The Twig and the Tree

The development of an early foreign language phonology: a dynamic approach

Maeike Kiers (1628046) & Hannah Rosing (1652001)

MA thesis
Department of Applied Linguistics
Faculty of Arts
Rijksuniversiteit Groningen

Supervisor:
Dr. W. M. Lowie

Second reader:
Prof. Dr. C. L. J. de Bot

June 30, 2010
Table of Contents

ACKNOWLEDGEMENTS ........................................................................................................ 6
LIST OF ABBREVIATIONS ............................................................................................... 7
ABSTRACT .......................................................................................................................... 8

CHAPTER 1 INTRODUCTION ........................................................................................ 9

CHAPTER 2 LITERATURE BACKGROUND .................................................................. 11

INTRODUCTION ............................................................................................................ 11
AGE AND THE ACQUISITION OF A SECOND LANGUAGE PHONOLOGY .................... 11
Introduction .................................................................................................................... 11
CPH-L2A pros and cons .................................................................................................... 12
Other explanations for the age effect .............................................................................. 14
CPH-L2A and the acquisition of an L2 sound system ....................................................... 16
Age-related effects: a DST perspective ......................................................................... 17
Summary ........................................................................................................................ 20

PERCEPTION AND PRODUCTION .............................................................................. 20
Introduction ..................................................................................................................... 20
Perception of an L2 phonology and four models on perception to explain the phonological age-related effects ............................................................................................ 21
L2 pronunciation and other factors ................................................................................. 24

ENGLISH AS A SECOND LANGUAGE PHONOLOGY FOR DUTCH LEARNERS ...... 25
Introduction ..................................................................................................................... 25
VOT ................................................................................................................................. 25
Vowels ............................................................................................................................ 29
Crosslinguistic influence ............................................................................................... 30

STATEMENT OF PURPOSE ....................................................................................... 32

CHAPTER 3 METHOD .................................................................................................. 33

PARTICIPANTS .............................................................................................................. 33
MATERIALS ..................................................................................................................... 34
PROCEDURE .................................................................................................................. 36
ANALYSES ...................................................................................................................... 37

CHAPTER 4 RESULTS ................................................................................................ 40

VOT RESULTS .............................................................................................................. 42
Izaline: Early learner ..................................................................................................... 42
Bram: Early learner................................................................. 48
Cora: Late learner................................................................. 53
Aaffien: Late learner............................................................. 59
Combined results.................................................................. 64
VOWELS RESULTS.............................................................. 70
Izaline: Early learner............................................................ 70
Bram: Early learner................................................................. 74
Cora: Late learner................................................................. 78
Aaffien: Late learner.............................................................. 81
Combined results.................................................................. 85

CHAPTER 5 DISCUSSION...................................................... 87
VOT...................................................................................... 87
VOWELS.............................................................................. 89

CHAPTER 6 CONCLUSION.................................................... 92
SUGGESTIONS FOR FURTHER RESEARCH.............................. 95

APPENDIXES........................................................................... 96
I – REMAINING FIGURES – VOT........................................... 96
II – REMAINING FIGURES – VOWELS................................. 107
III – VOWELS PERCEPTION TASK – PICTURES.................. 115
IV – VOWELS PERCEPTION TASK – RECORDING SETS........ 117
V – SCORE FORMS AX DISCRIMINATION TASK (VOT)........ 119
VI – SCORE FORMS VOWELS PERCEPTION TASK............... 123
VII – ENGLISH SHADOWING TASKS – RECORDING SETS.... 127
VIII – DUTCH SHADOWING TASK – RECORDING SETS........ 130
IX – PRODUCTION PICTURES (FOR PRACTICE AND FREE SPEECH) 131
X – PRAAT VOT-SCRIPT......................................................... 136
REFERENCES..................................................................... 137
Acknowledgements

First of all, we would like to thank our supervisor Dr Wander Lowie for guiding us through the process of setting up our experiment, analysing our data, and writing our thesis. We would also like to thank Dr Chris McCully and Dr Kevin Haines for recording our testing materials. Furthermore, we would like to thank our participants Cora Brand, Aaffien Bolt, Izaline Klamer, and Bram Veninga for their time and enthusiasm. In addition, we would like to thank Izaline’s and Bram’s parents for their time and their hospitality. We would also like to thank Sieuwke Reitsma for providing us with her AX-task, the PRAAT VOT-script, and the knowledge of working with PRAAT and VOT. Finally, we would like to thank Anke Kiers for her excellent Google skills.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASP</td>
<td>Automatic Selective Perception model</td>
</tr>
<tr>
<td>CLI</td>
<td>cross linguistic influence</td>
</tr>
<tr>
<td>CPH</td>
<td>critical period hypothesis</td>
</tr>
<tr>
<td>CPH/L2A</td>
<td>critical period hypothesis as applied to second language acquisition</td>
</tr>
<tr>
<td>DST</td>
<td>Dynamic Systems Theory</td>
</tr>
<tr>
<td>ERP</td>
<td>event-related potential</td>
</tr>
<tr>
<td>F1</td>
<td>formant 1</td>
</tr>
<tr>
<td>F2</td>
<td>formant 2</td>
</tr>
<tr>
<td>L1</td>
<td>First language</td>
</tr>
<tr>
<td>L2</td>
<td>Second language</td>
</tr>
<tr>
<td>L2A</td>
<td>Second language acquisition</td>
</tr>
<tr>
<td>ms</td>
<td>milliseconds</td>
</tr>
<tr>
<td>SLM</td>
<td>Speech Learning Model</td>
</tr>
<tr>
<td>SPRs</td>
<td>Selective Perceptual Routines</td>
</tr>
<tr>
<td>VOT</td>
<td>voice onset time</td>
</tr>
</tbody>
</table>
Abstract

It is generally believed that it is easier for early learners to learn a second language than it is for late learners, especially when it comes to the pronunciation of an L2. In addition, a relationship between perception and production is often assumed. Furthermore, the correct pronunciation of words in isolation is generally known to be acquired before words in context or in free speech. Our longitudinal study aims to provide insight into these issues from a DST perspective. Over an eight-week period, four beginning learners of English as a foreign language were tested twice a week. Two learners were adults over fifty, and two learners were around the age of five. The focus was on the development of their perception and production of voice onset time (VOT) and the trap/dress vowel; both mentioned in the literature as potentially problematic for Dutch learners of English. The results of our study showed the usefulness of a DST approach. DST argues that a language system is always changing and that change is preceded by variability. The early and late learners of our study did not show univocal results, but the early learners did show a larger amount of variability over the sessions as well as within the sessions. So even though our measurement period was limited and therefore no full patterns of development could be determined, the differences between the individual learners and the changing patterns towards the end clearly show the value of a process oriented dense longitudinal approach.
Chapter 1 Introduction

It is generally agreed upon that overall, early (or young) learners can more easily acquire a second language than older learners. One school of thought is that learners that are above a certain age have passed a ‘critical period’. Only during this period, the second language that is being acquired can develop to be nativelike. The other school of thought argues that the already established L1 influences the acquisition of the L2. So far, most research into this issue has been product oriented. Researchers compared learners at one moment in time, or when they had reached their ‘end state’ of L2 acquisition. Even though research has shown that age is an important factor in second language acquisition (SLA), many questions remain unanswered. We believe that a different approach might provide us with more insight into the age-issue. Instead of looking at the product, we believe we should focus on the process of language development, in line with a Dynamic Systems Theory (DST) approach to SLA.

From a DST point of view, which we adopted in our study, we can gain more insight from looking at the process of development of learning a second language. Looking at the data of the learning process itself can be of great help in discovering, for instance, the differences between young and late learners. According to DST, within the learning process, there are a number of interrelated factors that are constantly changing and constantly influencing each other. For each individual learner, the factors that are in interplay with each other differ, as well as the way in which they influence one another. Whereas most research considers variability to be ‘noise’, DST considers variability to be very important. Variability during the process of learning a second language is a signal that the learner’s language system is changing. Therefore, the variability that may occur within the collected data can give better insight into the learners’ learning process.

This study focuses on the acquisition of foreign language phonology by Dutch learners of English. By means of a longitudinal study we attempt to find an answer to the following research question: How does an early foreign language phonology develop over time? To answer this research question, we look at process of development of four Dutch learners of English, and to gain more insight into the age-issue, we included two early learners around the age of five, and to late learners with an age over fifty. In our study we will focus on the development of VOT and the trap/dress vowel contrast, both mentioned in the literature as potentially problematic for Dutch learners of English.
We will begin this paper with a literature background, in which we review previous research, introduce DST, and adopt a DST approach in discussing various topics. After that, we will explain more about our participants, materials, procedures and analyses, after which we will present our results. In the results we will first discuss the VOT results in relation to our three research questions, and then we will do the same for the trap/dress vowels results. In the discussion the results will be interpreted, and the conclusion serves to answer our research questions. However, the answers are not as straightforward as one might expect.
Chapter 2 Literature background

Introduction

In discussions about the acquisition of a second language in general, and of a second language phonology specifically, age is often considered as having an important role as far as the success of this learning process is concerned. The general thought is: the earlier the better. Related to this, is the idea that perception is a predictor of the quality of the foreign accent in production. It is thought that when we are young, we are still able to perceive sounds or sound contrasts from foreign languages that do not exist in our native language and that this ability disappears or declines as we get older. As a consequence of this loss or decline, we cannot perceive these unfamiliar sounds correctly, leading to incorrect pronunciation. With this study, we aimed to gain more insight into this assumed relation between perception and production and into the role of age-related factors in the acquisition process. In order to do this we traced the phonological development of English of two early and two late Dutch learners over an eight-week period. In our theoretical background we will discuss the age-issue, in which we will go into the critical period hypothesis as applied to second language acquisition and move on to discussing the age-issue from a Dynamic Systems Theory (DST) perspective. The next section will be about the assumed relation between perception and production in an L2 phonology, and the final section will discuss differences between the Dutch and English sound systems and difficulties for Dutch learners of English specifically. This final section will end with a discussion of crosslinguistic influence.

Age and the acquisition of a second language phonology

Introduction

Every child can acquire a first language if the child gets the opportunity to do so. Furthermore, a child can also acquire a second or foreign language with relative ease. When it comes to adults, however, acquiring a first, second, or foreign language is not as obvious: Acquiring a first language after puberty is often thought to be impossible (especially from a Chomskyan point of view), and acquiring a second or foreign language clearly requires greater effort from an adult than from a child. A causal relation between age and language acquisition success is generally assumed, and as explanation for the major differences between children and adults,
and it is generally accepted that after puberty, it is no longer possible to fully attain a first language. A popular hypothesis in relation to this issue is the Critical period Hypothesis (CPH). This hypothesis, which assumes brain maturation to have significant negative effect on language acquisition, was developed by Lenneberg in 1967(1967).

The idea behind Lenneberg’s hypothesis was that neural plasticity ended at puberty, after which it became impossible to acquire a first language. According to Lenneberg, when neural plasticity had ended, the ‘window of opportunity’ had been closed (Ioup, 2008). Lenneberg developed this hypothesis in relation to L1 acquisition. However, this same hypothesis, or variations of this hypothesis, has also been proposed in relation to acquiring an L2 (for an overview of variations of the CPH see (Birdsong, 1999). In this chapter we will first state the different views on the CPH-L2A (the CPH as applied to L2 acquisition) and CPH-L2A research in relation to L2 acquisition in general and to the acquisition of an L2 phonology in particular, pay attention to other explanations, and finally make my claim that we should not be looking at the end product as has been done in most research up till now, nor focus on solely a biological, neural explanation, but rather look at the process of language development from a DST perspective and consider a combination of various factors in order to better understand the ‘age-issue’.

**CPH-L2A pros and cons**

As can be concluded from Birdsong’s (1999) CPH-L2A overview, second language acquisition critical period hypotheses assume a biological, neural cause to explain the experience that children acquire a second language with more ease than adults do. However, as different studies come with different results, there seems to be no clear cut-off point after which the acquisition of an L2 becomes more problematic, but rather a continuous decline. A study in support of the CPH-L2A is the one by Johnson and Newport (1989), which showed a linear decline in performance after the age of arrival of seven, and a random distribution of performance after the age of arrival of seventeen. However, re-analysing the results and a replication of the study showed very different results, contradicting the CPH-L2A and supporting the claim that age of arrival predicts success (Bialystok & Hakuta, 1994; Johnson & Newport, 1989); (Bialystok & Molis, 1998). Another, very different study which can be cited

---

1 Lenneberg does mention L2 acquisition briefly, stating that an adult can learn a second language. He also clearly states that this does not contradicts his CPH, as he assumes that “the cerebral organization for language
as supportive of the CPH-L2A, is the study by (Weber-Fox & Neville, 1999). Using ERP’s they found that late learners show slower linguistic processing and also different locus and function when it comes to language-related neural systems. However, opponents could argue that there is no hard evidence for a causal relation between these neural differences and age, and between these neural differences and L2 acquisition success.

The main argument against the CPH-L2A is that nativelike attainment is possible by adults. Birdsong (1999), for example, states that to him, the major counter evidence was his own research from 1992, in which adults performed within the native norm. Long, in his (1990) article, already stated that one adult learner with nativelike proficiency would suffice to falsify the CPH-L2A. Since then, various studies found nativelike proficiency among adult learners (Van Wuijtswinkel, 1994; White & Genesee, 1996; Bongaerts, 1999; Ioup et al., 1994; Bialystok, 1997; Birsdsong, 1992). In addition, in his 1992 study, Birsdsong also found that age of arrival had been of influence, even though all participants were adults. Furthermore, various studies showed that the age of arrival is a precursor of success, also after the end of maturation (Birdsong, 1992; Bialystok & Hakuta, 1999; Flege, 1999). Clearly, the results that show that the age effect continues after the end of maturation, contradict the CPH-L2A. Birdsong (1999) is careful to state, however, that the effect of age of arrival can be the result of different factors before and after the end of maturation. He does not exclude the possibility that developmental factors cause the effect up till the end of maturation, and non-developmental factors, such as exogenous factors, attitude, and motivation thereafter. We should not here, however, that this end point of maturation is again not clear. Studies with different results define different ages as the end point of maturation. In our opinion, however, their might not be a clear cut-off point and the ‘end of maturation’ does not exist as one age for all people. It is more likely that there is a continuum, in which all these factors mentioned by Birdsong interact with each other over a continuum, interlinked with and inseparable with developmental factors.

In his 2006 overview, Birdsong further explores the possibility of Maturationaly Based Critical Period effects, and concludes the explanations do not fit research findings. In his overview, Birdsong (2006) starts by stating that various studies have shown a negative correlation and linear function between age of acquisition and success in L2 acquisition. As an explanation for the linear function between age of acquisition and L2 acquisition success, Birdsong (2006) considers a Maturationaly Based Critical Period, in which age effects are considered to be critical period effects. He proposes three possibilities, namely the stretched learning as such has taken place during childhood, and since natural languages tend to resemble one another in many fundamental ways […] the matrix for language skills is present” (Lenneberg, 1967, p.176).
“Z” shape, the stretched “L” shape, and the stretched “7” shape. The first one was first proposed by Johnson and Newport (1989) and expanded upon by Pinker (1994). The stretched “Z” shape shows two flat periods, the first one a period where age has no effect yet and the last one where age no longer has effect on L2 acquisition. The point where the declining line changes into the second flat period is the end of maturation. Birdsong (2006), however, states that no data supports the second flat period. The stretched “L” shape has the same flat period starting at the end of maturation, for which no evidence exists. Different from the “Z” shape, it has no flat period in which age has no effect at all yet. Finally, the stretched “7” starts with a flat period, followed by a downward slope. The flat period represents the “window of opportunity”, a period in which full attainment is expected. Evidence for the “L” shape comes from a study by Birdsong and Molis (2001). However, the study did not show a decline at the end of maturation, but later, around the age of 27.5 years. This means that even though there is some evidence for a period of peak performance, evidence on maturational-effects is not consistent as results do not show a corresponding cut-off point.

Other explanations for the age effect

DeKeyser and Larson-Hall (2005), besides the main argument of nativelike attainment by adults, mention a number of other counterarguments as well, interpreting the age effect by environmental variables instead of assuming a maturational interpretation. The first of those counter arguments is input and practice, targeting the fact that adults and children receive different input, both quantitatively and qualitatively. DeKeyser and Larson-Hall (2005), however, state that the age effect cannot entirely be explained by this. One of their reasons is that length of residence has proven not to be a predictor of L2 success (for example: (DeKeyser, 2000); Flege, 1999). In addition, they state, even if a correlation is to be found, this does not necessarily prove a causal relation. Important are also the findings by, for example, Yeni-Komshian, Flege, and Liu (2000), which hint at a trade-off between L1 and L2 proficiency. The idea behind it is that high language proficiency requires a lot of practice. If a learner speaks his or her L1 more often than his or her L2 he or she receives less practice than when this is the other way around. Correlations between proficiency and use of the L1 have been found by various studies, among which the one by Yeni-Komshian, Flege, and Liu (2000). Yet again, according to DeKeyser and Larson-Hall (2005), the correlation is not enough to explain the age effect.
Another counter argument discussed by DeKeyser and Larson-Hall (2005) is social-psychological variables. Motivation, self-consciousness, attitude to and identification with the L2 culture also predict L2 success. However, these variables were never found to be significant. Therefore, they state, they cannot fully account for the age effect. We believe, however, that because these factors are often confounded with and inseparable from age, the interaction of all these variables together causes the ‘age effect’ and thus, together, account for the ‘age-effect’ (see ‘Age-related effects: a DST perspective’ below for a more extensive discussion). As a third counterargument they mention maturation without a critical period (DeKeyser and Larson-Hall, 2005). This arguments targets the idea that if age effects continue after the end of maturation, this argues against the CPH-L2A. The final counter argument mentioned is lack of qualitative differences, assuming that differences in learning mechanisms are reflected in the age effect (DeKeyser and Larson-Hall, 2005). Little research has been done to compare the acquisition patterns of children and adults. There has been, however, some research into language aptitude by DeKeyser (2000) and Harley and Hart (1997; Yeni-Komshian et al., 2000). Results showed that verbal aptitude (DeKeyser, 2000) and aptitude in general were the best predictors for adults, and memory for children (Harley and Hart, 1997). These results clearly illustrate the usefulness of longitudinal process oriented research.

Bialystok and Hakuta (1999) consider the possibility that cognitive and linguistic factors play a role in L2 acquisition and interact with age, much like our view that different factors interact together and are inseparable from age. They do not accept the CPH-L2A, and even state that a causal relation between age and success should not be assumed without strong evidence. Research results should be carefully interpreted. They state that the CPH-L2A assumes a causal relationship between maturational changes in the brain and L2 acquisition success, without strong evidence.

Marinova-Todd, Marshall, and Snow (2000) also reject the CPH-L2A, but not in the first place by stating counter-evidence or arguments. Rather, they go as far as to say that all arguments in favour of CPH-L2A are subject to either misinterpretation, misattribution, or misemphasis. They even question the claim that early learners are better at acquiring a second language than adults, assigning the claim to misinterpretation. They state that various studies (Rivera, 1998; Snow & Hoefnagel-Hohle, 1977; Snow & Hoefnagel-Hohle, 1978) have shown adults to be (initially) faster and more efficient in learning a second language than children. They also state that research into neural localisation, like the one by Weber-Fox and Neville

---

2 They do state that there are many factors involved in L2 success, and expand on the role of environment and motivation. This, also, fits our view of interacting and interrelated variables.
(1999) or Kim et al. (1997; Yeni-Komshian et al., 2000) are subject to misattribution. To support their claim, Mariniva-Todd et al. mention studies by Wuillemin and Richardson (1994) and Furtado and Webster (1991). By misemphasis, Marinova-Todd et al. refer to the emphasis placed on unsuccessful L2 adult learners, whereas, in their opinion, successful L2 adult learners are often ignored. They state that whereas early learners perform alike, there is more individual variation among adult learners. Unfortunately, researchers often provide average scores per age group, resulting in a misleading picture of adult learners. They do, however, also acknowledge that some studies do focus on the fact that adults in their studies did obtain native like results.

**CPH-L2A and the acquisition of an L2 sound system**

Research has also been done on the effect of age on the acquisition of an L2 phonology specifically. Disagreement exists on the causes of a foreign accent, and, again, on the nature of the relation between accent and age (Flege, 1999). In this discussion too, opinions on whether or not the critical period exists for a second language is the core issue. In relation to phonology, the CPH-L2A implies that after a certain age, certain aspects of the capacity to learn to speak an L2 without a foreign accent are lost (Flege, 1999). Opinions on what this certain age is differ. Long(1990) states that before the age of six, a language can be learned without a foreign accent. After the age of twelve, a foreign accent will be acquired, and in between the ages of six and twelve it can go either way. He based his conclusions on a review of studies previously published. Patkowski (1990), believed the critical period to end a little later, namely at the age of fifteen. Again the results from different studies do not show agreement on a clear cut-off point.

Various studies, however, have shown that a nativelike accent is possible at an age after the supposed critical period. Flege, Munro, and MacKay (1995) did not find a discontinuity at any age, but rather a more or less straight decline. Over 70% of the variance in the data was accounted for by age, 15% by language use factors. Both findings argue against a CPH in relation to foreign accent, as there was no discontinuity at a certain age, and if accent were caused by a critical period, language use should be of no influence. Building on Flege’s statement that other factors that may also influence the degree of foreign accent are confounded with chronological age (1999; also see below), we go even further by stating that the factors age and language use are not necessarily two things. So far, it is unclear what the ‘age- factor’ exactly entails. The ‘age-factor’ is (or can be) interrelated and confounded with
many other factors, such as language use. In the 70% accounted for by age, age is thus, in our opinion, not one simple factor, but a complex factor that interacts with other factors, which are confounded and interrelated with age: age-related factors. Similar results, namely a near-linear relation between foreign accent and age of acquisition, were obtained in a study by Yeni-Komshian, Flege, and Liu (1997). In his 1999 article, Bongaerts (1999) reported on three of his studies, in which he aimed to find out whether or not a nativelike accent was possible for adult learners. In all three studies at least a few non-native participants were judged to be nativelike in their accents.

In his 1999 article, Flege expresses his doubts concerning the CPH and explains three alternative hypotheses to explain the age effect. Flege (1999) states that a problem with the CPH is that one cannot directly test the CPH, as other factors that may also influence the degree of foreign accent are confounded with chronological age. The first hypothesis Flege (1999) discusses as an alternative for the CPH, is the exercise hypothesis, which assumes that speech abilities remain intact as long as one does not interrupt learning speech. The second hypothesis mentioned by Flege (1999) is the unfolding hypothesis, which states that a foreign accent is not the result of some sort of loss, but rather an indirect consequence of previous phonetic development. The third alternative mentioned by Flege (1999) – which does not necessarily exclude hypothesis two – is the interaction hypothesis, which proposes that bilinguals cannot fully separate the phonetic systems of their L1 and L2, as they inevitably interact. This last hypothesis is supported by research results from the study by Flege, Frieda, and Nozawa (1997), where the participants using there L1 a lot had a stronger English foreign accent than the other group, which used their L1 very little. Another study mentioned by Flege (1999) which supports the third hypothesis, is the earlier mentioned study by Yeni-Komshian et al. (1997).

*Age-related effects: a DST perspective*

Earlier we mentioned that DeKeyser and Larson-Hall (2005), Yeni-Komshian et al. (2001), and Bialystok and Hakuta (1999) offer different explanations for the age effects than the CPH-L2A. Flege (1999) also questions the CPH in relation to L2 phonology. These researchers mention things like differences in input, a trade-off between L1 and L2 proficiency, social-psychological variables, differences in learning mechanisms, and cognitive and linguistic factors. However, according to these researchers, none of the these factors can account for the age effect on its own, even though they might add to it or might be confounded with age. The
only significant factor that can, they state, is age, bringing us back to the CPH-L2A, which we rejected. In the previous section we already stated that we believe that different factors interact, and that these factors are confounded and interrelated with age, and called them age-related factors. This approach to the ‘age-effect’ problem is in line with a Dynamic Systems Theory perspective (DST), which we believe might offer a solution.

DST starts from the idea that language, as a complex dynamic system, constantly changes over time as a result of self-organization (De Bot et al., 2007b; Van Geert, 2008; De Bot et al., 2007a). Different factors within the system interact with each other and with the environment, constantly changing the system, causing development in a non-linear way. This means that all those factors mentioned earlier, such as environmental factors, learner characteristics, cognitive factors, and biological change, all influence each other and are part of the system or its environment. The combination of all these dynamic factors causes the system to change. Most of these factors change as the learner gets older, building on Flege’s statement that factors other than age that may influence the degree of foreign accent are confounded with chronological age (1999). We can no longer speak of the ‘age effect’, but should rather speak of age-related issues or effects, as different factors together cause the age-related effects.

In addition to replacing the concept of ‘the age effect’ with the idea of age-related issues, DST has one other aspect that differs greatly from most L2 age-related research up till now. Most research looks at the (end) product of L2 acquisition. Researchers are interested in the level of proficiency at a certain age, or even the end product of L2 acquisition in adulthood. Birdsong (2004) writes that research into the ultimate attainment of learners provides us with insight into the potential of the learner and the limits of second language acquisition. Even though he acknowledges that it is difficult to determine when a learner has reached the end state, as this decision requires a longitudinal approach whereas most research base there data on one observation only, he does argue that, in relation to the CPH-L2A, evidence to support the CPH-L2A has to come from learners’ end state.

However, from a DST perspective, no such thing as an end state or ultimate attainment exists (Lowie, 2010). A learners’ language system is dynamic and always changing. Therefore, a second language never stops developing. This immediately undermines the arguments in favour of the CPH-L2A that draw on evidence based on the ‘end state’ of learners’ L2. If there is no such thing as end state, one can never say for certain that a learner is unable to obtain native-like proficiency at some point in the future. Other arguments, basing their conclusions on observed critical period effects, are also refuted from a DST approach. From the DST approach, it logically follows that there is not one (biological) cause only for the observed
critical period effect, but rather that it is a combination of dynamically interacting factors leading to these results.

Even though, according to DST, there is no end state, DST does not deny that a learner can reach a certain point at which he or she does not seem to make progress anymore. According to DST, the system can have preferred states, which are called attractor states (Kelso, 1995). Attractor states can differ for every individual or group, explaining why not every learner shows the same states of stability. In addition, different attractor states can be caused by the interaction of a different combination of factors.

Attractor states are reminiscent of the concept of fossilization, yet there is a clear difference. Unlike fossilized language (which from a DST point of view, does not exist), an attractor state is not obligatory or permanent (see also (Thelen, 1995) on preferred states). An attractor is no more than a preferred state. With the right energy, the system can always get out of an attractor state. This energy could, for example, be the right amount of language practice or input, or maybe the awareness of incorrect language use and the subsequent effort to improve. It could also be another, stronger, attractor. Furthermore, an attractor is not necessarily non-native or inaccurate (Larsen-Freeman, 2006). The overuse of correct forms can also be seen as an attractor state. Larsen-Freeman mentions, for example, the overuse of relative clauses by Russians and Bulgarians (Todeva, 1992) to the prejudice of attributive infinitives. This can be seen as an attractor state for them, presumably caused by a combination of interrelated factors among which crosslinguistic influence from their L1.

This line of arguing brings us to one of the main principles of a DST approach. Instead of looking at the product of L2 acquisition, the DST approach looks at the process of L2 acquisition. What is interesting is the development of the second language over time, the process and its variability. Variability can provide us with an insight into the interaction of different factors over time, changing the system, influencing and determining the language development. An important aspect of DST is that in any complex system, inter-subject variability continuously occurs. The degree of stability of the system at any given moment is a good indicator of the level of variability that is taking place. When a relatively unstable period presents itself, this often means that the system is changing. DST considers growth as an iterative process; that is, the current level of development heavily depends on the previous level of development (Van Geert, 1994). Moreover, the initial state of a complex system can be

---

3 We do not state, however, that what others call fossilization, per definition excludes the overuse of correct forms, as this depends on one’s definition of the concept. We merely want to clarify the concept of an attractor state by pointing out that it does not solely included non-native or inaccurate language use.
of extreme importance. Even minor changes that occur at the beginning may have major effects in the long run. The notion of non-linearity is strongly related to this. Within a complex system, there is a non-linear relation between the initial degree of a small change and the effects it may have on a longer term. A DST approach is useful in that it provides tools and constructs that can make clear the interaction over time of various independent variables and account for non-linear development (Port & van Gelder, 1995; Thelen & Smith L.B., 1994; De Bot et al., 2005c; De Bot et al., 2005b; De Bot et al., 2007a; De Bot et al., 2007a; De Bot et al., 2007c). Applying this view to the age issue could offer new insights as we would be looking at the process of language development over time, rather than at the product at one moment in time.

Summary

In sum, in this section we discussed the possibility of a CPH-L2A, leading to the conclusion that no hard evidence for such a critical period exists. Then, we also looked at other factors to explain the observed ‘age effect’, which led us to a DST approach to second language acquisition. We renounced the idea of ‘the age effect’, and replaced this term by age-related issues or effects, as from a DST perspective the combination of various factors coinciding and interrelated with age are causing these age-related effects: the age-relates factors. In addition, we moved away from the idea that the acquisition of a second language can reach an end state. According to DST, language is a dynamic system that always keeps developing. Therefore, we should no longer focus on the product but rather look at the process of language acquisition. It is the process and its variability (the interaction of different factors), that can tell us more about second language development.

Perception and Production

Introduction

In the previous section, we discussed the age issue in relation to second language acquisition in general, and to phonology specifically. We argued that, instead of speaking of ‘the age effect’, we should talk about age-related issues, as we believe many different factors interrelated with each other and interrelated and confounded with age contribute to this observed ‘age effect’. When it comes to second language pronunciation, age of onset is again believed to determine
the quality of the L2 accent. Age is believed to influence perception, and perception, in its turn, is believed to be a major factor in determining the quality of the L2 accent. As explained previously, it is thought that when we are young, we are still able to perceive sounds or sound contrasts from foreign languages that do not exist in our native language and that this ability disappears or declines as we get older. As a consequence of this loss or decline, we cannot perceive these unfamiliar sounds correctly, leading to incorrect pronunciation. It is thus believed that imperfect perception - the inability to perceive contrasts between non-native sounds (well enough) -, causes difficulty in producing those sounds (Rochet, 1995). Studies that support this claim are, for example, the study done by Miyawaki, Strange, Verbrugge, Liberman, Jenkins, and Fujimura, 1975, investigating the ability of Japanese speakers to distinguish between the English /r/ and /l/ (1975). A number of studies did a similar research, showing similar results (see Ioup, 2008). In this section we will start by discussing the definition of perception and the development of perception in our first year of life. After that, we will discuss three models which aim to explain the underlying difference in perception between children and adults, and how this leads to accented pronunciation. The starting point of these models is not a CPH-L2A, but interference from the L1 phonological system with new sound categories in the L2. Finally, we will approach the issue from a DST perspective and also shortly mention other factors.

**Perception of an L2 phonology and four models on perception to explain the phonological age-related effects.**

In the discussion on L2 phonology, CPH states that biological changes caused by maturation of the brain make the learner unable to perceive and produce new sounds, leading to accented speech. This, however, is only one school of thought. The other school of thought also targets perception as a major factor in the pronunciation of a second language, but argues that the already established L1 categories influence the perception of new L2 categories (Ioup, 2008). There are several theories on or models for L1 interference or influence, when it comes to L2 phonology. However, before we start discussing these models, we will first have a look at the definition of perception and the development of perception in our first year of life.

Strange and Shafer (2008, p.159) define perception as follows: “Perception is, by definition, an internal mental (and physiological) process by which the perceiver recognizes incoming stimulus events as instances of mental categories’. These mental categories are based on the sounds of the native language and established very early in life. Research has shown that
babies up to about a year can discriminate between any sounds on every phonological level, including non-native sounds (see (Ohala, 2008). After a year, however, babies move from being language-general perceivers to being language-specific perceivers, as they can no longer discriminate between sounds of which the discrimination is unfamiliar to the native language (Strange & Shafer, 2008). They have established mental categories based on their native language. Logically follows that it is very difficult for adults to discriminate between non-native sound contrasts.

Strange and Shafer (2008) continue to discuss the issue of second language perception further. They first explain that the reason adults cannot distinguish between non-native contrasting sounds is due to automatic patterns of categorisation, and not to basic auditory capabilities. Namely, if memory load is minimised and tasks are practised well enough, people can discern between non-native contrasts. Strange and Dittmann (1984) tested Japanese speakers of English on the /l/-/r/ contrast. After a training course, the Japanese performed native-like in an identification task. However, the post-task showed results not much different from the pre-rest results. They concluded that if the memory load was bigger and they were not familiar with the materials, difficulty in categorisation remained (see Strange and Shafer, 2008, for more studies like the Strange and Dittman 1984 study). This supports the idea that automatic patterns of categorisation prevent adults from being able to discern between non-native contrasts. In her Automatic Selective Perception model (ASP), Strange (2006) calls these language-specific automatic perception patterns Selective Perceptual Routines (SPRs). The model assumes that beginning learners of an L2 use their L1 SPRs. Yet, because the basic auditory capabilities stay undamaged, L2 perception can improve with experience. Distinguishing between non-native contrast thus becomes a possibility, Strange states, yet the differentiation may be based on other parameters than the ones used by native listeners (Strange and Shafer, 2008).

Strange’s model has various aspects in common with several other models of speech perception, which also start from the idea of L1 interference. These models of speech perception aim to “characterise the nature of the underlying perceptual representations of L1 and L2 phonological categories” (Strange and Shafer, 2008, p. 170), to explain why perception of an L2 changes as age progresses. The first model we will discuss was developed by Best (1994), and is called the Perceptual Assimilation Model. The main assumption is that once L1 phonological categories have been established, non-native categories will be assimilated to these L1 categories, based on articulatory similarities. A problem occurs when the L2 contains a phonemic contrast which is non-existent in the native language. This contrast, then, will not
be perceived. Rather, the two sounds will be perceived as one sound. This will make it very difficult to establish separate categories.

A second model is the Native Language Magnet Model by Kuhl (1992). This model starts from the idea of phonetic prototypes. Ioup (2008) explains it as follows: “These prototypes are idealized representations of phonetic categories and act as anchors that interfere perceptually with the acquisition of non-native higher-level phonemic categories.” The prototypes are established before speech is categorized into phonemic units, and the perceptual sensitivity is reduced near the prototype’s distributional peak. If a new L2 sound is encountered that is similar to an L1 sound, the sound is perceived as the prototype, because the prototype functions as a magnet (Ioup, 2008). Ioup (2008) remarks that the model by Kuhl explains the changes in perception that happen by age one, but does not explain the decline in ability to correctly pronounce an L2 when age progresses after the age of one. Nor does it, in our view, explain how it is possible that some people do attain nativelike pronunciation in an L2.

A third model is Flege’s Speech Learning Model (Flege, 1995). Flege assumes that perception changes as one gets older, but that the mechanisms one needs to produce new sounds stay intact. He also argues that the ability to perceive new contrasting sounds decreases as one gets older, because the native categories become more fixed as we get older. Unlike the other two models, this model does explain why children would still be able to discern new sounds, and why there is a decline as we get older. Flege also states that sounds similar to the L1 are more difficult to master than sounds very different from the L1, as the difference is more easily noticed (Ioup, 2008).

To sum up, the general idea is that after we have established native phonological categories at the age of one, learning an L2 phonology becomes far more difficult. The L1 categorisation interferes, but opinions on why and how exactly differ. However, Flege’s SLM is the only one that explains why children are still better at learning an L2 phonology than adults, as he states that categories become more fixed as we get older.

Looking at the perception issue from a DST perspective and drawing on the other models, we can draw our own conclusions. Up to the age of one, children can distinguish between all sounds, as their phonological system is still an empty canvas. In their first year, however, phonological categories emerge. These categories will start working as attractors, when listening to sounds from another language. The older the child gets, the stronger these attractors get and the more energy is needed to avoid or get out of these attractors. This can be compared to Strange’s automatic Selective Perceptual Routines (2006) and Kuhl’s prototypes or anchors (1992). If learners are trained to perceive non-native contrasts, resources are made
available (energy) to avoid or get out of these attractors and thus, to perceive the contrasts. The conclusion by Strange and Dittman (1984) that if the memory load was bigger and they were not familiar with the materials, difficulty in categorisation remained, means that not enough energy was available to avoid or get out of the attractors. Flege’s comment that sounds similar to the L1 are more difficult to master than sounds very different from the L1 as the difference is more easily noticed, can also be explained. Similar sounds are drawn to the attractors, whereas there is no attractor available for sounds that are different from the L1 sounds. So again, it is not ‘age’ that influences perception, but it is the entrenchment of phonological categories (the building of attractors) which is a factor confounded with age: an age-related factor.

**L2 pronunciation and other factors**

We have established that age is not a simple factor causing the age-related effects, but that many factors confounded with age cause the age-related effects, together forming the age-related factors. When talking about L2 pronunciation, perception is believed to be influenced by these age-related factors, and in its turn, to be one of the factors to cause the age-related effects noticed in the quality of foreign language accents. However, there are other factors, not necessarily age-related factors, that can influence L2 pronunciation.

Moyer (2004), discusses individual variables that can affect the quality of a second language accent. From a DST approach, some of these variables would be viewed as age-related factors, others not necessarily. Still, they all interact with each other. Factors she mentions that show a correlation with the quality of an L2 accent are: cognitive variables, aptitude, motivation, length of residence, instruction/training, amount of L1 use, amount of native speaker input, and the phonological structure of the L1. Ioup (2008) also mentions other factors that proved to influence accent, such as amount of L2 use (Flege et al., 1999), target language input (Flege & Liu, 2001), instruction or training (Bongaerts et al., 1995; Elliot, 1995; Moyer, 1999), attitude (Moyer, 1999; Moyer, 2004; Purcell & Suter, 1980), and social identity (Hansen Edwards, 2008).

These factors and the interaction of these factors can differ from individual to individual. This explains the amount of individual variation Ioup writes about (2008). She mentions that there is a lot of individual variation when it comes to L2 pronunciation. Some adults excel and outperform children, and some adult-onset learners can even pass for native speakers. Ioup (2008) ends her article with directions for future research, in which she states
that what is needed are longitudinal studies comparing the learning processes of children and adults. She argues that this will give insight into processes used by learners, and by what the change from one stage to another stage is influenced. These directions are very much in line with the DST approach, and it is exactly what we aimed to do in our study. We believe that a longitudinal study comparing children and adults will shed light on the learning processes of both children and adults and on the differences between those processes.

**English as a second language phonology for Dutch learners**

*Introduction*

Dutch people who are learning English as a second language (L2) often experience trouble in the pronunciation of the language. This may be caused by the different stress patterns that English words in some cases have, or perhaps because of the intonation of the English words. Also, Dutch people often experience difficulties when it comes to the pronunciation of English speech sounds. All together, this leads to a kind of pronunciation that may be addressed as “Dutch-English”. That is, English words are pronounced in the way that Dutch words are normally pronounced. While both languages have descended from the Germanic language family, several phonetic elements clearly vary between the two languages. In other words, there are considerable differences between the laryngeal systems of English and Dutch phonology. In this chapter, two of these phonetic differences will be discussed. They include the voicing or aspiration of word-initial consonant stops and the vowel contrast /ɛ/ - /æ/. For both these features, Dutch people have to establish a new phonemic category. During the learning process, this will often lead to crosslinguistic influence between the L1 Dutch and the L2 English. The notion of crosslinguistic influence will also be dealt with in more detail in this chapter.

*VOT*

Both English and Dutch have a two-way laryngeal contrast in consonant stops, but the voice contrast is not the same. In English, the pronunciation of initial consonants can be either aspirated or unaspirated. In Dutch, consonants can be pronounced either voiceless and
unaspirated or voiced\(^4\) (Collins & Mees, 2003). This type of categorization of initial stops was first listed by Lisker and Abramson (Lisker & Ambramson, 1964), which they called Voice Onset Time (VOT). VOT is known to indicate the voiceless time interval between the release of the stop and the moment, or onset, at which the voicing starts (Ladefoged, 1975). Both the amount of prevoicing and the amount of aspiration can be referred to by using VOT values. The stops that are referred to in this chapter are stop consonants. In English, stop consonants can be voiceless, such as /p,t,k/ or voiced, such as /b,d,g/ (Zampini, 2008). Whereas voiceless stops are normally aspirated in English, they are unaspirated in Dutch. When it comes to voiced stops, they are usually phonetically voiceless in English, while in Dutch, they are often produced with prevoicing. Prevoicing, or voicing lead, points to vibration of the vocal cords before the stop is released (Rietveld & Van Heuven, 2001).

According to the average duration of the VOT, stops can be arranged into three categories: long-lag stops, short-lag stops and prevoiced stops. With long-lag stops, there is relatively much time between the release of the stop and the beginning of voicing (usually more than 35 milliseconds [ms.]) and aspiration is also audible. Short-lag stops only have short VOT durations (0-35 ms.) (Lisker & Ambramson, 1964). Both long-lag and short-lag stops have positive numbers. The moment that the obstructed airflow is being released, is the reference point, and is therefore equated with zero. With both these measurements of VOT, voicing happens after the release and thus, they are positive. With prevoiced stops, or lead stops, voicing already occurs before the release phase of the stop, which causes a negative VOT duration (with negative values) (Rietveld and van Heuven, 2001). Because Dutch contrasts between stops that are either prevoiced or unaspirated, it is often called a ‘voicing language’. English, with its contrast between short-lag and long-lag aspirated stops, is therefore called an ‘aspiring language’ (Simon, 2007). The three categories of VOT are illustrated below in a table, as well as the different VOT values that occur for Dutch and English /t/ and /d/.

<table>
<thead>
<tr>
<th>VOT</th>
<th>Lead (prevoiced)</th>
<th>Short-lag</th>
<th>Long-lag</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0 ms</td>
<td>0 – 35 ms</td>
<td>&gt; 35 ms</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Three categories of VOT values

\(^4\) Since VOT has to do with initial stop consonants, ‘voicing’ in this case actually refers to prevoicing in Dutch.
Table 2 VOT values for /d/ and /t/ in L1 Dutch and L1 English

<table>
<thead>
<tr>
<th></th>
<th>L1 Dutch</th>
<th>L1 English</th>
</tr>
</thead>
<tbody>
<tr>
<td>/d/</td>
<td>&lt; 0 ms</td>
<td>0 – 35 ms</td>
</tr>
<tr>
<td>/t/</td>
<td>0 – 35 ms</td>
<td>&gt; 35 ms</td>
</tr>
</tbody>
</table>

Table 2 VOT values for /d/ and /t/ in L1 Dutch and L1 English

In English, *voiceless stops* are aspirated when initially in a stressed syllable (Collins and Mees, 2003) and followed by a vowel. For example with the words ‘pen’ [pʰEn] and ‘toy’ [tʰɔi], the initial consonants are produced without voicing. In Dutch, on the other hand, voiceless stops are unaspirated, also in word-initial position. For instance, the initial consonant in the word ‘pet’ [pet] is not only produced without voicing, but also without aspiration. Though when it comes to *voiced stops*, in English they are usually phonetically voiceless in word-initial position. The /b/ in ‘bark’, for example, may therefore almost sound like a /p/, because there is no voicing involved. In Dutch, voiced stops are most often pronounced with prevoicing, or voicing lead (Rietveld and van Heuven, 2001). These VOT differences are clarified in the table below.

<table>
<thead>
<tr>
<th></th>
<th>/p/, /t/, /k/</th>
<th>/b/, /d/, /g/</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L1 Dutch</strong></td>
<td>- voiceless</td>
<td>- voiced</td>
</tr>
<tr>
<td></td>
<td>- unaspirated</td>
<td></td>
</tr>
<tr>
<td><strong>L1 English</strong></td>
<td>- voiceless</td>
<td>- voiceless</td>
</tr>
<tr>
<td></td>
<td>- aspirated</td>
<td></td>
</tr>
</tbody>
</table>

Table 3 The different characteristics for voiced and voiceless stops in L1 Dutch and L1 English

Although English and Dutch share the phoneme /t/ in their languages, English provides two ways of pronouncing it versus a single manner of pronunciation in Dutch. As a result, early Dutch learners of L2 English may not be aware of the difference between [t] and [tʰ]. According to Flege (Flege, 1990), this is due to “equivalence classification”. This mechanism would prevent learners from “noting acoustic phonetic differences” between Dutch [t] and English [tʰ]. Therefore, a phonemic category for English /t/ will not be established (Flege, 1990). The notion of equivalence classification is a central idea in Flege’s Speech Learning Model (SLM) of L2 acquisition (Flege, 1990). According to this model (Flege, 1990), a child whose native language is Dutch, for example, will be able to establish a separate phonetic category for English /t/ besides the one already established for Dutch /t/.
That is, if the child is first exposed to English around the age of 5 or 6 years. A native Dutch speaker who starts learning English at a later age is unlikely to establish such an extra phonemic category for “similar” L2 sounds (Flege, 1990). Thus, the SLM states that after phonemic categories in the L1 have been established in early childhood, learners will most probably identify L2 sounds as realizations of an L1 category, because they partly resemble the L1 sound. Whereas late learners will continue to identify such L1 and L2 sounds as being similar, early learners are more likely to eventually notice the phonetic differences between them. Therefore, early rather than late learners will establish separate phonetic categories for such similar L2 sounds and will also be able to produce them like native speakers. Although late learners are more likely to arrive in a so-called “attractor state” in which they are unable to identify differences between similar L1 and L2 sounds, from a DST point of view, their language acquisition process is still developing. Their age may interfere with the ease with of perceiving the L1 and L2 sounds as different, yet the interconnectedness and changing of several variables within their language system may cause them to notice the differences eventually. Thus while late learners may not notice the differences between the similar L1 and L2 sounds at first, there is a possibility that their language system develops in such a way that they do establish a new phonetic category for the L2 sound.

Flege (Flege, 1987) also states that especially adult L2 learners may produce English /p,t,k/ with noticeably shorter VOT values than English monolinguals, though with longer VOT values than monolingual native speakers of the learners’ L1. The VOT values that these learners produce are thus intermediate to the values observed for monolingual speakers of the L1 and the L2. If this is the case, learners are said to produce “compromise” VOT values in the L2 (Williams, 1980). Although compromise VOT values are a common characteristic for late L2 learners, Flege (1990) argues that English /p,t,k/ have also been produced with short-lag VOT values by some learners. These short-lag VOT values resemble those of the learners’ L1, which means that they produce English words with L1 sounds. This is probably also often the case with Dutch learners of L2 English, who fail to produce the correct VOT values for English words. As a result, their pronunciation comes across less native-like. It is even known that strength of foreign accent in sentences is inversely related to VOT in English /p,t,k/ (Flege & Eefting, 1987; Major, 1987), but this also depends on the learner’s L1. Also, VOT tends to move in different directions; using an L2 may also affect L1 VOT values.
Aspiration of initial stop consonants and the lack of prevoicing are two examples of new phonemic categories that Dutch learners have to acquire when learning English as an L2. Another instance of English speech sounds which are known to cause difficulty for Dutch learners to acquire is the vowel contrast of epsilon /ɛ/ and ash /æ/. These vowels occur, for instance, in the English words “dress” and “trap” respectively. In Dutch, the exact English dress and trap vowels do not exist, though the Dutch ‘e’ and the English /ɛ/ are very much alike. According to Broersma, a sound somewhere in between these two does, namely the vowel in the Dutch word “pet” (Broersma, 2005). The difference between the two English vowels is not very salient, which will make it more difficult for Dutch learners to perceive as well as to produce them as two different vowels. Moreover, in case of exposure to other varieties of English that have “shifted vowels”, such as South-African English or Australian English, the difficulty of perceiving and producing the vowels may be increased. They may be apt to produce the Dutch <e> in English words, rather than the English /ɛ/ or /æ/. Vowels can be partly phonologically arranged according to tongue height (e.g., high, mid, low) and frontness / backness (e.g., front, central, back). To a certain degree, “formant frequencies” acoustically reflect these characteristics of each vowel (Zampini, 2008). The resonance characteristics of the vocal tract determine the position of formants on the frequency-axis. The two principal formants that distinguish vowels are the first formant (F1) and the second formant (F2). It is generally understood that F1 increases as vowel height decreases, and F2 decreases as vowel backness increases (Ladefoged, 1975). In the table below, the average F1 and F2 values of English /ɛ/ and /æ/ are listed.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>F1 (Hz)</th>
<th>F2 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ɛ/ (dress)</td>
<td>1200</td>
<td>550</td>
</tr>
<tr>
<td>/æ/ (trap)</td>
<td>1000</td>
<td>700</td>
</tr>
</tbody>
</table>

Table 4 The frequencies of F1 and F2 (in Hz) of English vowels /ɛ/ and /æ/

Although this method works well for measuring the quality of clearly pronounced and steady vowels, it may be questionable whether spoken words in context will always be measured correctly. Such vowels may be pronounced less clearly and not as stable, which could
possibly influence its formant frequencies. The length of the vowels may also differ, which reflects vowel quantity. English distinguishes between long and short vowels. Moreover, the long-short pairs (such as /i/-/ɪ/) also differ in their formant frequencies (Zampini, 2008). As far as L2 speech research is concerned, Zampini (2008) argues that comparing measures of both formant frequencies and duration of articulation is a common methodological approach study the pronunciation of L2 vowels by learners.

**Crosslinguistic influence**

Grosjean (e.g. (Grosjean, 1982) argues that the two language systems of bilinguals will never be able to operate completely independently of one another. The reason for this – if such a thing as two separate language systems even exists – is that both language systems are always activated, at least to a certain degree. According to this view, no-one would ever be able to prevent pronunciation characteristics of the L1 from influencing their pronunciation of the L2 (Flege, 1990). But what if the two languages of bilinguals are not stored as two separate language systems in the brain? Most research today agrees upon the idea that there is a single lexical network for all languages with the individual items in the lexicon being tagged for the language to which they belong. These individual items are functionally grouped into language subsets, which are groups of lexical items that share a certain characteristic, such as ‘verb’ (De Bot et al., 2005a). Not only the subsets that belong to one language are connected, but also the subsets of one language are connected to those of the other language. When speaking in the L2, language subsets of the L1 may also be activated, as well as the other way around. The subsets are likely to change over time, which may be caused by a number of factors that are in their turn interrelated.

Many phonological studies are in line with this view; it is generally agreed upon that there is always to some extent interaction between early and later acquired languages, which is known as ‘crosslinguistic influence’ (Sharwood Smith, 1986). The term crosslinguistic influence (CLI) can be used to refer to other phenomena such as ‘transfer’, ‘interference’ and ‘borrowing’ (Sharwood Smith, 1986). Sharwood Smith (Sharwood Smith, 1994) argues that CLI does not refer directly to the influence that one language system may have on another. Rather, it refers to what goes on in a language user’s head. According to Sharwood Smith (1986), three relevant kinds of CLI exist. The first kind involves using some element of a previously acquired processing system to control knowledge belonging to the new,
developing system. The second kind of CLI is similar to the first kind, in that it also entails using old processing mechanisms to control a new linguistic system. The difference between these two kinds is that the first is habitual whereas the second may occur sporadically. With the third kind of CLI, L2 learners may utilize competence knowledge from the L1 in the acquisition of the L2. This division of the different types of CLI has been called the ‘competence/control model’ (Sharwood Smith, 1986). This model is further elaborated upon by Sharwood Smith, discussing the phenomenon of fossilization within this model. In a study by Krashen (Krashen, 1981), his subject was aware of the errors she made and was able to correct them. However, she was not yet moving towards the target norms, neither did she have acquired the structures in question. The subject could have been showing either ‘control fossilization’, which means that she could have been correcting be means of the already acquired competence in L2, or ‘competence fossilization’, which means that she was able to correct without having acquired the competence. From a DST perspective, fossilization may be regarded as an attractor state, as was discussed in the section ‘Age and the acquisition of a second language’ of this chapter. If both types of fossilization, control and competence, were to be ignored, Krashen’s subject could arguably have reached a preferred state of her L2 system, or an attractor state. As the L2 language acquisition process further develops, variability will again present itself, but will ultimately again settle into control, or another attractor state.

In contrast to what may seem reasonable, similar sounds are more often transferred to the L2 sound system than dissimilar sounds. This is because larger differences between languages are more easily noticed and thus acquired (Major, 2008). This has also been explained by Flege’s (1990) SLM. As was mentioned earlier, the central notion of this model – equivalence classification – explains why “equivalent” or similar sounds are more difficult to acquire. Whereas a learner will rather perceive a similar L2 sound as an equivalent in the L1, he or she will more easily notice salient differences between dissimilar sounds (Major, 2008). When this is approached from a Dynamic Systems Theory (DST) point of view, the notion of CLI could occur in quite a different way during the development of an L2 phonological system. From a DST perspective, the language system of any individual might be looked at as being dynamic. Within this system, there is complexity and interconnectedness between its components – either directly or indirectly. The system is constantly changing and it is self-organising (De Bot et al., 2005). As was introduced earlier, the process of second language acquisition (SLA) might also be viewed as a dynamic system, with the L1 being a part of the learner’s system as well. Interaction between the different
languages is likely to take place, which means that the L1 will probably become affected by learning the L2, as is evident from VOT observations (De Bot et al., 2005). In other words, during the process of L2 development, it is not just characteristics of the learners’ L1 that may influence, for instance, pronunciation of the L2, but also newly established categories of the L2 that may influence elements of that learners’ L1. In his linguistic interdependence hypothesis, Cummins predicts the same argument (Cummins, 1981). He expects the role of interdependence to be as follows: “To the extent that instruction in a certain language is effective in promoting proficiency in that language, transfer of this proficiency to another language will occur, provided there is adequate exposure to that other language (either in the school or environment) and adequate motivation to learn that language.” Furthermore, the interdependence hypothesis puts forward the idea that if there is sufficient input in one language and motivation to learn that language, this leads to better skills in that language as well as a facilitation of the transfer of several cognitive and academic language skills across languages (Sparks et al., 2009). It seems appropriate to say that not one, but many other factors together contribute to the level of pronunciation of a learner. Just as with age-related effects, for instance, perception is only one variable within the interconnectedness of a number of variables that causes the effects. Age itself is an important variable to determine the level of pronunciation of a second language phonology as well; we hope to find out whether this variable truly makes for the greatest difference between early and late learners.

**Statement of purpose**

As we stated earlier in this chapter, we have conducted a study from a DST perspective, meaning a longitudinal study in which the focus is on the process of language development. To gain more insight into the age-issue, the relation between perception and production, and crosslinguistic influence, we tested four participants over an eight-week period. All four participants were Dutch, and to include the age-issue, we tested two early learners (age five) and two late learners (over fifty). In the next section, we will discuss our methodology. We will specify the participants, testing materials, and procedures of our study. The method will be concluded by a section in which we will elaborate on the programs and analyses we used to analyze our data. After that, we will present our results, discuss and interpret them, and finally we will conclude our study by considering our results in relation to our expectations and previous literature.
Chapter 3 Method

In order to answer our research question (How does an early foreign language phonology develop over time?), we have formulated three sub questions. The first question asks whether it is better to start early to learn a language than it is to start late. The second question focuses on the assumed relationship between perception and production, and whether the perception of a foreign language develops before the production does. The third sub question goes into the expectation that a correct L2 pronunciation is first acquired for words in isolation, before this is the case for words in context or in free speech. (The words in isolation and context being elicited by means of a shadowing task, and the free speech by means of picture naming.) The idea behind the third sub question is that of limited resources (Lowie, 2010). Namely that more attention can be paid to pronunciation for words in isolation than for words in context, and more attention can be paid to pronunciation for words in context than for words in free speech. Therefore, it is to be expected that a correct pronunciation in isolation will precede a correct pronunciation in the other two conditions, and that a correct pronunciation on context will precede a correct pronunciation in free speech. From this it logically follows that a correct pronunciation on isolation might be viewed as a precursor to the right pronunciation in the other two conditions.

In order to answer these questions from a DST approach, we conducted four longitudinal case studies: two early learners around the age of five and two late learners over fifty. We tested our participants twice a week, which makes the data quite dense. During each session, several language tasks were conducted with each of the participants individually. All of their speech was recorded and in the end, the relevant recordings were analyzed and interpreted. In this section we will explain more about the design of our study, about our participants, our materials, the procedures, and finally the analyses we conducted.

2.1 Participants

Our study is based on the data retrieved from a total of four participants. All of them were native speakers of Dutch, and they all had a good sense of hearing. All participants were beginning learners of English. The late learners were two women of fifty-nine and seventy-one years old, Cora and Aaffien. Cora was taught English in secondary school, but did not practice any English after that. Aaffien did not receive any English at school, but had just started following beginners’ lessons in English at a community centre once a week at the start
of our experiment. The young learners were a boy and a girl around the age of five, Bram and Izaline. Bram was 5:3 and Izaline was 4:9 at the start of the experiment. They both received an English lesson at school once a week. They all participated in the experiment voluntarily.

2.2 Materials

For the experiment we used a total of four tests; two perception tests and two production tests. The perception tests consisted of a VOT discrimination task and a vowels perception test. One of the production tests was conducted to measure both the VOT and the vowels, the other production test measured elements of Dutch pronunciation, which was a control test. The VOT discrimination task that we used was an already existing test developed by Siewukke Reitsma (Reitsma, 2010). This task was developed to test categorical perception. Each of the 14 items that this test consisted of contained two sounds. With each item, the participants had to indicate whether the two sounds they heard were similar or different. The following sounds were included in this test: pa/ba, ta/da, ko/go. The /p/, /b/ and /k/ occurred alternately with and without aspiration, as well as with or without prevoicing. Prior to the 14 test items were two practice items, for which the correct answer was provided immediately after. Also, several fillers were added in between the regular test items. Four different versions of the same test were used, to avoid a learning effect. For all four versions, a scoring form was designed with the columns ‘same’ and ‘different’ per item. On this form, the researcher filled out the answer given by the participant (see appendix V).

For the dress-trap vowel perception task we designed four lists of words, a sheet with pictures of all those words, and a scoring form for each list of words (see appendix IV, III, and VI). We chose to alternate between four lists to avoid a learning effect. Each list contained 19 words in random order, consisting of six minimal pairs with the target vowels, six fillers (three minimal pairs), and one repetition of one of the words of the target minimal pairs. The fillers served to mask the targeted vowel distinction, and the repetition of one word served to make the participants aware that a word could occur twice: if they heard ‘man’ first, it did not necessarily mean that the second time they heard a similar word it had to be ‘men’. The lists did not contain a higher number of fillers and repeated words, because they could not be too long for the five-year-old participants. The four lists were quite similar to each other on the whole; however, the fillers and the repeated words were different in each list. All the words on the lists had to be comprehensible for five-year-olds. In addition, for each word we needed a picture, so we mainly had to look for content words, nouns, and non-abstract items.
All four lists were recorded in a studio by a native speaker of English. The picture sheet contained the pictures of all four lists plus a separate picture of a question mark, in case the participant did not know the answer. The sheet consisted of two A4 pieces of paper, with a total of 24 pictures. The four scoring forms contained the words of the four word lists, with a ‘right’ and ‘wrong’ column attached to them. Next to the ‘wrong’ column was room to write down an ‘s’ in case the mistake was semantic instead of phonological. In addition, we designed a Microsoft PowerPoint presentation and picture cards to practise the vocabulary used in this test, as it was necessary for the participants to know the meaning of the words.

The English production test consisted of a shadowing test of words in isolation and words in context (see appendix VII). As with the vowel perception test, we created four sets – all of them containing single words as well as words in context – which gave us the opportunity to switch between the sets to avoid a learning effect. In each set, 18 words were included in random order. Furthermore, each set contained three minimal pairs to test VOT, three minimal pairs with the target vowels, and three minimal pairs serving as fillers. A total of five filler sets was divided over the four sets for more variety. The same 18 words offered in isolation in each set, returned in very short sentence - or in context - in the same set. We chose to use sentences that were as easy as possible, because the participants had to reproduce them immediately after hearing them once. Also, we tried to place the target words in sentence-final position as much as possible, because in such short sentences, the stress is often on the final word. We hoped that the participants would not lose focus in pronouncing the sentences. Just like the words for the vowel discrimination task, all of these words and short sentences were recorded in a studio by a native speaker of English. We created powerpoint slides and picture cards in order for the participants to practice the words before they had to actually reproduce them. An overview of the kind of tests that were used is presented below in table 1. Several examples of minimal pairs that were used in the tests are listed in table 2.

<table>
<thead>
<tr>
<th>VOT</th>
<th>trap</th>
<th>dress</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perception</strong></td>
<td>- selection task: point to pictures</td>
<td>- discrimination task: AX-task</td>
</tr>
<tr>
<td><strong>Production</strong></td>
<td>- shadowing: words in isolation and context</td>
<td>- shadowing: words in isolation and context</td>
</tr>
<tr>
<td></td>
<td>- free speech: picture naming</td>
<td>- free speech: picture naming</td>
</tr>
</tbody>
</table>

Table 1 The different tasks that were used to test the perception and production of English VOT and trap / dress vowel distinction.
<table>
<thead>
<tr>
<th>Minimal Pair</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>pet / bet</td>
<td>VOT</td>
</tr>
<tr>
<td>toy / boy</td>
<td>VOT</td>
</tr>
<tr>
<td>marry / merry</td>
<td>vowels</td>
</tr>
<tr>
<td>bad / bed</td>
<td>vowels</td>
</tr>
</tbody>
</table>

Table 2 Examples of minimal pairs that were used in the production and perception tasks. The right column indicates whether the minimal pair aimed to test English VOT or trap / dress vowel distinction.

For the Dutch production test we only used one set of words in isolation plus these same words in context (see appendix VIII). In total, 10 words were included. They were to be subdivided in three minimal pairs to test Dutch VOT and four single fillers to minimize the attention on the minimal pairs. As with the English production test, the Dutch sentences were kept as short and easy as possible, with the target words in sentence-final position as much as possible. Of this test, two different versions were used to avoid predictability.

2.3 Procedure

Participants were tested approximately twice a week over an eight week period. They were tested in a quiet room. Cora, Bram, and Izaline were tested at home. During the first five weeks, Aaffien was tested at the community centre where she was also following her English lessons. During the last three weeks, the testing also took place at her home. The recordings and powerpoint slides used in the tests were presented and controlled using a laptop. Output from the participants was recorded using a Marantz Solid State audio recorder.

The sessions of the late learners slightly differed from the sessions of the young learners. Each session with the late learners started with going over the powerpoint presentations of the perception and production words of that week, to see if the meanings were known. Participants were asked to name the pictures on each powerpoint slide in English. Especially for the vowel perception test it was very important that the meaning of the words was known. In addition, this exercise elicited free speech. After this ‘powerpoint check’ a small ‘lesson’ took place, of which the theme varied from session to session. The lesson was always focused on speaking, and no reading or writing was involved. For instance, the participants would be asked to tell about their weekend, clothes or family. Then, the VOT discrimination test was conducted. After this test, the participants were asked to perform the English shadowing test, both for the single words as well as the words in context. The last
item of each session was the vowel perception test, for which the participants had to listen to words and point to the correct accompanying picture on the sheet that was placed in front of them. In the first week and in the last week, the participants also took the Dutch production control test.

The sessions of the young learners started by practising the words that would come up in the vowel perception test. This was done by using the picture cards. This was also used to elicit free speech. The children were asked to name the pictures or to point to the right picture. After practising the words, the children would do the VOT discrimination task. Participants were instructed to listen to two sounds, and tell the researcher if the sounds were the same or different. The researcher wrote down their answers on the appropriate scoring form. After the VOT test they would do the vowel perception test. During the test, the researcher would play the words of a set on the laptop, pausing after each word. The participants were instructed to point to the picture corresponding to the word they heard. The researcher wrote down their answers on the appropriate scoring form. After this some more practise games took place (using the picture cards, but also story telling, counting etc.). Finally, the participants would take the Dutch shadowing test and the English shadowing test. They were instructed to repeat the word or sentence that they heard. There was no time constraint. Each session lasted a minimum of thirty minutes and a maximum of fifty minutes.

2.4 Analyses

At the end of the eight weeks of testing, we had ended up with approximately four different recordings (of the different tests) per session of each individual participant. To be able to analyse the useful recordings, we started by cutting the recordings of the shadowing tasks per word. The cutting of the words was done in the programme Praat, version 5.1.01 (Boersma & Weenink, 2009). For each participant, we created a folder for each session in which we copied the shadowed words that were cut. There were separate folders for words in isolation, context, and free speech. The recordings were now ready to be analysed in Praat.

A Praat VOT-script (see appendix X; (Boersma & Weenink, 2009) was utilised to analyze the voice onset time of our participants. In this script, several things had to be filled out: the directory in which the sound files were saved, the name for the text file with the VOT results, and the name of the files that were to be analysed. For each shadowed word, the area from the release of the stop till the onset of voicing was selected. An example of a sound file with the VOT selected follows:
The selections of VOT that were made for each word were then transferred to a text file by the script. The results could now easily be copied to an Excel sheet. Since the results of the VOT were given in milliseconds by the script, in Excel they were multiplied by a thousand. Hereafter, we used Excel to calculate the average VOT value per session. Also, we computed the minimum and maximum VOT values over a moving window of three sessions. Furthermore, we created several graphs. First of all, we made a scatter diagram that showed the dispersion of VOT values per word and per session. Then we made a line graph of the average VOT values per session, in which we added a linear trendline. A min-max graph (Verspoor, 2008) of the already calculated minimum and maximum values was created as well. After we created these graphs for all of our four participants, we started comparing the early and late learners. This was done by pasting the graphs of the four participants on one another. We made comparison graphs for the average VOT values and for the min-max graphs. When it comes to perception, we made two graphs: one line graph that showed that number of right answers per session with a trendline in it and one line graph with the VOT perception scores per combination, also with trendlines.

For the analysis of the vowels, we did not use a script. We did try at first to work with a script that indicated vowel formants at the midpoint of an interval, though this appeared to be too time-consuming. Therefore, we decided to measure the formants manually, using PRAAT. First, we selected the appropriate vowel of a word and zoomed in, so that only the selected vowel was visible on the screen. Of this vowel, we selected the midpoint, after which
PRAAT could give a listing of the vowel formants. The figure below shows such a selected vowel, of which the midpoint is selected.

![Figure 2 Selection and midpoint of the vowel of ‘pan’ produced by a native Dutch speaker](image)

We copied the values of the first two formants to an Excel sheet, though we only used the F2 values to work with in the results section. The reason for this is that the difference between the F2 of the trap and the dress vowel is bigger than the difference between the F1 of these vowels. We calculated the average trap and dress F2 value per session by using Excel. Then, we computed the minimum and maximum trap and dress F2 values over a moving window of three sessions. We also created several graphs, just like we did with the VOT data. First of all, we made a scatter diagram that showed the dispersion of trap and dress F2 values per word and per session. Then we made a line graph of the average F2 values per session for both vowels, in which we added a linear trendline. A min-max graph of the already calculated minimum and maximum values was created as well. For each participant, we also compared words in isolation, context and free speech for the vowels individually in three kinds of graphs: dispersion, line graph with average values, and a min-max graph. As far as perception is concerned, we made four graphs. The first two were line graphs, one with the number of right answers per session and one with the number of wrong answers per session, both with a trendline. Another graph was a bar graph that showed the general perception of the vowels with the number of right and wrong answers per session. The trap / dress distinction was not included in this graph. The last graph was also a bar graph, which illustrated the number of trap and dress vowels that were guessed wrong per session.
Chapter 4 Results

In this chapter we will present the results in the following order. First, we will discuss the results in relation to the development of VOT. We will discuss each participant separately. For each participant, we will start with the Dutch control task. Next, we will present the results concerning the question whether production in isolation comes before production in context and free speech. Then, we will show the results concerning the question whether perception comes before production. Finally, we will compare the participants to each other, in relation to the final question whether it is better to start early than it is to start late. We will do the same for the trap/dress-vowels results, except that the Dutch control task does not apply here. For all graphs holds that the scales in have been adjusted to the values, which makes a direct comparison difficult, but clarifies the development within the graph.

Before presenting our results, we will first elaborate on the three sub questions a little further, to clarify what we will be looking for in our results.

**Question 1: Does production in isolation precede production in context and free speech?**

*In relation to VOT*

In the results relating to this research question, we will be looking for increase and/or variability in isolation that does not occur (yet) in context and free speech. If the values in isolation are overall higher than the values in context and free speech, it supports the idea that isolation precedes the other two conditions which might be due to limitation of resources. Furthermore, if there is an increase in VOT values in isolation but not in context and free speech, and the values in context and free speech are overall lower than the final values in isolation, this means that the learner is making progress in isolation, but not (yet) in the other conditions. In addition to looking for an increase, there is also a need to focus on variability. According to DST, variability is a precursor of change, so before there is an increase in a certain condition, we would expect a lot of variability in that condition. So even if there is no increase, variability already signals that the system is changing and moving out of an attractor state.

*In relation to the trap/dress vowels*

In the results relating to this research question, we will be looking for different F2 vowel values for trap and dress and/or variability in isolation that does not occur (yet) in context and
free speech. If the dress F2 values are higher than those of the trap vowel around the correct F2 value of our native speaker, it supports the idea that isolation precedes the other two conditions. It is important that the participants make a distinction between the two vowels, rather than producing them as if it were a single vowel (probably close to the Dutch ‘e’). Also, variability is an aspect to focus on, since this is regarded as a precursor of change.

**Question 2: Does perception precede production?**

*In relation to VOT*
In the results relating to this research question, we will be looking for an increase and/or variability in perception that does not occur (yet) in production. If there is a lot of variability in perception, but not in production, this supports the idea that perception precedes production, as variability is a precursor of change. Equally so, if there is an increase in perception, but not in production, it supports the idea that perception precedes production.

*In relation to the trap/dress vowels*
In the results relating to this research question, we will be looking for an increase and/or variability in perception that does not occur (yet) in production. Ideally, we would want the participants to perceive a difference between the trap and the dress vowel. If they do not perceive this difference (yet), they might perceive one vowel incorrectly more often than the other vowel. If there is variability or an increase in perception, but not in production, it supports the idea that perception precedes production.

**Question 3: Is it better to start early than it is to start late?**

*In relation to VOT*
In the results relating to this research question, we will be looking for an increase and/or variability in the results of the early learners that do not occur in the results of the late learners or vice versa.

*In relation to the trap/dress vowel*
In the results relating to this research question, we will be looking for a distinction between the two vowels and/or variability in the results of the early learners that do not occur in the results of the late learners or vice versa.
**VOT results**

**Izaline: early learner**

*Dutch control task*

Izaline’s average VOT value in Dutch at the start of the experiment is shown together with the average VOT value of Kevin, the recorded native speaker.

![Figure 1 Dutch VOT compared to native speaker's VOT at the start of the experiment](image)

**Question 1: Does production in isolation precede production in context and free speech?**

*Isolation*

The development of Izaline’s VOT production in the word repetition task, or ‘isolation’ in short, is shown in the graph below.

![Figure 2 Dispersion of the VOT values in isolation](image)
Figure 2 shows the dispersion of the VOT values in milliseconds for each word in each session. The graph shows that there is some variation over time and within sessions, with values that vary from a minimum of eight milliseconds to a maximum of forty-seven milliseconds. The amount of variation over time varies and overall there is no increase in VOT values over time (see appendix $, figures $ and $).

**Context**

The development of Izaline’s VOT production in the sentence repetition task, or ‘context’ for short, is shown in the graphs below.

Figure 3 shows the dispersion of the VOT values for each word in each session. There are three outliers with a VOT value exceeding 250 milliseconds. If we exclude these three outliers, like we have done in figure 4 the picture of the variation among the other values becomes clearer.

Figure 3 Dispersion of the VOT values in context

Figure 4 Dispersion of the VOT values in context without outliers
Again figure 4 shows that there are two outliers, one in session three and one in session eleven. The values vary from eleven milliseconds to seventy-three milliseconds. There is some variation over time and within sessions.

Figure 5 Line graph of the average VOT values per session

Figure 6 Line graph of the average VOT values per session without outliers

Figure 5 and 6 illustrate of the average VOT values per session. The trendline in figure 5 shows an increase in VOT over time. In figure 6, in which the three outliers were disregarded, there is no increase in VOT.
Figures 7 and 8 depict a min-max graph. In figure 7 variation seems to increase significantly over time. Figure 8, in which the three outliers were disregarded, is very different. Variation decreases in session five, increases after session five, decreases in session eight, and increases again after session nine.

*Free speech*

The development of Izaline’s VOT production in free speech is shown in the graphs below.
Figure 9 Dispersion of the VOT values in free speech

Figure 9 shows the dispersion of the VOT values for every word in every session. There is some variation over time and also within sessions, but not a lot as values vary from eight milliseconds to thirty milliseconds. Neither is there a significant increase in VOT values over time or in the amount of variation (see appendix I, figures 3 and 4).

Comparison isolation, context, and free speech

In the figures below the results were added together for comparison. We plotted the context graphs without the three major outliers.

Figure 10 Dispersion of the VOT values

Figure 10 shows that there is the least variation in free speech, and the most variation in context.
Figures 11 shows the average VOT values per session for all three conditions. The trendlines show that the VOT values do not significantly increase or decrease over time. It does show that, overall, the values in context are highest, followed by the values in isolation. The values in free speech are lowest. In all three conditions, most values do not exceed the 35 milliseconds (values below the 35 ms are normally not considered to be aspiration).

**Question 2: Does perception precede production?**

In the figures below the development of the discrimination of different VOT contrast is depicted.
In figure 13 the scores per combination per session are depicted. Again there is some variation. The combination prevoiced / aspirated, of which the contrast between the two sounds is biggest, is correctly perceived for five sessions in a row towards the end. However, in the final sessions it was perceived incorrectly again. The combination short / prevoiced, which is the Dutch contrast, was perceived correctly only twice; in session two and in session nine. The combination aspirated / short, which is the English contrast, was perceived correctly three times towards the end, in sessions six, nine, and ten.

**Bram: early learner**

*Dutch control task*

Bram’s average VOT value in Dutch at the start of the experiment is shown in the figure below, together with the average VOT value of Kevin, the recorded native speaker.
**Question 1: Does production in isolation precede production in context and free speech?**

**Isolation**

The development of Bram’s VOT production in isolation is shown in the figures below.

![Figure 15 Dispersion of the VOT values in isolation](image)

Figure 15 shows the dispersion of the VOT values for each word in each session. The graph shows a lot of variation over time, but also within sessions. The values vary from sixteen to sixty milliseconds.

![Figure 16 Line graph of the average VOT values per session](image)

Figure 16 depicts a line graph of the average VOT values per session. The trendline shows that there is a very slight increase in the average VOT value over time. The increase exceeds the 35 ms.
Figure 17 Min-max graph of the average VOT values per session

Figure 17 shows the amount of variation over time. Towards the end the amount of variation increases and the values seem to exceed the 35 ms.

Context

The development of Bram’s VOT production in context is shown in the figures below.

Figure 18 Dispersion of the VOT values in context

Figure 18 shows that there is quite some variation over time and within sessions. The values vary from eighteen milliseconds to seventy-nine milliseconds.
In figure 19 a line graph of the average VOT values per session is depicted. There was no data for session seven. The trendline shows a slight decrease in average VOT length.

**Free speech**

The development of Bram’s VOT production in free speech is shown in the figure below.

Figure 20 illustrates that over time, there has been a very slight increase in VOT length. However, there is very little variation over time, nor within session and none of the values exceeds the 35 ms (see appendix I, figure 10). The amount of variation over time does not vary much either (see appendix I, figure 11).

**Comparison isolation, context, and free speech**

In the figures below the results were added together for comparison.
Figure 21 shows that there is the least variation in free speech, and the most variation in context (see also appendix I, figures 13, 14 and 15).

Figures 22 shows the average VOT values per session for all three conditions. The trendlines show that the VOT values in isolation and free speech very slightly increase, whereas the values in context slightly decrease. It also shows that, overall, the values in context and isolation are highest. The values in free speech are lowest. Furthermore, isolation seems to show a delayed effect of the same pattern as context.

**Question 2: Does perception precede production?**

In the figures below the development of the perception of VOT is depicted.
Figure 23 Number of right answers per session

Figure 23 shows the number of right answers per session, three correct answers being the maximum possible score for each session. The figure shows that more correct answers seem to be given towards the end.

Figure 24 Scores per combination per session

In figures 24 shows the scores for the combination aspirated / short. The combinations prevocalized / aspirated and short / prevocalized showed no variation, but were perceived wrong every session except for the combination prevocalized / aspirated, which was perceived correctly in the first session. The combination aspirated / short does show some variation as it was perceived correctly four times, in sessions one, five, nine, and eleven.

Cora: Late learner

Dutch control task
Cora’s average VOT value in Dutch at the start of the experiment is shown together with the average VOT value of Kevin, the recorded native speaker.

![Figure 25 Dutch VOT compared to native speaker’s VOT at the start of the experiment](image)

**Question 1: Does production in isolation precede production in context and free speech?**

**Isolation**
The development of Cora’s VOT production in isolation is shown in the figures below.

![Figure 26 Line graph of the average VOT values per session](image)

Figure 26 depicts a line graph of the average VOT values per session. The figure shows that there is some variation over time, which seems to increase towards the end (see also appendix I, figure 19). The trendline shows that there is a slight increase in the average VOT value over time, in the end exceeding the 35 ms.

**Context**
The development of Cora’s VOT production in context is shown in the figure below.
Figure 27 Line graph of the average VOT values per session

Figure 27 depicts a line graph of the average VOT values per session. There is not much variation and the amount of variation does not seem to vary much either (see also appendix I, figures 20 and 21). The trendline shows that there is a very slight increase in the average VOT value over time, but stays below the 35 ms.

*Free speech*

The development of Cora’s VOT production in free speech is shown in the figure below.

Figure 28 shows the dispersion of the VOT values for every word in every session. There is quite some variation over time and also within sessions. Values vary from thirteen milliseconds to sixty-four milliseconds.
Figure 29 Line graph of the average VOT values per session

Figure 29 illustrates the average VOT values per session and shows a very different picture. There is not a lot of variation between the average values. The trendline shows no decrease or increase and values stay below 35 ms. The amount of variation does not vary a lot either (see appendix I, figure 22).

Comparison isolation, context, and free speech

In the figures below the results were added together for comparison.
Figures 30 and 31 show the average VOT values per session for all three conditions. The trendlines show that the VOT values in isolation slightly increase, and that the VOT values in context very slightly increase. The graphs also shows that all values for all conditions are close together, and the values in isolations show the most variation, especially towards the end.
Figure 33 Min-max graph of the average VOT values per session in context and free speech

Figure 34 Min-max graph of the average VOT values per session in free speech and isolation

Figures 32, 33 and 34 depict min-max graphs for the three conditions combined. Isolation clearly shows most variation, and the amount of variation increases towards the end.

**Question 2: Does perception precede production?**

In the figures below the development of the perception of VOT is depicted.
Figure 35 shows the number of right answers per session, three correct answers being the maximum possible score for each session. The graph shows some variation.

In figure 36 the scores per combination per session are depicted. The combinations prevoiced / aspirated and short / prevoiced sow not variation. The combination aspirated / short, which is the English contrast, does show variation as it was perceived correctly three times, in sessions four, eight, and twelve.

**Aaffien: Late learner**

*Dutch control task*

Aaffien’s average VOT value in Dutch at the start of the experiment is shown together with the average VOT value of Kevin, the recorded native speaker.
**Question 1: Does production in isolation precede production in context and free speech?**

**Isolation**

The development of Aaffien’s VOT production in isolation is shown in the figures below.

Figure 38 shows the dispersion of the VOT values for each word in each session. The graph shows quite some variation over time and within sessions. The values vary from a minimum of fourteen milliseconds to a maximum of ninety-three milliseconds. There are many values exceeding the 35 ms.
Figure 39 Line graph of the average VOT values per session

Figure 39 depicts a line graph of the average VOT values per session. The trendline shows a slight decrease in the average VOT value over time.

**Context**

The development of Aaffien’s VOT production in context is shown in the figures below.

Figure 40 Dispersion of the VOT values in context

Figure 40 shows the dispersion of the VOT values for each word in each session. The graph shows considerable variation over time and within sessions. The values vary from a minimum of seventeen milliseconds to a maximum of sixty-one milliseconds. Various values exceed the 35 ms. There is no significant increase or decrease over time in average values (see appendix I, figure 25).
Figure 41 Min-max graph of the average VOT values per session

Figure 41 shows the development of the amount of variation over time of the average values. There is not a lot of variation, but variation does seem to increase towards the end.

Free speech

The development of Aaffien’s VOT production in free speech is shown in the figures below.

Figure 42 Dispersion of the VOT values in free speech

Figure 42 shows the dispersion of the VOT values for every word in every session. There is quite some variation over time and also within sessions. Values vary from sixteen milliseconds to sixty-four milliseconds. There is no significant increase or decrease in average values per session (see appendix I, figure 26).
Figure 43 Min-max graph of the average VOT values per session

Figure 43 depicts a min-max graph. The graph shows little variation in the amount of variation. Variation also decreases after session six.

*Comparison isolation, context, and free speech*

All conditions show about the same amount of variation. In none of the conditions there is a significant increase or decrease (see appendix $, figures $ - $).

*Question 2: Does perception precede production?*

In the figures below the development of the perception of VOT is depicted.

Figure 44 Number of right answers per session

Figure 44 shows the number of right answers per session, three correct answers being the maximum possible score for each session. The graph shows that there is some variation in the number of correct answers.
In figures 45 and 46 the scores per combination per session are depicted. The combination prevoiced / aspirated had to be plotted separately for clarity. All combinations show quite some variation.

**Combined results**

*Question 3: Is it better to start early than it is to start late?*

In the figures below the four participants were compared to each other, to see if the early learners have an advantage over the late learners or vice versa.
Graphs 47, 48, and 49 show the dispersion of the VOT values in all three conditions for all four participants. In graph 47, the three major outliers in Izaline’s data were excluded. In graph 49 we see that Aaffien seems to show the most variation and Izaline the least. In graph
48 both early learners show most variation. In graph 49 the late learners show the most variation.

Figure 50 VOT values in isolation averages per session for all participants

Figure 51 Trendlines VOT values in isolation averages per session for all participants

Figures 50 and 51 show the average VOT values in isolation for all four participants. Bram and Aafien seem to have the highest values overall. Aaffien and Cora show the most variation in averages per session. In figure 51 the trendlines show that Cora shows the biggest increase, which is still only around 12 milliseconds. Aaffien, on the other hand, shows a decrease. Izaline is more or less stable, and Bram slightly increases.
Figures 52, 53, and 54 show the average VOT values in context for all four participants. Izaline’s outliers caused by three words in three separate sessions, are left out in figure 53 and
provide a clearer picture. Bram and Aafien seem to have the highest values overall. The early learners show the most variation. Bram shows a slight decrease.

Figures 55 and 56 show the average VOT values in free speech for all four participants. Aafien clearly has the highest VOT values and Izaline the lowest. Aafien and Izaline show most variation. The trendlines do not show any significant in- or decreases.

We can also compare the min-max graphs of the four participants to get a clearer picture on the amount of variation of the four participants. If we compare the min-max graphs of the values in isolation (appendix I, figure 2; figure 17; appendix I, figure 19; appendix I, figure 24) we see that both Bram and Cora show more variation towards the end. Aafien shows the least variation. The min-max graphs of values in context (figures 7 and 8; appendix I, figure 9; appendix I, figure 21; figure 41) show that Izaline and Aafien show more variation towards the end. Overall the early learners show more variation than the late learners. The
min-max graphs of values in free speech (appendix I, figure 4; appendix I, figure 11; appendix I, figure 22; figure 43) show that Izaline shows the most variation.

The graphs below show the perception scores of all four participants.

Figure 57 Perception scores for the combination prevoiced / aspirated for all participants

Figure 58 Perception scores for the combination short / prevoiced for all participants
Figure 59 Perception scores for the combination aspirated / short for all participants

Figure 57 depicts the results of the biggest contrast, namely prevoiced / aspirated. Cora gave the right answer almost every time, and Bram gave the wrong answer almost every time. Aaffien shows a lot of variation, and so does Izaline. Figure 58 depicts the scores on the Dutch contrast. Cora scores a hundred percent. Bram gave the wrong answer every time. Aaffien shows quite some variation, and a Izaline shows a little variation. Figure 59 depicts the scores on the English contrast. All learners show a similar amount of variation.

**Vowels results**

The F2 vowel formants of the trap and dress vowels in the table below are of the native English speaker that our participants shadowed. The F2 values of the Dutch ‘e’ are of the native Dutch speaker that our participants shadowed.

<table>
<thead>
<tr>
<th></th>
<th>Trap</th>
<th>Dress</th>
<th>Dutch ‘e’</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2 (Herz)</td>
<td>1506</td>
<td>1738</td>
<td>2071</td>
</tr>
</tbody>
</table>

Table 1 F2 vowel formants in Hz of the English trap and dress vowel and the Dutch ‘e’

**Izaline: early learner**

*Question 1: Does production in isolation precede production in context and free speech?*

**Isolation**

The development of Izaline’s vowel production in isolation is shown in the figure below.
In figure 60, the development of the amount of variability over time is portrayed. For the trap vowel, variability seems to increase after session four and to decrease again after session nine. For the dress vowel, variability seems to increase slightly after session six, to decrease after session seven, and to increase again after session eight. Overall, but especially after session five, the amount of variability of the trap vowel is higher than of the dress vowel. There is a slight increase in the average F2 vowel formants of the trap vowel (from 2250 Hz to 2500 Hz), but this does not happen for the dress vowel, which stays around 2500 Hz (see appendix II, figure 2).

**Context**

The development of Izaline’s vowel production in context is shown in the figures below.

In figure 61, the development of the amount of variability over time is portrayed. For the trap vowel, variability seems to increase after session four and to decrease again after session nine. For the dress vowel, variability seems to increase slightly after session six, to decrease after session seven, and to increase again after session eight. Overall, but especially after session five, the amount of variability of the trap vowel is higher than of the dress vowel. There is a slight increase in the average F2 vowel formants of the trap vowel (from 2250 Hz to 2500 Hz), but this does not happen for the dress vowel, which stays around 2500 Hz (see appendix II, figure 2).
Figure 61 shows the average F2 values of the trap and dress vowels per session. The trendlines of both the trap and the dress vowels show a slight increase in F2 values. On the whole, the F2 values of the dress vowel are higher than those of the trap vowel.

![Figure 61: Vowels in context min-max averages per session](image)

**Figure 62 Min-max graph of the average F2 vowel formants in Hz per session**

In the min-max graph that figure 62 depicts, the variability of the trap vowel seems to increase after the first session, to decrease after session six, and to increase again after session eight. For the dress vowel, variability seems to increase after session three and to decrease again after session four. From session six onwards, there seems to be hardly any variability.

**Free speech**

The development of Izaline’s vowel production in free speech is shown in the figures below.

![Figure 63: Vowels in free speech min-max averages per session](image)

**Figure 63 Min-max graph of the average F2 vowel formants in Hz per session**

Figure 63 depicts a min-max graph of the average F2 values for both the trap and dress vowel per session. As there is no data for the trap vowel in session five, the graph is incomplete. The
values of the trap vowel seem to start with relatively a lot of variability, but in session six, there is hardly any variability visible. After session seven, variability increases again, but decreases after session nine. For the dress vowel, variability seems to decrease after session three, and to increase after session six. This vowel begins with relatively much variability as well. Overall, the F2 values for the trap and dress vowel are very close together.

**Question 2: Does perception precede production?**

In the graphs below the development of the perception of vowels is depicted.

Figure 64 Number of right answers per session in the vowels perception task

Figure 64 shows the number of right answers per session. The figure shows that there is an increase in the number of right answers.

In figure 65, the number of wrong answers per session for the trap and dress vowel are showed. There are only two sessions in which more trap vowels are answered wrong against
ten sessions in which more dress vowels are answered wrong. When a dress vowel was perceived incorrectly, it was rather perceived as a trap vowel, and vice versa.

**Bram: early learner**

**Question 1: Does production in isolation precede production in context and free speech?**

**Isolation**

The development of Bram’s vowel production in isolation is shown in the figures below.

![Vowels in isolation averages per session](image)

Figure 66 Line graphs of the average F2 vowel formants in Hz per session

Figure 66 shows the average F2 values of the trap and dress vowels per session. There is quite some variability between the average F2 values over the sessions, though the trendlines show some development. The average values of the trap vowel seem to decrease whereas those of the dress vowel seem in to increase. In the beginning, the average values of the trap vowel were higher, but in the end this holds for the dress values. Both vowels – especially the trap vowel – seem to develop more towards the native speaker distinctions.
In figure 67, the min-max graphs of the average F2 values of the trap and dress vowels per session are depicted. The variability of the trap vowel slightly decreases until session six, after which it increases, but decreases again after session seven. The variability of the dress vowel decreases at first, but increases after session four. After session six, there is again a slight decrease of variability. Overall, the F2 values of the trap and dress vowel are quite close together.

**Context**

The development of Bram’s vowel production in context is shown in the figure below.

Figure 68 shows the development of the amount of variability over time for the trap and dress vowels. The amount of variability for the trap vowel is quite constant; the biggest increase already occurs after session one. For the dress vowel, the amount of variability decreases up to session four, after which it increases. After session six, there is a decrease of variability, but
after session eight there is a very slight increase again. The F2 values of both vowels are again quite close to each other, though overall, there is variability going on (see appendix II, figure 10).

**Free speech**

The development of Bram’s vowel production in free speech is shown in the three graphs below.

![Vowels in free speech averages per session](image)

**Figure 69** Line graphs of the average F2 vowel formants in Hz per session

Figure 69 illustrates that over time, there has been a slight increase in F2 values for both the trap and the dress vowel. The average F2 values for both vowels are very much alike.

![Vowels in free speech min-max averages per session](image)

**Figure 70** Min-max graphs of the average F2 vowel formants in Hz per session

Figure 70 shows that from session four onwards, the maximum values for the trap and dress vowels are very similar, there is almost no difference between the vowels. For the trap vowel, there is an increase of variability after session two and a decrease again after session seven.
The dress vowel begins with relatively much variability. This decreases after session three, increases after session five, and decreases again after session 7.

**Question 2: Does perception precede production?**

In the graphs below the development of the perception of vowels is depicted.

![Figure 71 Number of right answers per session in the vowels perception task](image)

Figure 71 shows the number of right answers per session in the vowels perception task. The figure shows that the number of right answers increases over time.

![Figure 72 Number of wrong answers for the trap and dress vowel per session](image)

In figure 72, the number of wrong answers per session for the trap and dress vowel are showed. There are two sessions in which there were more wrong answers for the trap vowel; there are five sessions in which there were more wrong answers for the dress vowel. In two other sessions, there were the same amount of wrong answers for the trap and dress vowel.
Cora: Late learner

Question 1: Does production in isolation precede production in context and free speech?

Isolation

The development of Cora’s vowel production in isolation is shown in the figures below.

Figure 73 Dispersion of the F2 vowel formants in Hz in isolation

In figure 73, the dispersion of the F2 values in isolation for the trap and dress vowel for each word in each session are showed. There is neither much variation within the session, nor much variation over time. This is the case for both the trap and the dress vowel. The values are all very close together.

Figure 74 Min-max graphs of the average F2 vowel formants in Hz per session
In figure 74 is showed that there is almost no variation in the sessions, both for the trap and the dress vowel. The values of the dress values are only a bit higher than those of the trap vowel.

**Context**

The development of Cora’s vowel production in context is shown in the figures below.

![Figure 75 Dispersion of the F2 vowel formants in Hz in context](image)

Figure 75 shows that for words in context, there is again little variability within the sessions as well as over time. Only in session six and twelve, there is some variation.

![Figure 76 Min-max graphs of the average F2 vowel formants in Hz per session](image)

Overall, figure 76 shows little variability in the sessions for the trap and dress vowels. For the trap vowel, variability only seems to increase clearly after session ten. For the dress vowel, variability increases a bit after session four and decreases again after session seven.
Free speech

The development of Cora’s vowel production in free speech is shown in the figures below.

Figure 77 Dispersion of the F2 vowel formants in Hz in context

Figure 77 shows that again for free speech, there is little variability for the trap and dress vowels in each session. In sessions five and ten, there is a bit more variability for the dress vowel.

Figure 78 Min-max graphs of the average F2 vowel formants in Hz per session

Figure 78 also shows that there is hardly any variability for both vowels within the sessions.

Question 2: Does perception precede production?

In the graphs below the development of the perception of vowels is depicted.
In figure 79, there seems to be a slight increase in the number of right answers per session in the vowels perception task, though the number of right answers in the first and in the last sessions are similar.

Figure 80 shows the number of wrong answers for the trap and dress vowel in each session. In eleven of the fourteen sessions, the trap vowel has been perceived wrong more times than the dress vowel. The dress vowel has been perceived correctly more often in only one session.

**Aaffien: Late learner**

**Question 1: Does production in isolation precede production in context and free speech?**

**Isolation**
The development of Aaffien’s vowel production in isolation is shown in the three graphs below.

In figure 81, the dispersion of the trap and dress F2 values for each word in each session are depicted. Up till session nine, there is hardly any variability. From session ten onwards, there is more variability within the sessions, both for the trap and the dress vowel.

Figure 82 shows that the minimum and maximum F2 values of the dress vowel are both higher than those of the trap vowel. Overall, there is not much variability within the sessions. Only after session eight, there is an increase in variability for the trap vowel, but this decreases again after session eleven.
Context

The development of Aaffien’s vowel production in context is shown in the three graphs below.

**Figure 83** Dispersion of the F2 vowel formants in Hz in context

Figure 83 shows that there is not much variability in F2 values for the trap and dress vowel for words in context. In session twelve, there is an increase in variability for the dress vowel. In session thirteen there is more variability for both vowels.

**Figure 84** Min-max graphs of the average F2 vowel formants in Hz per session

In figure 84, the min-max graphs of the average F2 values for the trap and dress vowels for words in context are illustrated. Up till session nine, the minimum and maximum values of the dress vowel are higher than those of the trap vowel. After session nice, variability increases for the dress vowel. Overall, there is little variability.
Free speech
The development of Aaffien’s vowel production in free speech is shown in the three graphs below.

Figure 85 Dispersion of the F2 vowel formants in Hz in free speech

Figure 85 shows that for free speech as well, there is little variability going on. In sessions ten, eleven and twelve, there is an increase in variability for the trap vowel.

Figure 86 Min-max graphs of the average F2 vowel formants in Hz per session

Figure 86 shows that after session nine, there is an increase in variability for the trap vowel only.

Question 2: Does perception precede production?

In the graphs below the development of the perception of vowels is depicted.
Figure 87 Number of right answers per session in the vowels perception task

Figure 87 shows that there has been some increase in the number of correct answers per session in the vowels perception task.

Figure 88 Number of wrong answers for the trap and dress vowel per session

Figure 88 shows that there are three sessions in which the trap vowel was perceived wrong more often and four sessions in which the dress vowel was perceived wrong more often. There are five sessions in which the trap and dress vowel are perceived wrong the same amount of times.

Combined results

**Question 3: Is early better than late?**
In the graphs below the four participants were compared to each other, to see if the early learners have an advantage over the late learners or vice versa.

We have only included the combined results for the vowels perception task here, because the combined graphs for both vowels in production as well as all four participants would ultimately be too unclear.

![Percentage of right answers per session](image)

**Figure 89 Percentage of right answers per session in the vowels perception task for all participants**

In figure 89, the instances where the meaning was guessed wrong are excluded from the 100%. In this way, fair percentages of the number of right and wrong answers are created. Overall, the lines in figure 89 all seem to be increasing. Whereas the graphs of the late learners are higher than those of the early learners in the beginning, the early learners seem to catch up with the late learners in the end.
Chapter 5 Discussion

In this chapter we will discuss our results. We will first discuss our three sub questions in relation to our VOT results, and then in relation to our trap/dress vowels results.

VOT

_**Question 1: Does production in isolation precede production in context and free speech?**_

The results from the four participants do not show a unified picture. Izaline’s results show, unlike we predicted, most variation and the highest values occurs in context, not in isolation. The variation is a signal that her system is changing. Striking in Izaline’s context results were the three major outliers in context, which exceeded the 250 ms. The outliers could mean that she has picked up on the existence of VOT and that she is trying out different things. It could also mean she is consciously thinking about her pronunciation. Least variation and lowest values were found in free speech, which is in line with our expectations. It seems that context precedes isolation and free speech.

Bram’s results are similar to Izaline’s results. Like Izaline, he shows the least variation and the lowest values in free speech. The values in isolation and context cross over and isolation seems to show a delayed effect of the same pattern as context. Again, context seems to precede isolation.

Cora and Aaffien show a different picture. Their values in all three conditions lie close together and so does the amount of variation. For Cora, most variability and an increase in VOT is shown in isolation. It does seem that production in isolation precedes production in context and free speech. Yet, for Aaffien, even though free speech shows the least variation, isolation does not seem to really precede context or free speech.

Overall, free speech does seem to be the condition that shows the least variation, which is what we expected. However, isolation does not always seem to precede context, as for the early learners it is the other way around. All learners apart from Aaffien seemed to make progress. The variation in their data signals that their language systems are changing. Bram and Cora even showed many values above the 35 ms in the dispersion graphs. For Bram this was only the case in context and isolation, again showing that free speech is the condition that comes last. In free speech, the learner’s attention first of all goes to naming the picture,
knowing the right answer is the priority. In isolation and context the learners had to repeat after a native speaker, which automatically puts the focus more on pronunciation.

**Question 2: Does perception precede production?**

In the perception results, all participants showed variability in the number of right answers, to a certain extend. If we look at the different combinations separately, all learners show some variability. Izaline and Aaffien show variability in all three conditions, and Bram and Cora only in the English contrast aspirated / short. Bram perceives the other combinations wrong and Cora perceives them correctly.

If we compare this to the production results we can argue that both in perception and production something seems to be happening as there is a lot of variation. In production the (increased) variability signals that the learners are developing their VOT in some conditions, but if we look at the perception results we see that the aspirated / short contrast has not been acquired yet. Hence, our results do not support the assumption that perception precedes production.

**Question 3: Is it better to start early than it is to start late?**

From our results we cannot conclude that it is better to start early than it is to start late, as results are not univocal. If it is better to start early than it is to start late, we would expect to find most variability and progress in the results of the early learners and this is not always the case. In isolation, Bram and Aafien seem to have the highest values overall, but the late learners show the most variation. In context, Bram and Aafien seem to have the highest values overall, but the early learners show the most variation. In free speech Aaffien has the highest VOT values and Izaline the lowest. Aaffien and Izaline show most variation. The results do seem to show that late learners are better at pronouncing words in isolation, whereas the early learners seem to provide the context condition. One possible explanation for this is that the early learners’ knowledge of English was so little, that they often only understood the final word of the sentence, thus the target word, in the repetition task. Consequently, they would mumble the first part of the sentence and pronounce the final, familiar word very clearly. Their focus would thus be on that final word. For the late learners, this was not the case as they did understand the entire sentence.
Vowels

*Question 1: Does production in isolation precede production in context and free speech?*

One thing that is remarkable when it comes to the production of the trap and dress vowel for all four participants, is that while both early learners show variability in all three conditions, both late learners hardly show any variability. On the whole, both late learners produce trap and dress vowels with F2 values that are higher than those of our native speaker. Also, they generally do not make a distinction between the trap and the dress vowel. Both of them do begin to show a bit of variability in the last sessions. The varying data points of the early learners of both vowels is sometimes at F2 values similar to those of our native speaker.

The dispersion graphs in isolation, context, and free speech of the four participants are univocal, as both early learners show an equal amount of variability, while the late learners hardly show any variability.

The graphs with the averages per sessions for isolation, context and free speech do not show univocal results. Overall, the trap and dress values of Izaline are higher than those of Bram. Also, Izaline’s F2 values for trap and dress slightly increase, while those of Bram slightly decrease. The results of the late learners are quite similar; both show a slight increase in average F2 values for both vowels, starting around 2000 Hz. In the end, the late learners still seem to produce the trap and dress vowels with a Dutch ‘e’, not making a distinction between the two vowels.

The thing that is most notable when it comes to the min-max graphs of isolation, context and free speech of all four participants, is that the early learners do show variability, whereas the late learners scarcely do. Towards the end, the late learners do begin to show some variability, though still very little.

The variability that the early learners already show from the beginning may indicate a change in their language systems. On the other hand, it may also indicate that they do not have one manner of producing the vowels, which leads to the variation in the data points. It is clear that the late learners have not picked up on the difference between the trap and the dress vowel yet. They produce the vowel as if it is one single vowel, of which the F2 values point to a Dutch ‘e’.
**Question 2: Does perception precede production?**

The individual perception graphs of our four participants illustrate that while the late learners started with a higher percentage of right answers, the early learners catch up with them in the end. Both early learners have a bigger increase than the late learners. This may indicate that the early learners are faster in developing a new phonetic category for the perception of the English trap and dress vowels. On the other hand, the early learners started lower at first, so they were able to have a bigger increase in perception than the late learners.

Especially the late learners seem to have picked up the difference between the trap and dress vowels when it comes to perception, as they generally make less mistakes than the early learners. However, Cora clearly seems to make more mistakes in perceiving the trap vowel correctly than the dress vowel on the whole. This means that in many cases, she does not perceive the trap vowel as being different from the dress vowel. Nevertheless, this number does decrease over the sessions. The opposite is the case with the early learners; they have perceived the dress vowel incorrectly more often during the sessions than the trap vowel.

It is rather difficult to compare the results of the vowels perception task with the produced data of the participants, since the figures are quite different from each other. Though, for the early learners we could state that perception comes before production, since there is a positive development in their perception, while there is hardly any development yet in their production. In their production, both late learners do not distinguish between the trap and the dress vowel, but rather keep producing them as if it were a Dutch ‘e’. Their perception is already at a relatively high level in the beginning. Overall, there is more variability when it comes to the early learners, but their perception does improve during the second half of the sessions. In their production, there is quite some variability going on, though some of these data points do approximately come close to the correct F2 formant values.

**Question 3: Is it better to start early than to start late?**

Combining the two early learners with the two late learners when it comes to the acquisition of the trap and dress vowel, we might argue that the early learners have more advantages over the late learners. When it comes to production, the late learners are a bit stuck in pronouncing the English trap and dress vowel as if it were a Dutch ‘e’. Furthermore, they do not make a distinction between the two English vowels in producing them. The early learners do not perform much better yet, though their amount of variability may point to a change in their
language system. They may be trying different things, which could ultimately lead to an attractor state during which they produce the vowels correctly. On the other hand, especially the early learner Bram does not yet make a clear distinction between the two vowels as well, which may illustrate an example of equivalence classification. When it comes to perception, the late learners make less mistakes overall. Nevertheless, the early learners catch up with them at quite a high pace.

Because of the rather short period of time that the participants were tested, we cannot yet convincingly state that it is better to start early than it is to start late, though the early learners’ results are more promising than the results of the late learners.
Chapter 6 Conclusion

In this study, we aimed at gaining more insight into the development of acquiring a foreign language phonology. We were specifically interested in three questions: first of all, is it better to start learning a foreign language phonology early than it is to start late? Secondly, does perception precede production? And lastly, does production in isolation precede production in context and free speech? In order to answer these questions we did four longitudinal case studies, analysing the development of two Dutch children and two Dutch adults over an eight-week period. To answer these questions, we adopted a DST approach, looking at the process of development, rather than looking at the end product.

The first research question, concerning age, is one of the core issues in discussions about second language acquisition. The general thought is that it is easier to acquire a second language at an early age, than it is at a later age. In addition, this is believed to be especially true for an L2 phonology. There are two schools of thought on why this is so. The first school of thought applied CPH to L2A, putting it down to a biological change after maturation. The second school of thought believes it is caused by the L1 influences the acquisition of the L2. In our second chapter we rejected the CPH-L2A and looked at the issue from a DST point of view. From a DST point of view we should speak of age-related issues. These issues consist of age-related factors causing age-related effects. Age is not longer seen as a simple concept, but as a complex concept: different factors interrelated and confounded with age are interacting with each other and never stop changing. The influence of the L1, which the second school of thought pinpoints as the cause of the age-related effects, is just one age-related factor.

This links to the second question which assumes a precursor relation between perception and production of non-native contrasts and sounds. After having discussed various models of perception in chapter 2, we again adopted a DST approach. We speculated that the native categories in our phonological system act as attractors when we perceive non-native sounds. These categories get more fixed as we get older which makes them an age-related factor. These categories effect perception, which then becomes an age-related effect. Perception in its turn becomes an age-related factor as it effects the L2 pronunciation. In chapter one we also stated that an attractor state is not obligatory, and that with the right energy the system can always get out of an attractor state. This is connected to limitation of resources, (if resources are limited there is not enough energy to get out of the attractor), which links to the third question.
The third question assumes that it is easier to adopt the right pronunciation for words in isolation than it is for words in context or free speech. (With the conditions isolation and context we mean repetition tasks.) The idea behind this is based on the idea of limitation of resources. In isolation, the learner’s attention is on the pronunciation of one single word. As it is a repetition task, all energy can go into pronouncing the word. In context, an entire sentence needs to be repeated, which lessens the attention that is available for the target word. Finally, in free speech – a picture naming task in our experiment – attention is paid to getting the answer right semantically leaving less energy for pronunciation.

The results of the study showed that the answers to our three research questions do not necessarily support the assumptions behind them. Over the eight-week period, the participants showed quite some variation and results are not always univocal. This was, for instance, the case with the VOT results in relation to the first research question. From a DST perspective, variability is seen as a precursor of change: a signal that the language system is changing. Therefore, we did not only look for an increase in VOT values, but we were also interested in the amount of variation. In Izaline’s case, most variability and the highest values were found in context. The values in isolation came second. Free speech, as expected, showed the least variation and the lowest values. Bram’s results also suggested that context precedes isolation and free speech. Cora’s results suggested that isolation precedes context, but there was little difference between context and free speech. Finally, Aaffien’s results showed yet another picture. Even though there was quite some variation which suggests a change in her language system, there was little difference in variation and values between the three conditions. Overall, isolation does not always seem to precede context, as it was the other way around for the early learners. Free speech does seem to come last in line.

The VOT results in relation to the second and third question do not lead to a decisive answer either. Even though the variability in Izaline’s, Cora’s, and Aaffien’s results suggests a change in their language systems, the did not provide us with convincing evidence for a precursor relation between perception and production. There is no clear indication that it is better to start early than it is to start late either.

The vowel results do not provide a clear answer to the first question either. Bram does show a positive development in isolation towards the right F2s, and Izaline does not make a clear difference between the two vowels, except in context where she does show some difference between trap and dress. Cora and Aaffien hardly show any variation, though Aaffien shows little variation in isolation later on. Aaffien therefore suggests isolation does come before context and free speech, but Bram’s and Izaline’s results do not entirely.
Whereas Bram’s variation does not suggest isolation comes before context and free speech, his positive development in isolation does. Izaline’s overall variation, on the other hand, does support the claim, but her development in context does not.

The vowel results on perception are univocal. All participants show an increase in the number of right answers, which suggests that their language systems are changing and new categories are being developed. As there is a clear increase in perception but not so much in production (although the amount of variability in the early learners does suggest change), it seems that perception indeed precedes production in this case.

The vowel results on the third research question do seem to suggest that it is indeed better to start early than it is to start late when it comes to production. The early learners show a lot more variability than the late learners, who hardly show any variability at all. In addition, whereas Izaline and Bram both show promising results in one of the conditions, the late learners use the two vowels interchangeably in all conditions, pronouncing them close to the Dutch 'e'.

In conclusion, we can state that we did not find any decisive answers to our research questions. The inconsistent results to our research questions are not surprising from a DST perspective. DST starts from the idea that language is a dynamic system which is always changing. More importantly, it is non-linear, which means that a small change at one point in time might have great effects later on. The variation we found in our data suggests that the language systems of our participants are definitely changing; however, a clear positive development is not always present yet, leading to inconsistent results. These results also seem to point to the fact that there are more factors that play a role than we could have included in our study. Since in DST all resources, subsystems, and also their interaction changes over time, it is sometimes difficult to account for the results found.

The important thing we learned from this study is that the development of the acquisition of a foreign language development is not as straightforward as some people may think. Looking at the process teaches us that variability gives us valuable information about their development. It tells us that values in isolation do not always need to be better than the ones in context or free speech, even if isolation proves to precede context and free speech in the long run (which we still do not know for sure). In addition, the results suggest that correct perception is not always required for production. Finally, it tells us that the age-issue is not as simple and clear-cut as some people like to think. Our study showed us that late learners can learn a foreign language phonology, because this is what their variability in VOT values suggests. Whether early learners are indeed better at this than late learners is something future
process-oriented research can tell us. But who knows, maybe the tree bends just as well as the twig does, just in a different way.

**Suggestions for further research**

For future research it would be interesting to do a similar experiment over a longer period of time and less often than two times a week. The eight-week period was fairly short, and a longer experiment would provide even more insight into the process of development. The study could also be extended by following not only the development of the VOT and trap/dress vowel distinction, but also of final voicing, another difficult aspect for Dutch learners of English. Another interesting element to add to future research would be the development of their native language phonology: is the perception and production of their native language influenced by the acquisition of a foreign language phonology, and, if so, how does this develop? Also, it might be interesting to add one early and one late learner who receive explicit instructions, and compare their development to the development of the other participants. Another final suggestion concerns the amount of exposure to the target language. The participants in our study were minimally exposed to English, though for phonological acquisition to really take off, you will need a lot more exposure to the language, close to immersion. This is one thing that we learned from our study in relation to early foreign language education.
Appendices

I Remaining figures – VOT

Izaline: early learner

**Question 1: Does production in isolation precede production in context and free speech?**

*Figure 1 Line graph of the average VOT values per session*

*Figure 2 Min-max graph of the average VOT values per session*
Comparison isolation, context, and free speech

Figure 3 Line graph of the average VOT values per session

Figure 4 Min-max graph of the average VOT values per session

Figure 14 Line graph of the average VOT values per session
Figure 6 Min-max graph of the average VOT values per session in isolation and context

Figure 7 Min-max graph of the average VOT values per session in context and free speech

Figure 8 Min-max graph of the average VOT values per session in free speech and isolation

Bram: early learner

Question 1: Does production in isolation precede production in context and free speech?
Comparison isolation, context, and free speech
Figure 12 Line graph of the average VOT values per session

Figure 13 Min-max graph of the average VOT values per session in context and free speech

Figure 14 Min-max graph of the average VOT values per session in free speech and isolation
Figure 15 Min-max graph of the average VOT values per session in isolation and context

**Question 2: Does perception precede production?**

![VOT perception scores prev/asp](image)

Figure 16 Scores of the prevoiced / aspirated combination per session

![VOT perception scores per combination](image)

Figure 17 Scores of the short / prevoiced combination per session

Cora: Late learner

**Question 1: Does production in isolation precede production in context and free speech?**
Figure 18 Dispersion of the VOT values in isolation

Figure 19 Min-max graph of the average VOT values per session

Figure 20 Dispersion of the VOT values in context
Comparison isolation, context, and free speech

Aaffien: Late learner
Question 1: Does production in isolation precede production in context and free speech?

Figure 24 Min-max graph of the average VOT values per session

Figure 25 Line graph of the average VOT values per session

Figure 26 Line graph of the average VOT values per session

Comparison isolation, context, and free speech
Figure 27 Dispersion of the VOT values

Figure 28 Line graph of the average VOT values per session

Figure 29 Line graph of the average VOT values per session with trendlines
Figure 30 Min-max graph of the average VOT values per session in isolation and context

Figure 31 Min-max graph of the average VOT values per session in context and free speech

Figure 32 Min-max graph of the average VOT values per session in free speech and isolation
II Remaining figures – vowels

Izaline: early learner

Question 1: Does production in isolation precede production in context and free speech?

Figure 1 Dispersion of the F2 vowel formants in Hz in isolation

Figure 2 Line graph of the average F2 vowel formants in Hz per session
Figure 3 Dispersion of the F2 vowel formants in Hz in context

Figure 4 Dispersion of the F2 vowel formants in Hz in free speech

Figure 5 Line graph of the average F2 vowel formants in Hz per session

**Question 2: Does perception precede production?**

Figure 6 Number of wrong answers in the vowels perception task per session
Bram: early learner

**Question 1: Does production in isolation precede production in context and free speech?**

Figure 7 Number of right and wrong answers and number of meaning answered wrong

Figure 8 Dispersion of the F2 vowel formants in Hz in isolation

Figure 9 Line graphs of the average F2 vowel formants in Hz per session
Figure 10 Dispersion of the F2 vowel formants in Hz in context

Figure 11 Line graphs of the average F2 vowel formants in Hz per session

Figure 12 Dispersion of the F2 vowel formants in Hz in free speech

**Question 2: Does perception precede production?**
Perception number of wrong answers per session

Figure 13 Number of wrong answers in the vowels perception task per session

Perception vowels

Figure 14 Number of right and wrong answers and number of meaning answered wrong

Cora: Late learner

Question 1: Does production in isolation precede production in context and free speech?

Vowels in isolation averages per session

Figure 15 Line graphs of the average F2 vowel formants in Hz per session
Figure 16 Line graphs of the average F2 vowel formants in Hz per session

Figure 17 Line graphs of the average F2 vowel formants in Hz per session

**Question 2: Does perception precede production?**

Figure 18 Number of wrong answers in the vowels perception task per session
Aaffien: Late learner

*Question 1: Does production in isolation precede production in context and free speech?*
Question 2: Does perception precede production?

Figure 22 Line graphs of the average F2 vowel formants in Hz per session

Figure 23 Number of wrong answers in the vowels perception task per session

Figure 24 Number of right and wrong answers
III Vowels perception task – pictures
IV Vowel perception task – recording sets

Perception test trap/dress vowel distinction

Set 1:
man
key
bread
pen
chair
kettle
men
Brad
Boy
band
cattle
pan
tea
vat
toy
bend
men
hair
vet

Set 2:
pan
House
Men
cattle
Mouse
band
Boy
bend
pan
Toy
vet
kettle
bread
Key
man
Tea
Brad
vat
pan

Set 3:
Band
House
Vat
bend
Cattle
pan
Mouse
Key
vet
Pen
Tea
man
Kettle
Car
Kettle
Brad
Star
bread
men

**Set 4:**
Vat
Brad
Hair
Bend
Chair
kettle
men
Coat
Bread
Goat
Band
Vet
cattle
Car
pan
Bread
Star
Man
pen
V Score forms VOT perception AX-task

Scoreformulier Perceptie VOT

Naam:
Datum:

<table>
<thead>
<tr>
<th>Deel</th>
<th>1</th>
<th>Versie</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>oefen1</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>oefen2</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
</tbody>
</table>

Grijs = goede antwoord
**Scoreformulier Perceptie VOT**

**Naam:**

**Datum:**

<table>
<thead>
<tr>
<th>Deel</th>
<th>2</th>
<th>Versie</th>
<th>A</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>hetzelfde</strong></th>
<th><strong>verschillend</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>oefen1</td>
<td><strong>hetzelfde</strong></td>
</tr>
<tr>
<td>oefen2</td>
<td><strong>hetzelfde</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4</th>
<th><strong>hetzelfde</strong></th>
<th><strong>verschillend</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>6</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>7</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>8</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>9</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>10</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>11</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>12</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>13</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>14</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>15</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>16</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
<tr>
<td>17</td>
<td><strong>hetzelfde</strong></td>
<td><strong>verschillend</strong></td>
</tr>
</tbody>
</table>
## Scoreformulier Percepie VOT

**Naam:**

**Datum:**

<table>
<thead>
<tr>
<th>Deel</th>
<th>Versie</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>hetzelfde</th>
<th>verschillend</th>
</tr>
</thead>
<tbody>
<tr>
<td>oefen1</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>oefen2</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>hetzelfde</th>
<th>verschillend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>
Scoreformulier Perceptie VOT

Naam: 
Datum: 

<table>
<thead>
<tr>
<th>Deel</th>
<th>Versie</th>
<th>B</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>hetzelfde</th>
<th>verschillend</th>
</tr>
</thead>
<tbody>
<tr>
<td>oefen1</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>oefen2</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>4</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>5</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>6</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>7</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>8</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>9</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>10</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>11</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>12</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>13</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>14</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>15</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>16</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
<tr>
<td>17</td>
<td>hetzelfde</td>
<td>verschillend</td>
</tr>
</tbody>
</table>
VI Score forms vowels perception task

Scoreformulier Perceptie dress/trap

Naam:  
Datum:  

Set 1.

<table>
<thead>
<tr>
<th>Woord</th>
<th>Goed</th>
<th>Fout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>Bread</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chair</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>Kettle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boy</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>Band</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cattle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tea</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>Vat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toy</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>Bend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hair</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>vet</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Scoreformulier Perceptie dress/trap**

**Naam:**

**Datum:**

**Set 2.**

<table>
<thead>
<tr>
<th>Woord</th>
<th>Goed</th>
<th>Fout</th>
</tr>
</thead>
<tbody>
<tr>
<td>pan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>House (f)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cattle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse (f)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boy (f)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toy (f)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kettle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bread</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key (f)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tea (f)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pan</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Scoreformulier Perceptie dress/trap

Naam:  
Datum:  

Set 3.

<table>
<thead>
<tr>
<th>Woord</th>
<th>Goed</th>
<th>Fout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band</td>
<td></td>
<td></td>
</tr>
<tr>
<td>House</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>Vat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cattle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>Key</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>vet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tea</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>man</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kettle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>car</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>kettle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>star</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>bread</td>
<td></td>
<td></td>
</tr>
<tr>
<td>men</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Scoreformulier Perceptie dress/trap

Naam:  
Datum:  

Set 4.

<table>
<thead>
<tr>
<th>Woord</th>
<th>Goed</th>
<th>Fout</th>
</tr>
</thead>
<tbody>
<tr>
<td>vat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hair</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>bend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chair</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>kettle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>men</td>
<td></td>
<td></td>
</tr>
<tr>
<td>coat</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>bread</td>
<td></td>
<td></td>
</tr>
<tr>
<td>goat</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>band</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cattle</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>car</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>pan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bread</td>
<td></td>
<td></td>
</tr>
<tr>
<td>star</td>
<td>(f)</td>
<td></td>
</tr>
<tr>
<td>man</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pen</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
VII English shadowing tasks – recording sets

Set 1
pet
toy
pie
marry
path
men
boy
key
bath
hair
man
bad
bet
merry
tea
buy
chair
bed

She’s taking a bath.
I want some tea.
I like pie.
I like his hair.
The boy is merry.
I have a pet.
She’s taking a bath.
I need a chair.
We have a bet.
It’s his toy.
She will marry him.
He is a man.
This is my key.
I’m walking down a path.
He is bad.
This shirt I want to buy.
He’s a boy.
I see two men.

Set 2
dad
big
house
beach
sad
pig
beg
tea
pan
mouse
peach
toy
bag
It’s like you said.
This is my key.
Look at that pig.
Her dog is dead.
It’s his toy.
I don’t want to beg.
He’s eating a peach.
This is my house.
I don’t like this ban.
I want some tea.
He is my dad.
Let’s go to the beach.
She cooks with a pan.
It’s a nice bag.
I have a mouse.
He’s a boy.
She is sad.
Our house is big.

Set 3
pin
car
pan
star
band
house
park
head
pee
had
bee
key
bin
bend
mouse
pen
tea
bark

It’s a blue pen.
I like the one I had.
This is my car.
He’s got a big head.
This is my house.
This is my key.
The dog says bark.
I need a pan.
I see a star.
Look at that bend.
I want some tea.
I have to pee.
I have a mouse.
I need a pin.
We are in the park.
She sings in a band.
Put it in the bin.
I saw a bee.

**Set 4**
land
mat
c coat
lend
pay
goat
pen
met
set
car
pack
hair
Ben
sat
bay
chair
back
star

Our house has a mat.
We have to go back.
I see a star.
I want a new pen.
Some people lend money.
I like this set.
I need a chair.
His name is Ben.
I like her pack.
He lives at the bay.
It’s a goat.
We just met.
This is my car.
I own some land.
I like his hair.
I still have to pay.
You need a coat.
I know where he sat.
VIII Dutch shadowing task – recording sets

**Set 1**
peer
knoop
pet
stroop
bed
fruit
beer
poot
slim
boot

De hond heeft een gebroken poot.
Ik eet veel fruit.
Dat is mijn beer.
Zij is slim.
Ik heb een rode pet.
Dat is een mooie boot.
Ik hou van stroop.
Ik wil een peer.
Dit is een knoop.
Ik wil nog niet naar bed.

**Set 2**

boot
slim
poot
beer
fruit
bed
stroop
pet
knoop
peer

Ik wil nog niet naar bed.
Dit is een knoop.
Ik wil een peer.
Ik hou van stroop.
Dat is een mooie boot.
Ik heb een rode pet.
Zij is slim.
Dat is mijn beer.
Ik eet veel fruit.
De hond heeft een gebroken poot.
IX Production pictures (for practice and free speech)

bark
Ben
set
pack
pay
sat
goat
back
bay
had
<table>
<thead>
<tr>
<th>boy</th>
<th>man</th>
<th>tea</th>
</tr>
</thead>
<tbody>
<tr>
<td>chair</td>
<td>pet</td>
<td>pie</td>
</tr>
<tr>
<td>bed</td>
<td>bath</td>
<td>men</td>
</tr>
<tr>
<td>merry</td>
<td>key</td>
<td>bad</td>
</tr>
</tbody>
</table>
beg     dead       said

mouse      pig     peach

bag                                                      ban    head

bend     pen     car
X Praat VOT script

# VOT script

# Geen spatie's in bestandsnamen
# Bestandsnamen niet met 'r' beginnen
# Eerst een keer zelf Write to binary file doen, daar map goed zetten!!!!

# Directory hier mogen alleen geluidsbestanden staan
dir$="C:\Documents and Settings\Hannah Rosing\Bureaublad\Master TTW\Scriptie\Testmap\Opnames\E2Izaline\Shadowing geknipt\EN\Zinnen\E2Shad1205 Set 3"

# Naam voor bestand met resultaten
results$=dir$+"results VOT E1Izaline.txt"
fileappend "results$" file\tab$beginning\tab$end\tab$VOT\newline$

# Invullen: 'dir$'[filename]
Create Strings as file list... fileList 'dir$'E2Shad1205pen*.*

select Strings fileList
nfiles=Get number of strings
for i from 1 to 'nfiles'
    select Strings fileList
    fileName$=Get string... i
    Read from file... 'dir$'fileName$' .wav'
    To TextGrid... "Mary bell" bell
    utt$=fileName$".wav"
    select Sound 'utt$'
    plus TextGrid 'utt$'
    Edit
    pause Select consonant voice onset time
    editor TextGrid 'utt$'
        aTime=Get start of selection
        bTime=Get end of selection
        cTime=Get selection length
    Add interval on tier 1
endeditor
    select TextGrid 'utt$'
    fileappend "results$" 'utt$'tab$"aTime:3"tab$"bTime:3"tab$"cTime:3"newline$
    select Sound 'utt$'
    plus TextGrid 'utt$'
#     output$ = "TG'utt$"
#     Write to binary file... 'output$'
    Remove
endfor
Reference List


Ref Type: Unpublished Work


*Language Learning, 30*, 271-287.
Reitsma, S. VOT AX-task. 2010.


Rivera, N. F. (1998). Effecto de la edad de inicio de aprendizaje de la lengua extranjera (ingks) y cantidad de exposicidn a la LE en la perception de las fonemas del inglis por hablantes de espariol y catalan.

Ref Type: Unpublished Work


