abroad), adult education

Other L2 participants
school for interpreting and translation (higher professional education), university, private language school, Berlitz-school, primary school, secondary school, European school, higher professional education, Dutch institute, private lessons, "Internationale Schakelklas" (secondary school for beginning learners from abroad), adult education

Type of Dutch course:

Within NS range
intensive course/language lab, preparation for state exam, general proficiency, speaking

Other L2 participants
higher professional education for interpreting and translation, undergraduate studies Dutch language and literature, Dutch studies, undergraduate studies, preparation for state exam, preparation for secondary school exam, secondary school, general proficiency (for advanced learners), spelling/ writing

Educational background of the L2 speakers:

Within NS range
higher professional education for interpreting in court, English language (and literature), French language and literature, teacher training college French, classics, psychology, social and economic history

Other L2 participants
higher professional education for interpreting and translation, general linguistics, lexicology, Dutch language and literature, English language and literature, French language and literature, German language and literature, Germanic studies, Turkish language and literature, language and culture studies, speech therapy, teacher training college (for secondary education; teaching in the L1), training college (for primary education), European school, building, chemistry, creative therapy, economics (and administration), environmental technology, history, lathe operator, management studies, mathematics, pedagogy, politics of the job market and staff policy, psychology, social service, sociology, theology

Country in which the L2 speakers took their highest level of education:

<u>Within NS range</u>
country of birth, Netherlands, United States

<u>Other L2 speakers</u>
country of birth, Netherlands, United States

Summary

In general it is true that the younger people start learning a language, the higher their level of proficiency in that language will ultimately be. People who start acquiring a new language after the age of twelve, therefore, usually do not reach the same level of proficiency in that language as native speakers do. According to the critical period hypothesis this is due to maturational changes in the brain before puberty, which make people less and less sensitive to language input. Because of this reduced sensitivity, a native-like level of proficiency should not be attainable after puberty. This hypothesis, applied to the domain of syntax, is the basis of the first research question in this dissertation:

- Are there any late second language learners who fall within the native speaker range in their command of grammatical constructions that are known to be very difficult for second language learners and which can only be acquired on the basis of the input?

This dissertation also contains an investigation into the relationship between a native-like level of proficiency (if attainable at all) on the one hand and the typological distance between the language pairs involved and the background characteristics of the participants on the other hand. This is expressed in the following research questions:

- How is the level attained in L2 grammar after the age of twelve related to the typological distance between the L1 and the L2?
- What are the input and background characteristics of late learners who perform within the native speaker range (if they exist)?

For pronunciation, there are a number of previous studies that have identified second language learners who could not be distinguished from native speakers (see e.g. Bongaerts, 1999). For morphosyntax, results have been less clear and more controversial. (compare for example Coppieters, 1987; Birdsong, 1992; Hyltenstam, 1992; Ioup e.a. 1994; White & Genesee, 1996). Moreover, there were methodological problems with many of these studies and little attention had been paid to the role of the mother tongue. In the study presented in this dissertation, these problems were addressed and the relation between proficiency level at the end state and differences between the L2 and the first languages involved was systematically investigated.

SUMMARY

In this dissertation, a study is presented in which 43 native speakers of German, French and Turkish participated, who arrived in the Netherlands after the age of twelve and who were highly proficient in Dutch. Their performance on two grammar tests was compared to the performance on the same tests of (highly educated) native speakers of Dutch.

To assess the (implicit) grammatical knowledge of these participants, their command of dummy subject constructions in Dutch was tested. In these constructions the logical subject is not in its normal syntactic position for semantic or pragmatic reasons. Instead, this position is occupied by *het, er* or *0.*

In our study, we distinguish three types of dummy subject constructions:
- (active) sentences with *er* or *0* in which the (logical) subject is a noun phrase (DP) (DP-type)
- active sentences with *er, het* or *0* and a sentential (logical) subject (AS-type)
- passive sentences with *er, het* or *0* and a sentential (logical) subject (PS-type)

The native speakers of Dutch in this study revealed a general preference pattern for each type (a preference for *het, er* and/or *0*). This pattern is disturbed by certain factors. Therefore, each type consists of two or three categories with different judgement patterns. Examples are presented in (1)-(6):

(1) Men beseft niet altijd dat *0* <u>een pinguïn</u> een vogel is. (DP-type, general pattern)
"One does not always realise that a penguin is a bird."
(2) Ik vind het vervelend dat *er* boven <u>een raam</u> open staat. (DP-type, non-specific subject in intransitive sentence)
"It bothers me that there is a window open upstairs."
(3) Meestal valt *het* niet mee <u>om kaartjes voor een concert te krijgen</u>. (AS-type, general pattern)
"Usually it is not easy to get tickets for a concert."
(4) Nu schiet *0* mij ineens te binnen <u>dat ik nog boodschappen moet doen</u>. (AS-type, change of state)
"Now it suddenly occurs to me that I still have to go out shopping."
(5) In de krant wordt *0* beweerd <u>dat hij dronken achter het stuur gezeten heeft</u>. (PS-type, general pattern)
"It is claimed in the newspaper that he was drunk while he was driving."

178

(6)   Door haar vrienden wordt *het* bewonderd <u>dat ze ook in moeilijke</u> <u>tijden vrolijk blijft</u>. (PS-type, dummy object in equivalent active sentence)
        "Her friends admire her for remaining cheerful, even in difficult times."

There were two important reasons for choosing dummy subject constructions. First, they are known to be very difficult to acquire for second language learners. Second, they are hardly covered in Dutch grammars and L2 text books. This means that learners, having no access to explicitly formulated rules, can only acquire these constructions on the basis of processing language input.

Tot test the participants' command of dummy subject constructions in Dutch, two tasks were used in this study: a sentence imitation task and a sentence preference task. In the sentence imitation task participants had to repeat orally presented sentences literally. It has turned out that participants often unconsciously change elements that are phonologically non-salient and ungrammatical (from the point of view of the participant). In the sentence preference task participants had to indicate on a scale which sentence of a minimal pair they preferred. We also gave all participants a questionnaire with questions about background characteristics, such as age of arrival in the Netherlands, self-reported proficiency in various languages, level of education and questions about usage of Dutch and the L1.

The results on the tasks described above show that there are second language learners in each L1 group who have reached a native level in L2 grammar after the age of twelve. On the sentence preference task there were eight second language learners who performed within the native speaker range: three native speakers of German, four native speakers of French and one native speaker of Turkish. For the sentence imitation task there were eleven second language learners who performed within the native speaker range: seven native speakers of German, three native speakers of French and one native speaker of Turkish. As can be seen from these results, the role of the typological distance between the L1 and Dutch seemed to be greater for the sentence imitation task than for the sentence preference task.
        A comparison of the learners who fell within the native speakers range on the sentence preference task according to our (strict) criteria with the other second language learners suggested that the role of factors such as input, attending Dutch classes and age of arrival (after the age of twelve) were rather limited. At the same time, there did seem to be a meaningful relation with level of education, proficiency in some other language and

pleasure in learning languages. In addition, it appeared that many participants within the native speaker range had a linguistic background.

On the basis of these results it was concluded that reaching a native level after the age of twelve is possible for constructions that are difficult to learn and for which no explicit knowledge is available. The results thus falsify the critical period hypothesis. We also established that reaching this level is even possible for second language learners with an L1 which is very different from the L2 (both typologically and with respect to the constructions investigated). It should be noted, though, that the people who reach a native level constitute only a small percentage of second language learners. One should, therefore, exercise caution and not have unrealistic expectations for the majority of second language learners.

Finally, it was argued that the results with respect to the background characteristics of the second language learners suggest that factors on which learners can exert most influence seem to play a rather limited role, while something like language aptitude or language awareness seems to play a more important role. It seems plausible that people with higher aptitude or language awareness should be better able to notice and process details in the form of the L2 input than average L2 learners. This might have contributed to their greater success in acquiring difficult constructions that are phonologically non-salient and do not contribute much to the meaning of a sentence, as is the case for dummy subject constructions.

Samenvatting

Over het algemeen geldt dat mensen een hoger eindniveau halen in een (nieuwe) taal naarmate ze jonger beginnen met het leren van die taal. Mensen die na hun twaalfde beginnen met het leren van een nieuwe taal, bereiken dan meestal ook niet meer hetzelfde niveau als moedertaalsprekers. Volgens de kritiekeperiodehypothese komt dit doordat er tot de puberteit biologische veranderingen plaatsvinden in de hersenen die ervoor zorgen dat mensen steeds minder gevoelig worden voor taalaanbod. Door die verminderde gevoeligheid zou na de puberteit een hoog niveau niet meer haalbaar zijn. Deze hypothese staat centraal in de eerste onderzoeksvraag van mijn proefschrift. Die luidt:
- Zijn er late tweedetaalleerders die binnen het bereik van de moedertaalsprekers vallen voor wat betreft hun beheersing van constructies die erom bekend staan dat ze erg moeilijk zijn voor tweedetaalleerders en die alleen verworven kunnen worden op basis van het taalaanbod?

Ook heb ik gekeken naar de relatie tussen dit niveau (als dat mogelijk is) enerzijds en taalafstand en achtergrondkenmerken anderzijds. Dit leidde tot de volgende onderzoeksvragen:
- Hoe is het niveau dat bereikt kan worden in de grammatica van een tweede taal na de leeftijd van twaalf jaar gerelateerd aan de typologische afstand tussen de eerste en de tweede taal?
- Wat zijn de kenmerken van het taalaanbod en de achtergrond van late leerders die binnen het bereik van de moedertaalsprekers vallen (als die er zijn)?

In eerder onderzoek zijn er al mensen gevonden die door beoordelaars niet van moedertaalsprekers onderscheiden konden worden op het gebied van uitspraak (zie bijvoorbeeld Bongaerts, 1999). Voor grammatica waren de resultaten echter onduidelijker en tegenstrijdiger (zie bijvoorbeeld Coppieters, 1987; Birdsong, 1992; Hyltenstam, 1992; Ioup e.a. 1994; White & Genesee, 1996). Bovendien waren er een aantal methodologische problemen met veel van deze studies en was er weinig aandacht besteed aan de rol van de moedertaal. In het onderzoek dat in dit proefschrift wordt beschreven, is geprobeerd om deze problemen aan te pakken en systematisch aandacht te besteden aan de relatie tussen het bereikte niveau in grammatica en verschillen tussen de moedertalen en de taal die wordt geleerd.

In dit proefschrift beschrijf ik een onderzoek dat ik heb gedaan met 43 moedertaalsprekers van het Duits, Frans en Turks die na hun twaalfde naar Nederland zijn gekomen en zo goed Nederlands spreken dat je - behalve eventueel aan hun accent en lidwoordfouten - (bijna) niet kunt merken dat het geen Nederlanders zijn. Ik heb het gedrag van deze proefpersonen op twee grammaticataken vergeleken met dat van (hoogopgeleide) moedertaalsprekers van het Nederlands.

Om de grammatica van de proefpersonen te beoordelen is gekeken naar de beheersing van constructies met vervangende subjecten. Dit zijn zinnen waarin het eigenlijke (logische) onderwerp om semantische of pragmatische redenen niet op zijn normale syntactische positie staat en waarin die positie wordt bezet door *het, er* of *0*. Deze definitie komt ongeveer overeen met die voor het voorlopig onderwerp, zij het dat *0* meestal niet als voorlopig onderwerp wordt beschouwd.

In ons onderzoek worden drie typen constructies met vervangende subjecten onderscheiden:

- (actieve) zinnen met *er* of *0* waarin het (logische) onderwerp een zelfstandignaamwoordsgroep (DP) is (DP-type)
- actieve zinnen met *er*, *het* of *0* en een bijzin als (logisch) onderwerp (AS-type)
- passieve zinnen met *er*, *het* of *0* en een bijzin als (logisch) onderwerp (PS-type)

De moedertaalsprekers van het Nederlands in dit onderzoek hadden voor ieder type een algemeen voorkeurspatroon (een voorkeur voor *het, er* en/of *0*). Dit patroon wordt doorbroken door bepaalde factoren. Daardoor bestaat ieder type uit twee of drie categorieën. Voorbeelden van zinnen van deze categorieën zijn:

(1)   Men beseft niet altijd dat *0* <u>een pinguïn</u> een vogel is. (DP-type, algemeen patroon)

(2)   Ik vind het vervelend dat *er* boven <u>een raam</u> open staat. (DP-type, non-specifiek subject in intransitieve zin)

(3)   Meestal valt *het* niet mee <u>om kaartjes voor een concert te krijgen</u>. (AS-type, algemeen patroon)

(4)   Nu schiet *0* mij ineens te binnen <u>dat ik nog boodschappen moet doen</u>. (AS-type, overgang van situatie/toestand)

(5)   In de krant wordt *0* beweerd <u>dat hij dronken achter het stuur gezeten heeft</u>. (PS-type, algemeen patroon)

(6)   Door haar vrienden wordt *het* bewonderd <u>dat ze ook in moeilijke tijden vrolijk blijft</u>. (PS-type, voorlopig lijdend voorwerp in equivalente actieve zin)

Er waren twee belangrijke redenen waarom we hebben gekozen voor constructies met vervangende subjecten. Ten eerste staan ze erom bekend dat ze erg moeilijk te verwerven zijn en tweede is er (bijna) geen informatie over te vinden in grammatica's. Dat betekent dat leerders deze constructies uitsluitend kunnen verwerven op basis van het taalaanbod en dat ze geen expliciete kennis hierover hebben.

Om de beheersing van vervangende subjecten van de proefpersonen te toetsen heb ik twee taken gebruikt: een imitatietaak en een zinsoordelentaak. In de imitatietaak moesten proefpersonen zinnen letterlijk nazeggen. Het is gebleken dat proefpersonen elementen die fonologisch niet opvallen en ongrammaticaal zijn - vanuit het perspectief van de proefpersoon - vaak onbewust veranderen. Ook hebben we een zinsoordelentaak gebruikt waarin proefpersonen steeds van twee zinnen op een schaal aan moesten geven welke ze beter vonden. Bovendien hebben we bij alle proefpersonen een vragenlijst afgenomen waarin hen werd gevraagd naar achtergrondkenmerken, zoals de leeftijd waarop ze naar Nederland waren gekomen, hun beheersing van verschillende talen, hun opleiding en de mate en situaties waarin ze hun moedertaal en Nederlands gebruiken.

Uit de resultaten van de proefpersonen op de bovengenoemde taken bleek dat er in alle moedertaalgroepen tweedetaalleerders waren die na hun twaalfde nog een moedertaalniveau voor vervangende subjecten hebben bereikt. Voor de zinsoordelentaak waren en acht tweedetaalleerders die binnen het bereik van de moedertaalsprekers van het Nederlands vielen: drie moedertaalsprekers van het Duits, vier van het Frans en één van het Turks. Voor de imitatietaak waren er elf tweedetaalleerders die binnen het bereik van de moedertaalsprekers vielen: zeven moedertaalsprekers van het Duits, drie van het Frans en één van het Turks. Voor de imitatietaak leek de rol van de typologische afstand tussen de moedertaal en de doeltaal dus groter dan voor de zinsoordelentaak.

Vergelijking van de achtergrondkenmerken van de proefpersonen die volgens onze (strenge) criteria een moedertaalniveau hadden bereikt met de overige proefpersonen suggereerde dat de rol van factoren als taalaanbod, het volgen van Nederlandse les en leeftijd bij aankomst in Nederland beperkt was. Tegelijkertijd leek er wel een verband te zijn met opleidingsniveau, beheersing van een andere taal en plezier in het leren van taal. Bovendien viel het op dat veel proefpersonen met een moedertaalniveau een taalkundige achtergrond hadden.

Op grond van deze resultaten heb ik geconcludeerd dat het bereiken van een moedertaalniveau (ver) na de leeftijd van twaalf jaar nog mogelijk is voor moeilijk leerbare constructies waar geen regels voor beschikbaar zijn en dat

er geen kritieke periode is voor de verwerving van grammatica door tweedetaalleerders. Bovendien heb ik vastgesteld dat deze mogelijkheid ook openstaat voor tweedetaalleerders met een moedertaal die sterk verschilt van de tweede taal. Hierbij heb ik de kanttekening gemaakt dat het om een zeer klein percentage gaat en gewaarschuwd tegen onrealistische verwachtingen ten aanzien van het merendeel van de tweedetaalleerders.

Tenslotte heb ik betoogd dat de resultaten met betrekking tot de achtergrondkenmerken van de proefpersonen in dit onderzoek suggereren dat factoren waar leerders veel invloed op hebben een beperkte rol lijken te spelen in mijn onderzoek, terwijl iets als taalaanleg of taalbewustzijn wel een belangrijke rol lijkt te spelen. Het lijkt aannemelijk dat mensen met meer taalaanleg of taalbewustzijn beter in staat zijn details in de vorm van het taalaanbod waar te nemen en te verwerken dan de meeste taalleerders. Hierdoor zouden ze regels voor moeilijke constructies die fonologisch onopvallend zijn en weinig bijdragen aan de betekenis, zoals constructies met vervangende subjecten, beter kunnen verwerven.

Curriculum vitae

Sonja van Boxtel was born in Rotterdam on 5 March 1975. In 1993 she received her gymnasium diploma from the C.S.G. Comenius in Capelle aan den IJssel and went to the University of Nijmegen to study English language and literature for a year. In 1998 she graduated in General Linguistics and in Applied Linguistics at the same university. After having worked at Cito for a few months as a test item developer, she was appointed by the University of Nijmegen to carry out the research project which resulted in this dissertation. From February to June 2005 she worked as an educationist for ISM.